NUCLEAR MAGNETIC RESONANCE SPECTROSCOPIC CHARACTERIZATION OF PEPTIDE BASED NANOCARRIERS AND CARGO COMPLEXES

BANWO INUMIDUN DAMILOLA
Bachelor of Science, Lagos State University, Nigeria, 2012

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

MASTER OF SCIENCE

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

THESIS EXAMINATION COMMITTEE MEMBERS
NUCLEAR MAGNETIC RESONANCE SPECTROSCOPIC CHARACTERIZATION OF PEPTIDE BASED NANOCARRIERS AND CARGO COMPLEXES

BANWO INUMIDUN DAMILOLA

Date of Defence: December 15, 2017

Dr. François Eudes Co-Supervisor	Research Scientist	Ph.D.
Dr. Paul Hazendonk Co-Supervisor	Associate Professor	Ph.D.
Dr. Igor Kovalchuk Thesis Examination Committee Member	Professor	Ph.D.
Dr. Peter Dibble Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. Micheal Gerken Chair, Thesis Examination Committee	Professor	Ph.D.

DEDICATION

This thesis is dedicated to my son David Banwo for being so patient and understanding. To my parents Mr & Mrs Banwo and family members, for your prayers and motivations. To Bowlyn Daniel and friends, for your support, sacrifice, inspiration, and challenging me to think positive through out this academic journey. And above all to God who has granted me both wisdom, and knowledge to make it this far.

ABSTRACT

Cell penetrating peptides (CPPs) are short peptide 8-30 amino acids polypeptides that can deliver various molecules (nanoparticles to large fragment DNA) to subcellular locations such as the plastid organelles, cytoplasm, and nucleus etc. CPPs are characterized by containing clusters of cationic side chains that allow them to interact directly with the polar membrane surface, which enables then to enter the cell. One of the function of CPPs are to deliver cargo molecules such as DNA and RNA into cells. Most application have been developed for animal cells; however, the use of CPPs in plant cells remain a small researched field because of insufficient understanding of their mode of uptake. The present research describes firstly, the use of NMR technique to elucidate the structural properties of CPP-DNA complexation. Secondly the dynamic of the complexation. Finally, to relate the observations to their translocation mechanism.

For this study, short peptides such as; arginine (R), tri-arginine (R_3), nano-arginine (R_9), and Tat_2 (RKKRRQRRRRKKRRQRRR) were complexed with single-stranded and double-stranded DNA 5’ AGTCC 3’. The interaction between the single-stranded and double-stranded DNA 5’ AGTCC 3’ with the peptides was compared to understand how different DNA strands can influence complex formation. These complexes are used at relatively high concentrations; hence, measurements were at millimolar concentration. Therefore, for the study of DNA-CPP complexes, both non-isotopically enriched DNA and CPP samples were used. The reason for using such samples is to see if complexation can be observed using NMR spectroscopy, which will reduce the use of costly materials.

ACKNOWLEDGEMENTS

This work would have been impossible without the support of many individuals and collegues. I would like to express my gratitude to my supervisors Prof. Francois Eudes and Dr. Paul Hazendonk, without your support none of the work presented herein would have been possible. Special thanks to Laboratory technicians Dr. Fengying Jiang and also Kimberley Burton-Hughes, for their guidance and hands-on mentorship during sample preparations.

I would also like to thank my supervisory committee members, Dr. Peter Dibble, and Dr. Igor Kolvalchuk for the criticism and direction provided in my introductory chapter. Appreciation also goes to Dr. Nora Foroud for giving me the opportunity to work at her lab while I raise funds to support myself and my son.

I would like to thank the NMR manager and his assistant Tony Montina and Michael “Mike” Opyr. This two individuals are responsible for the knowledge and experience I developed operating the spectrometer. Special thanks to Tony Montina for pardoning me so many times when I break the intrumentation rules and for creating time to address experimental concerns.

I would like to thank my friend Daniel Bowlyn for proof reading my thesis even though you don’t have any idea of what I was doing, for your time, criticisms and encouragement. How could I forget my colleagues Dayo, Samira, Faroud, Mohammed, and Yesmin for making every Tuesday group meeting so fun and inspiring.

Lastly, I would like to thank my son David Shotonwa for his understanding, even when mum has to stay late at school and could not put you to bed every night. Thanks for been co-operative with your babysitters. To my parents Adebola Banwo and Toyin Banwo, for your prayers and encouragement. To my lovely brother Bamise Banwo and friend

Osakpolor Egbon. They happened to do most of the babysitting, thanks for your support and understanding. To my best friend Geraldine for your understanding and encouragement. And most of all to God for wisdom, knowledege and understanding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFM…… Atomic Force Microscopy
AMP……Antimicrobial Peptide
ARID…… AT-Rich Domain
CDS……. Circular Dichroism Spectroscopy
CHEOPS…… Chemical Shift *de* Nova Structure Derivation Protocol
COSY……... Correlation Spectroscopy
CPMG……... Carr-Purcell-Meiboom-Gill
CPPs …… Cell Penetrating Peptides
CSA……... Chemical Shift Anisotropy
DEPT……... Distortionless Enhancement by Polarization Transfer
DNA……. Deoxyribonucleic Acids
DQC…… Double Quantum Coherence
dsDNA……. Double-Stranded DNA
EPR…… Electron Paramagnetic Resonance
FID ……..……Free Induction Decay
FT…… Fourier Transform
GTP……... Guanosine Triphosphate
HIV……. Human Immunodeficiency Virus
HMQC……...……..... Heteronuclear Multiple Quantum Coherence
HOX…… Homeobox
HSQC…… Heteronuclear Single Quantum Coherence
INEPT……... Insensitive Nuclei Enhancement by Polarization Transfer
MAP…… Model Amphipathic Peptide
NLS……. Nuclear Localization Sequence
NMR……. Nuclear Magnetic Resonance
NOE ……...Nuclear Overhauser Effect
NOESY……... Nuclear Overhauser Effect Spectroscopy
PNA……. Peptide Nucleic Acids
ppm……. part per million
PRE…… Paramagnetic Relaxation Enhancement
RDC…… Residual Dipolar Coupling
REDOR……. Rotational-Echo, Double Resonance
RF or r.f.…… Radio Frequency
RNA……. Ribonucleic Acid
SEM…… Scanning Electron Microscopy
siRNA……. Small Interfering RNA
ssDNA……... Single Stranded DNA
TAR……..……...……...Torpedo Acetylcholine Receptor
TAT……. Trans-Activator of Transcription
WURST…… Wideband Uniform Rate, Smooth Truncation
ZQC……. Zero Quantum Coherence

CHAPTER 1

1. INTRODUCTION

1.1. Overview

Peptides, smaller than proteins, can still exhibit complex behaviour owing to their dynamics. Cell penetrating peptides (CPPs) are known for containing clusters of cationic residues that interact directly with the polar membrane surface and are thus essential to translocate into the cell.[1] Various biotechnologies exploit this property to transfer cargo molecules, such as DNA and RNA into cells.

The plasma membrane of a living cell has a unique ability to allow certain ions and molecules to pass through it and protects the cell membrane from its surroundings. The inability of bioactive molecules to enter the membrane has been a major drawback in drug targeting, therapeutics, and diagnostic areas.[2]

The plasma membrane includes proteins that can perform specific membrane functions.[3] These proteins, some of which are channel proteins allow the passage of a substance across the membrane, and carriers protein that binds to the substance and helps it transfer to the membrane. Substances such as water can pass through the plasma membrane easily without a carrier protein, and others, such as ions, macromolecules, and charged molecules need a carrier to cross the membrane. Macromolecules such as polysaccharides, polynucleotides, and polypeptides, that are too large to be transported by carrier proteins, are transported into cells by vesicles. These molecules are usually transported by endocytosis, the process in which the molecules are captured from the surroundings of the cell by engulfing it with the cell membrane and exocytosis in which the vesicles formed by Golgi apparatus binds

to the plasma membrane and secrete their waste to the exterior of the cell membrane. The ability of these proteins to transport molecules, bind, and interact with the plasma membrane make it possible for peptides to form complexes with oligomers such as DNA and RNA.

The delivery of biomolecules such as DNA, RNA, and proteins to plant cells has been achieved using various agricultural biotechnology.[1] These techniques use plant genetic engineering methods such as particle bombardment,[4] *Agrobacterium*,[5] microinjection[6] and recently cell penetrating peptides[5] to deliver genetic material into plant and animal cells. Some of these techniques were shown to be more effective in mammalian cells as compared to plant cells. Possible limitations could be the structural barrier in the form of the plant cell wall. More so, the presence of the large central vacuole that does exist in the plant cells and therefore makes it harder for genetic material to pass through the plant cells.[7]

The two most common plant transformation methods known are *Agrobacterium* mediated transformation and particle bombardment. The *Agrobacterium* mediated transformation, as the name imply is a transformation method that uses *Agrobacterium tumefaciens*.[8] The ability of *Agrobacterium* to transfer biomolecules such as DNA to plant cells has been used extensively for plant genetic engineering. *A.tumefaciens* is a bacterium that causes crown gall disease in plant cells, which can result in the growth of tumors in infected plants. Studies have shown that a plasmid referred to as tumor-inducing (Ti) plasmid allows these bacteria to target a broad range of plant hosts.[9]

Although *A. tumefaciens* has been shown to infect many plant hosts, the process of infection and the delivery of Transferred-DNA (or T-DNA) is determined by the genetic control factors from both plant host and the bacterium.[10] For instance, *Agrobacteria* have been shown to genetically transform monocot plant species such as rice, wheat, and maize.

A.tumefaciens has also been effectively used for nuclear transformation.[11]The VirD2 protein attached to the 5'-end of the T-strand is presumed to aid the transformation of *A.tumefaciens* into the nucleus.[12] The VirD2 contains two NLS located at the amino and carboxylic-terminal ends, that helps in the infection of agrobacterium. It was reported that the ability of VirE2 to direct fluorescently labelled ssDNA into plant cell nucleus was achieved using a microinjection of In vitro formed T-complexes.[9] The result shows the use of microinjection of florescence ssDNA and dsDNA alone results in the cytoplasmic florescence with no possibility of transfecting the nucleus. However, using microinjection of In vitro-formed VirE2-ssDNA complexes, results in the efficient nuclear transformation. Although the efficiency of transformation is not at par when compared to transformation of dicots.[13] This limitation led to the development of other transformation method such as the particle bombardment.

Particle bombardment, also known as the biolistic method is one of the most popular gene transfer techniques for monocots. In this biotechnology transformation method, the use of tungsten microprojectile and plasmid DNA are used to transform *Chlamydomonas reinhardtii, Saccharomyces cerevisiae,* and *Candida glabrata* mitochondria.[14] The use of particle bombardment involves the extensive coating of heavy metals such as tungsten or gold particle with plasmid DNA, which is fired at high velocities into the target cells.[15] This method has been used to attain a stable plasmid transformation in dicot plant species including potato, tomatoes, tobacco, soybean, lettuce, cabbage and sugar beet. Particle bombardment has also been used for the delivery of DNA into the nucleus. A study has reported the delivery of plasmid DNA into cultured tobacco cells using a pneumatic particle gun.[16] The result shows that 90% of the gene expression cell after bombarding the cell with

the gold particle received a DNA-coated particle in the nucleus, which results in the highest gene expression efficiency of the DNA.

To overcome the limitation of host range, difficult target of T-DNA transfer into the cell, and the inability to transform different types of monocots in both *Agrobacterium,* and particle bombardment gene transfer technique, a new peptide transfection technology referred to as cell penetrating peptide, that was formally developed to work for animal cells, have been considered for plant cells.

1.2. Cell Penetrating Peptides: Definition and Properties

The cell membrane is composed of two layers of phospholipids which contain both hydrophilic head and hydrophobic tail.[17] The hydrophilic region has the ability to interact with water due to the presence of polar phosphate groups, while the hydrophobic region of the molecule does not interact with the water because it is non-polar. [18] These basic structural features of the plasma membrane suggest that cell penetrating peptides can directly interact with the phospholipid chains, by altering the structure of the membrane and lead to peptide translocation.[17]

In the late 1980's, proteins that can enter plant cells were first discovered, from the understanding that hydrophilic molecules cannot pass through the plasma membrane.[19] However, studies have shown that Trans-Activator of Transcription (Tat) protein of the human immunodeficiency virus has the ability to translocate cell membrane.[20] In addition, a study has shown that a synthesized 60 amino-acid similar to *Drosophila Antennapedia homodox melanogaster* could enter nerve cells and control neural morphogenesis.[21] These result in the development of new peptides which can translocate through the plasma membrane, such as cell penetrating peptides.

Cell penetrating peptides are short peptides that can transfer polar hydrophilic molecules such as DNA and protein through the plasma membrane.[22] The cell penetrating peptide sequences contain residues such as arginine (R) and lysine (K). Arginine and lysine residues, have a large positive charge due to their large amount of basic amino acids that can bind easily to the surface the cell membrane.[18] Electrostatic interactions between the anionic phosphate group of the plasma membrane and the cationic peptide results in the binding of the peptide at the surface of the cell.[23-24] The mechanism of CPP translocation into living cells have been shown to follow mostly the endocytotic pathway.[25] The endocytotic pathway mechanism has been used to study the internalization of CPP-cargo complexes.[7] This pathway is categorized by forming a vesicle to transfer fluids, and macromolecules into living cells.[1] Cationic CPPs such as poly-arginine, poly-lysine, and Tat can mediate endocytosis through caveolin mediated, macropinocytosis, and clathrin. During the internalization of CPP into the cell, the vesicle content can be affected either by transporting to the endoplasmic reticulum or stored in the caveosome in order to get the final target, which depends solely on the type of endocytotic pathway involved.[5] Therefore, CPPs must be able to cross the vesicle, be released and able to reach the cellular compartment.

Another mechanism known for CPP internalization into the living cell is the direct membrane translocation. These mechanisms involve the disruption of the lipid bilayer by pore formation and micelles. In the "inverted micelle model" according to Derossi et al.[26] the interaction of the cell penetrating peptide with the lipid membrane will lead to an interaction of hydrophobic residue within the membrane core followed by the destabilization of the cell, and formation of an invaginated membrane. These destabilization results in an inverted micelle formation that confine CPP in the hydrophilic

core. It was shown that the interaction of amino acids arginine and lysine with the cell membrane using the inverted micelle model.[27] The interaction of the amino acids with the membrane was suggested to be due to interaction of guanadinium and amine group with the hydrophilic head group of the plasma membrane. However, the limitation of this method is the inability to translocate higher molecular weight peptides.

In addition, the interaction of peptide with the phospholipid chain can lead to pore formation in the membrane. During these formation, the peptide which is assumed to have an amphipathic alpha-helix structure inserts into the membrane, where the aliphatic chain of the phospholipid interacts with the hydrophilic face of the peptide to form a pore. Although, compared to the inverted micelle model, larger molecular weight peptides can be transported in this mechanism, the disadvantage, however, is the destabilization of the plasma membrane.[28] It is likely that the above-mentioned models may not completely explain the mechanism of penetration of different CPPs. More data using different CPPs and cargo molecules are needed to complete our understanding of the mechanism of CPP translocation.

One of the advantages of CPPs is their ability to interact with biological active cargoes. The cargo molecules have been reported to be effectively delivered by CPPs, including small molecules, fragment DNA and PNA etc.[29] The thiol group of cysteine was used to bind cargo to transporatan.[30] The thiol group was cleaved to a disulphide bridge between the cargo and CPP, which results in the release of the cargo into the cell.

CPPs are classified into four categories cationic, primary amphipathic, secondary amphipathic, and hydrophobic CPPs. Cationic peptides were the first to be discovered and was derived from H1V-1 protein, specifically the Tat (RKKRRQRRR) domain.[31] For adequate uptake, studies have shown that the length of the cationic CPPs required for

cellular uptake starts from chains consisting of at least eight amino acids. In addition, larger amounts of arginine amino-acid increases the chances of cellular uptake.[32] Cationic CPPs such as nuclear localization sequence (NLS) which consist of lysine, arginine, and proline can translocate into the nucleus through the nuclear pore complex. Although, most NLS have less than eight positively charged amino residues and as such are not good CPPs, they can attach to a hydrophobic peptide sequence to form an amphipathic CPPs useful for cellular uptake.[33]Another CPPs that was discovered is the primary amphipathic CPPs.

Primary amphipathic CPPs were derived from proteins that are covalently attached to a hydrophilic peptide.[34] They have about fifty amino acid residues in the sequence rich in tryptophan, leucine and isoleucine.[34] Unlike the primary amphipathic CPPs that have hydrophilic residues, secondary amphipathic CPPs have both hydrophobic and hydrophilic faces separated in the helix.[35] Using a model amphipathic peptide MAP (KLALKLALKALKAALKLA), it was shown that by adding lysine along with polar residues will increase uptake if only the amphipathic feature remains inert.

1.3. Characterization of CPPs

Cell penetrating peptides contain positively charged polar amino acids, including arginine, lysine, glutamine, serine, and alanine etc. (Figure 1.3 a, b, c).

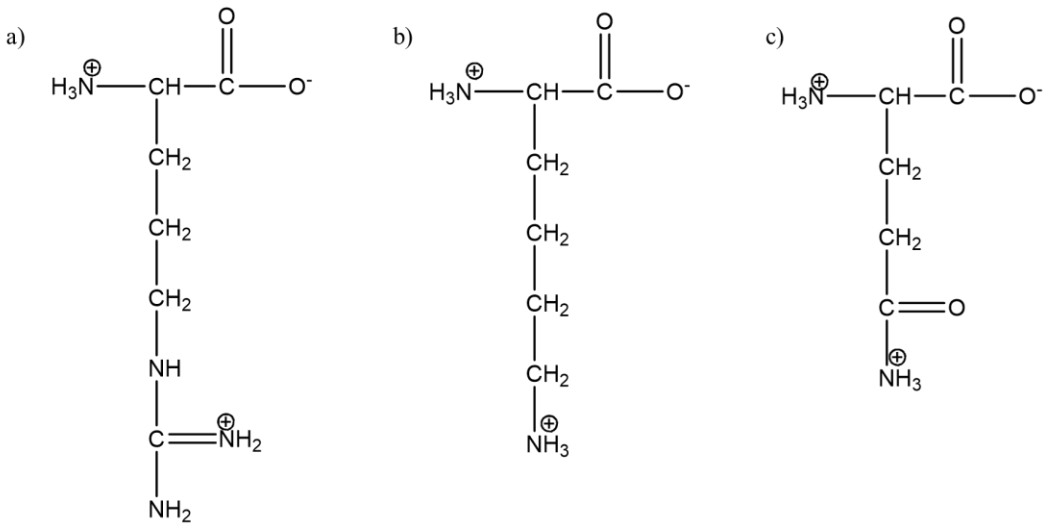


Figure 1.1. Structural examples of CPP a) Arginine b) Lysine c) Glutamine at pH 7

Many studies have been carried out to characterize CPPs including, methods such as scanning electron microscopy (SEM), atomic force microscopy (AFM), light scattering, X-ray crystallography and NMR, amongst others. These methods have been used to measure parameters such as electrophoresis assay, overall size of nanoparticle in AFM and SEM, structural, and conformational studies.

Solid-state NMR was used to observe the structure and dynamic interactions of CPPs and antimicrobial peptides (AMP) with lipid membranes.[36] Tat and penetratin were chosen as the CPPs, while Protegin-1 (PG-1),[36] and Neutrophil-peptide -1(HNP-1), as the AMPs. 2D ^{13}C detected ^{1}H spin diffusion experiment was used to observe the interaction between both arginine guanidinium- water phosphate and lysine-lipid. The result shows that Tat has a cross peak with the water-phosphate, which confirm Tat strong interaction with the cell membrane.[37] In addition, another study show a short distance of 4.3 Å between ^{13}C ξ-^{31}P of Arg10 and 4.0 Å between the side of ^{13}Cε-lipid ^{31}P in Lys13; the measurement using a ^{13}C-^{31}P REDOR NMR prove guanidinium- phosphate complexation.[36]

Fluorinated amino-acid has also been shown useful for protein NMR. [38] This is because ^{19}F nucleus is a sensitive nucleus of ½ spin with 100% isotopic abundance, naturally not common in most biomolecules, which result in their good spectral interpretation. Fluorinated proteins have been widely used to study biomolecules of high molecular protein and protein-ligand binding.[38]

The use of 5-fluoroindole instead of 6-fluorotryptophan is about 15-fold less expensive and allows for the study of more fluorinated amino-acids .[39] In addition, it has been shown that *Escherichia coli* BL21 use fluoroindole for biosynthesis of protein that has ^{19}F-tryptophan.[40] 6-fluro-tryptophan and 5-fluoroindole shows multiplets peaks at -46.1 and -50.1ppm in the ^{19}F NMR spectra, respectively. The ^{19}F nuclei in the B1 domain of protein G (GB1) in both 6-fluro-tryptophan and 5-fluoroindole lead to signals at -45.1 and -48.2 ppm which shows the mobilisation of the starting material. For this reason, fluorine NMR is a power technique in determining the interaction of large molecular weight amino-acids with cell membrane.

1.4. Deoxyribonucleic Acid (DNA) – Definition and Properties

DNA can be thought of containing the necessary commands that aid the production of proteins and other biomolecules crucial to cell function. Biomolecules such as proteins can interact with DNA including, participating in structuring chromatin, methylating DNA, transcribing DNA sequences, replicating and repairing it. An understanding of the DNA structure is important to know how DNA can encode the functions of genes.[41]

The first structure of DNA as proposed by Watson and Crick suggested that the two DNA strands called the polynucleotides consist of smaller units called *nucleotides*. The *nucleotides* consist of three groups which are bases, sugar-ribose, and the phosphate group.

DNA consists of four bases which include, Adenine (A), Guanine (G), Thymine (T), and Cytosine (C). These bases are connected covalently to the sugar-deoxyribose phosphate backbone, which has a phosphate at the 5'end. According to Chargaff, the nitrogenous base is bonded together to formed a double stranded DNA. This follows the rule (A-T and G-C), were A-T are formed by two hydrogen-bonds, binding the nitrogenous bases of each separate polynucleotide, and G-C by three hydrogen-bond binding the nitrogenous bases.

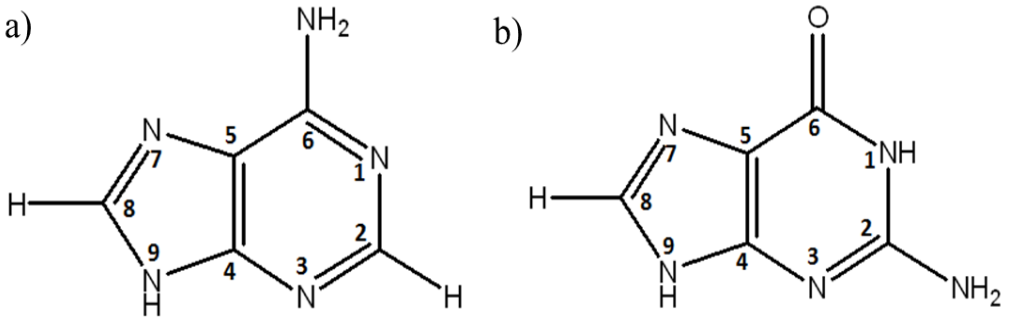


Figure 1.2. DNA Purine Rings a) Adenine b) Guanine

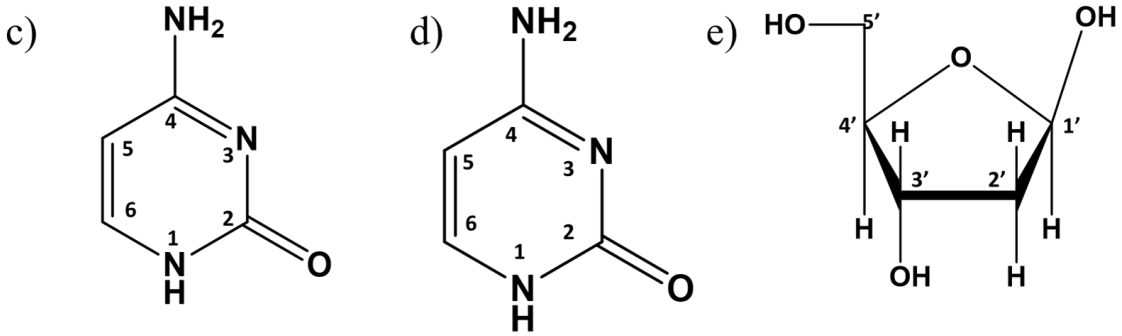


Figure 1.3. DNA Pyrimidine and Ribose Rings c) Cytosine d) Thymine e) 2''-deoxy-β-D-ribose sugar

Bases in DNA consist of both ***purines*** **(**Figure 1.2); and ***pyrimidines*** (Figure 1.3c,1.3d). The *purine* (A and G) are heterocyclic organic compounds that consist of *pyrimidine* joined with an *imidazole* ring while the *pyrimidines* (C and T) consist of a six-membered aromatic heterocyclic organic ring with two nitrogen atoms at position 1 and 3

of the ring. Sugar-phosphate backbone consists of deoxyribose (5-membered carbon) sugar. These sugars are connected to the phosphate groups, which forms phosphodiester bond between the 3', and 5'carbon atoms of the next sugar rings as shown in Figure 1.3e.

1.5. Characterisation of DNA

Over the last decade, the structure of DNA has been studied to get a better understanding of how the DNA helix are formed, and their interaction with other molecules. Recently, different techniques have been used to study the structure of DNA, which includes X-ray, Crystallography, NMR, etc.

For instance, a crystal structural characterisation of an exocyclic DNA adducts: 3, N^4-etheno-2'-deoxycytidine in the dodecahedra 5'-CGCGAATTεCGCG-3' at 1.8Å resolution has shown that exocyclic DNA base adduct can influence the structure of oligonucleotide.[42] This is due to the easy rotation around the sugar-phosphate backbone which may cause a displacement of large base adduct into the major groove of the DNA helix. In light of this, a high-resolution NMR spectroscopy of the same DNA has been observed and suggested an anti-orientation around the glycosidic torsion angle which confirms a right-handed helical conformation and aligns with the Watson-Crick canonical base pair model.[43] The cross-peak volumes were integrated and plotted against the mixing times using a three jump-and-return pulse. The result of this method was obtained by comparing build-up curves for cross peaks between two protons in which their distance is fixed. However, the limitation of this approach is the inability to integrate over-lapping cross-peaks that are not well resolved.

Electron paramagnetic resonance spectroscopy (EPR) has also been used to study the structure and dynamics of proteins, DNA, and RNA.[44]. EPR was used to determine the

relative motion of nucleotides in a C/G hairpin loop 5'-d (GCGC) and 5'-d (CGCG). In these molecules, it was shown that the second nucleotide at the 5'end has more mobility than the other nucleotides.

An X-ray diffraction and NMR study of a heterogeneous mixture of 3-dimensional DNA molecules have been presumed to have the same monotonous structure of 360^{o} angle of helical twist between the adjacent nucleotide pairs.[45] Shiyan Xiao et al. have also used a neutralization method to study the changes within the cationic region of the phosphate group in a DNA duplex.[46] In their findings, they observed that neutralizing the net charge in the nucleic acid, and condensing the ion layer surrounding, the molecule can reduce repulsion through electrostatic screening which will affect the structural stability and dynamic of DNA. Moreover, the effect of this neutralization will increase in the major groove width, decrease in the minor groove width, and reduce the global direction preference.

NMR techniques such as 2D and 3D NMR spectroscopy have been used to characterize DNA sugar-phosphate backbone, geometry, and glycosidic bond dihedral angle. The sugar conformation of a DNA fragment d-ACATCGATGT has been studied using a 2D-NOESY and ^{19}F phase angle (P) experiment used in measuring the non-planar ring in the sugar phosphate.[47] Using this approach, it was shown in the that the distance between H_1'- H_3', and H_2'-H_4' in the intra-sugar H-H distance is usually greater than 3.7Å, and will, therefore, give rise to a NOESY cross peak at short mixing times. Whereas, the distance between H_1'H_2'', 2'-3', 1'-2', 2''-3', and 3'-4' which vary within a range of 0.2 Å, therefore, give a very strong NOESY cross peak. However, the distance between 1'- 4', and 2'-4' varies between 2.8 and 5.0Å, which makes them useful for sugar phosphate backbone conformations. For the DNA fragment stated above, it was observed that the distance between H_1'-H_2'' ranges from 2.6 to 3.1Å.

A method called CHEOPS (chemical shift *de Nova* structure derivation protocol exploring singular value decomposition and backbone restrained molecular dynamic) has also been used to determine the 3-dimensional structure of RNA.[48] This method uses the proton chemical shift parameters which are either an experimental or simulated analysis to find errors in the data that have already been used to determine the structure of a helix. In their findings, they concluded that the experimental data from proton chemical shift was sufficient to determine the helix structure.

1.6. Complexation of CPP: Nucleic Acids

The understanding of the CPP-cargo complexation is of great interest in plant biotechnology. The first finding was reported for histone-DNA complexation.[49] In these studies, the binding of a highly positively charged histone to anti-dsDNA antibodies was shown to be due to the interaction of the anionic site of the glomerular basement membrane (GMB) with cationic histone.[49] Another study has shown the complexation between a single-stranded DNA and ethylene diamine dendrimers using a molecular dynamics simulation.[50] It was also reported that the binding of the DNA-dendrimer was due to competition between the enthalpic and entropic rigidity of the DNA. The complexation was also dependent on pH, salt concentration, and the DNA sequence which allows stable complexation. Amongst the various classes of CPPs, arginine-rich CPP such as HIV-1, TAT, and poly- arginine emerge as the most commonly used CPPs for the delivery of macromolecules. The successful binding of the CPPs depends on the basic properties of the guanadinium head group in the arginine molecule. The guanadinium head group is capable of forming hydrogen bond with both sulfate and phosphate groups, and thus improves the binding affinity of the CPPs at the surface of the cell membrane.[51] Another method of enhancing the internalization of arginine-rich CPP depends

on the size of both the cargo and CPPs molecule.[52] At a certain size of CPP alone and when complexed with DNA, CPP can instantly translocate through the membrane.

For instance, Tat has shown an intrinsic ability to cross the plasma membrane. The main characteristic of this CPP is the guanadinium group of its arginine residue and the overall high positive net charge.[1] Tat plays a very important role in promoting replication in HIV-1 virus, by increasing the binding ability of TAR hairpin at the 5' region on RNA.[5] The non-transcription function of Tat in HIV virus has also been shown to impact HIV-1 translation[53], capping, reverse transcription, and splicing.[27]

A study showed that the Tat_2 mediated the delivery of plasmid DNA to triticale isolated microspores. The mechanism of internalization of Tat_2 in plant cells is not known since the study of cellular internalization of DNA associated with Tat_2 has been only shown for mammalian cells.[54] Studies have reported the cellular internalization of the basic domain of Tat and Tat_2 in triticale mesophyll protoplast.[8] The result shows that the cellular internalization of the Tat_2 dimer was about 1.6 times higher to that of Tat monomer, which was because of the number of arginine residues in Tat_2 as compared to Tat.

Recently, to improve the binding properties of DNA and RNA with peptides, researchers came up with methods of modifying the molecules, which include labelling the part of the molecule or substituting them with other materials. For instance, the sugar-phosphodiester backbone of Peptide nucleic acids (PNA) is replaced with an N-(2-aminoethtyl) glycine peptide. PNA has been shown to have a very stable binding affinity with RNA and DNA oligomers, even more than their equivalent DNA-DNA, DNA-RNA duplexes.[55] The ability of PNA to bind strongly to complementary oligomer sequence has made them useful in pharmaceutical research.[6] Substituting charged and chiral groups into

PNA backbone have been shown to increase complexation in macromolecular uptake.[56]For instance, a molecular dynamics study was used by Ida Autiero et al,. to study the structural properties that drive the complexation and binding of chiral D-Lys-PNA and its achiral D-Lys-PNA with DNA and complementary strand of RNA.[57] It was shown that the presence of the chiral group in the PNA-DNA complexes affect the stability of the complex and does not affect the formation of the anti-parallel duplex.[58] They also discovered a structural rigidity at position 5, 6, and 7 in the D-Lys-PNA-DNA duplex which was due to the WC-has in the central region of the duplex. Thus, it was concluded that the substitution of the three D-Lys in PNA backbone could be explored for the expansion of a newer PNA-based molecule that can distinguish DNA and RNA duplexes.[57]

The complex formation between peptides and DNA or RNA is also due to the motion within the macromolecules. [59] The motion of the macromolecules plays a vital role in the dynamics of the complexes formed, and can improve the function of DNA and RNA process. The optimization of interfacial line broadening in the protein-DNA complex caused by dynamic changes in the aromatic ring, was solved by Junji and co-workers.[47] The solution structure of Dead ring protein (DRI-DBD ly262-Gly398) and its DNA binding site (DRI DBD: DNA,26 kDa) was investigated. The Dead ring protein (DRI-DBD ly262-Gly398) came from a *Drosophila melanogaster*, a regulatory protein used in the development of embryo that interacts with DNA in the nucleus through a domain named AT-rich domain (ARID).[60] 2D (F1, F2)^{13}C-filtered NOESY experiments which use a $^{1}J_{HC}$ against chemical shift optimized adiabatic ^{13}C inversion pulse were used. Results show that using a complementary WURST optimized 2D [F2] ^{13}C-filtered NOESY experiment, it is possible to get the intra-DNA cross peaks and intermolecular NOE's between the protein-DNA protons. Moreso, the complexation formed between the DRI-DRD and the DNA causes a line broadening in the ^{1}H

spectrum. These results in slow exchange dynamic which cannot be resolved by optimizing the spectrum. To resolve this problem of line broadening, they introduced the removal of the aromatic side chains, which could cause a large-ring current and induced distortion in the chemical shift close to the nuclei. Although, this work suggests a good method of optimizing line broadening problem that arises from a dynamic change in the close aromatic ring of protein-DNA complexes, however, this method can only be used if one of the molecules is labelled with ^{15}N or ^{13}C isotopes.

Crystallography studies have been carried out on specific protein-DNA complexes. However, these studies have been restrained to a specific binding of the protein on a single site of the DNA molecule. Although studies have shown the structural characterisation at atomic resolution, the biomolecular importance is not well understood. This is due to the fact the structure of the complex can be affected by crystal packing forces and cross-linking agents, and can, therefore, aid the binding at a single location site of the DNA.[61]

Recently, the structural characterisation of the non-specific protein-DNA complexes has been done without constraining the protein to a site of the DNA molecule.[62] In this study, different NMR techniques such as relaxation analysis, paramagnetic relaxation enhancement (PRE), and residual dipolar coupling (RDC) was used to study complexation of HOXD9 homeodomain with a non-specific, 24 base pair DNA duplex. Relaxation analysis shows that the transfer of protein from DNA to the other occurs directly. Also, comparing the RDC and PRE-analysis, the difference between the non-specific and specific DNA site shows that even at the same binding mode, the non-specific DNA of HOXD9 homeodomain diffuses continually as compared to the specific DNA. Surprisingly, in the heteronuclear $[^{1}H_{\varepsilon}]^{15}N_{\varepsilon}$ NOE analysis, arginine side chain that binds with the DNA move more rapidly in the non-specific site than the specific site.

Kwaku T. and his colleagues also looked into the dynamic process of HIV2.TAR RNA in the presence and absence of a ligand named arginiamide.[63] Using both ^{13}C and ^{15}N relaxation measurements, they were able to recognise a conformational change in TAR when it binds with arginine amide ligand. The measurement showed rigidity at the region where the complexation took place, as compared to results without complexation with the ligand with more flexibility. However, even with their successful findings in recognizing the dynamics of the complexation, they were not able to solve the exact mechanism for the conformation changes, and binding affinity.

1.7. Conclusion

The knowledge gained about the CPP has increased and become useful in nano-biotechnology to deliver cargoes such as DNA and RNA into plant cells. Further characterization of CPPs sequences and the mechanisms of their functions will allow to optimize their mode of interaction and facilitate the discovery of new CPPs that can be used for cellular uptake. The following thesis endeavours to illustrate the use of NMR technique in characterizing the structural features that give rise to complexation behaviour of CPP-cargo complexes and polyplexes, and the problems that must be overcome in other to make CPP an important biotechnology method for plant transformation.

Although, several factors affect the quality of NMR data which includes: molecular weight, spectral complexities and line broadening caused by chemical exchange. In this thesis, characterization of an unlabelled CPP-DNA complexation will be more focused on relaxation measurements and equilibrium concentration, which will give a better insight into their dynamic, and the stoichiometry of their complexation. The next chapter will introduce some NMR fundamentals, and the techniques employed in the later chapters.

CHAPTER 2

2. NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

2.1. Overview

Nuclear Magnetic Resonance (NMR) is a physical phenomenon in which the nuclei in a magnetic field absorbs energy. NMR spectroscopy was first carried out in 1946 by Bloch and Purcell, which won them a Nobel prize in 1952. NMR spectroscopy is used to characterise the structures of macromolecules at atomic resolution. Because NMR spectroscopy is good for the measurement of time dependence chemical phenomenon, it is used to investigate reaction kinetics, and intramolecular dynamics in biomolecular systems.

The application of NMR spectroscopy in this thesis will be based on the structural and dynamic characterisation of DNA and peptide complexation in solution state. This chapter will cover some fundamental techniques that have been used for this research and will also provide sufficient details of the theoretical concepts to explain the experimental methods.

In this chapter, the introduction to the NMR fundamentals was adapted from textbook by Keeler,[64] Jacobsen[65], Levitt[66], Hore[67], Günther[68], Zerbe, and Juot[69]. All figures shown herein also were generated by the authors.

2.1.1 Nuclear Magnetic Spin

The major properties of NMR spectroscopy are based on the magnetic characteristics of the atomic nuclei, which are independent of the type of the spectrometer. These nuclei possess an angular momentum ***L***, which give rise to a magnetic moment ***μ***, related with this equation:

$$\vec{\mu} = \gamma \vec{\boldsymbol{L}} \qquad |\boldsymbol{L}| = \hbar\sqrt{l(l+1)} \tag{2.1}$$

where γ is the *magnetogyric ratio* (in rad $T^{-1}s^{-1}$). Nuclear spin angular momentum *spin quantum number* (**I**), which varies from (0, 1/2, 1, 3/2… etc.) depending on the mass, atomic number and the nuclear angular momentum (**L**). The z-component of the angular momentum has associated quantum number m, that ranges from -I to +I integer steps. For example, the proton (^{1}H) has a spin quantum number and exhibit two spin states, α and β attributed to magnetic quantum numbers m_I = +1/2 or m_I = -1/2 respectively. It can be rewritten therefore from equation (2.1):

$$\mu_z = m_I\, \gamma\hbar \tag{2.2}$$

or

$$\mu_z = \pm\frac{1}{2}\,\gamma\hbar = \pm\gamma\hbar$$

In the presence of the magnetic field B_0, they have two different energies as described in Fig 2.1 below, only when the magnetic field direction defines the z-axis orientation. The energy is given as:

$$E = -\vec{\mu}.\vec{B} = -(\mu_x B_z + \mu_y B_y + \mu_z B_z) \tag{2.3}$$

where $B_z = B_0$

$$E = -m_I \gamma \hbar . B_0 \quad (2.4)$$

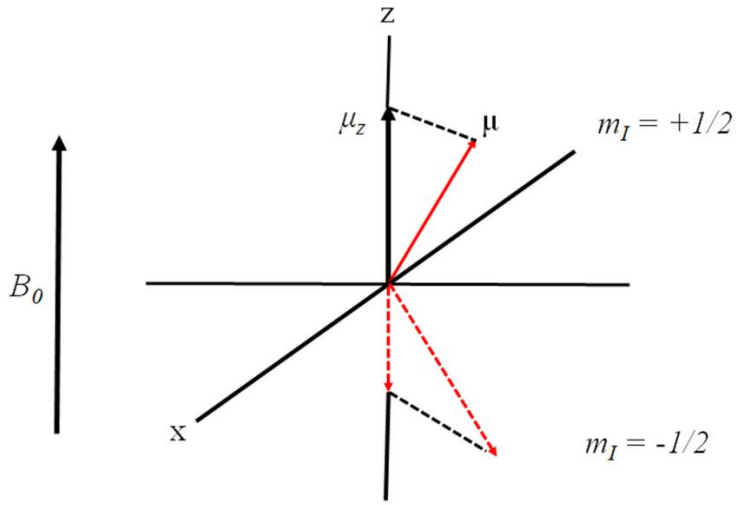


Figure 2.1. The proton spins states in the presence of an external magnetic field.

The energy of the lower spin state α (m_I =+1/2), is the more stable state with, $E_{+1/2}$ = $-\frac{1}{2}\gamma\hbar B_0$, and the higher spin state β (m_I =-1/2), is the less stable state with, $E_{-1/2} = \frac{1}{2}\gamma\hbar B_0$.

Thus, the energy difference between α and β are

$$E_\alpha - E_\beta = \frac{1}{2}\gamma\hbar B_0 - -\frac{1}{2}\gamma\hbar B_0 \quad (2.5)$$

Therefore, the energy difference $\Delta E = E_\alpha - E_\beta$

$$\Delta E = +\gamma\hbar B_0 = \omega_o \quad (2.6)$$

The number of nuclei in each spin state can be represented by Boltzmann distribution:

$$\frac{N_\beta}{N_\alpha} = e^{-\gamma B_0 / kT} \quad (2.7)$$

where N_β and N_α are the number of nuclei in the spin respectively, γ is the magnetogyric ratio, B_0 is the external magnetic field strength, k is Boltzmann constant and T is the absolute temperature.

Since $\Delta E \ll kT$

$$\frac{N_\beta}{N_\alpha} \approx 1 - \frac{\gamma B_0}{kT}$$

Therefore, the population of α and β are,

$\alpha = \frac{1}{2} + \frac{1}{2}\frac{\gamma B_0}{kT}$, $\beta = \frac{1}{2} - \frac{1}{2}\frac{\gamma B_0}{kT}$

$$\Delta n \cong \frac{\gamma B_0}{kT}$$

$$M_z \cong \Delta n . \mu_z \qquad (2.8)$$

2.2. Solution State NMR Spectroscopy

Solution sate NMR spectroscopy is a technique for studying the structural dynamic of biomolecules, even at high-resolution and residual level. In the analysis of this thesis, solution NMR was employed to characterise the structure of how peptide based nano-carriers bind to cargoes such as DNA. In the next paragraphs, focus will be on one-dimensional (1D) and two-dimensional (2D) NMR methods that were used for full structural assignment.

2.2.1. Chemical shifts and J-Coupling

In NMR spectroscopy, the frequency of the NMR line is directly proportional to the magnetic field strength, so increasing the magnetic field strength of the spectrometer increase the frequency scale. The magnetic dependence makes it difficult to compare frequencies of spectrometer with different field strength. To rule out this problem, the chemical shift scale is introduced, which is independent of the field strength of the spectrometer.

The magnetic field experience by the nuclei in a molecule are different for different electronic environment. For instance, the proton nuclei located in the -CH_2-Cl group in chloropropane will experience a different magnetic field than the protons located in -CH_3 and -CH_2 groups. This effect resulting to different chemical environment of the proton

nuclei in the molecule is known as the *chemical shift*. The chemical shift is connected to the resonance frequency which is proportional to the magnetic field

$$\upsilon_0 = \gamma B_0/2\pi \quad (2.9)$$

The chemical shift is caused by an induction of current from an external magnetic field B_0 in the electron clouds surrounding the nucleus. When this happens, a magnetic moment μ opposed to the B_0. This allows the local field at the nucleus to be smaller than the applied field by an amount equal of σB_0, Equation (2.18). Where σ is called the s*hielding constant*, σB_o is the magnitude of the other field induced at the nucleus.

$$B_{local} = B_0(1 - \sigma) \quad (2.10)$$

The shielding moves from left to right in its principle axis system, which describes the shielding as a factor of orientation. Mathematically, shielding is express as a 3×3 diagonal matrix given by:

$$\hat{\sigma} = \begin{vmatrix} \delta_{11} & 0 & 0 \\ 0 & \delta_{22} & 0 \\ 0 & 0 & \delta_{33} \end{vmatrix} \quad (2.11)$$

Due to the random re-orientational motion in solution state, averages the orientational dependence giving by $tr\sigma$ which is the trace of the shielding matrix:

$$tr\sigma = \frac{1}{3}(\sigma_{11}, +\sigma_{22} + \sigma_{33}) \quad (2.12)$$

The signals observed in the spectrum are due to various chemical shielding contributions, which includes: *diamagnetic, paramagnetic,* and *ring-current* effects.

Diamagnetic shielding effect is due to the motion of electrons induced by the external magnetic field B_0 thereby generating a local magnetic field. The new induced field is opposite to the external magnetic field B_0. This implies that the larger the electron density, the greater the shielding effect. This effect is very small in other nuclei except for

protons. It accounts for electrons that are spherically distributed; for instance, electrons from *s*-orbital of protons. The effect is related to the Biot-Savart law given by:

$$\sigma = \frac{\mu_0 e^2}{3m_e} \int_0^\infty r\rho(r)\, dr \tag{2.13}$$

where *r* is the distance of the electron from the nucleus, *ρ(r)* is the electron density, *e* is the charge and m_e is the mass of the electron. Alternatively, the *paramagnetic effect* is caused by the presence of unpaired electrons, induced by the external magnetic field, which leads to an increase in chemical shift of nearby nuclei. Paramagnetic effect is observed in the molecules p-orbital.

And lastly, the *ring-current effect* is mostly observed in π-orbitals of an aromatic ring system, which cause a deshielding effect in the proton resonances. For instance, in benzene were the π-electron moves freely around the ring, if an external magnetic field B_0 is applied, a diamagnetic effect is felt. In this case, the magnetic field produced by the current complements the external magnetic field, set at the middle of the ring and leads to deshielding of the protons outside the ring as shown in Figure 2.2. The proton above and below the benzene ring are aligned with the external magnetic field and as such are shielded.

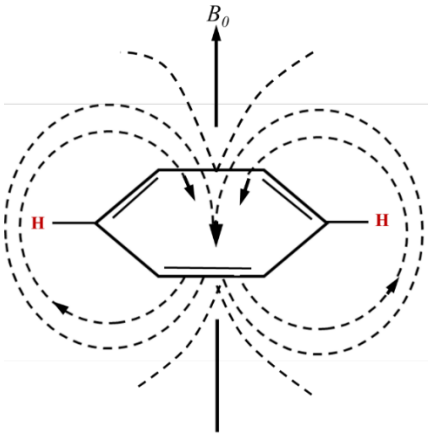


Figure 2.2. Diamagnetic effect of benzene ring in the presence of an external magnetic field B_0 nuclei opposing B_0 are deshielded, while those aligned with B_0 are shielded.

The resonance frequency of a particular nucleus is proportional to the strength of the magnet. Chemical shifts are usually the difference in the resonance frequency with respect to standard. The normalised value is very small and are multiplied by 10^6 which is represented in *parts per million* (ppm). For different spectrometers with varying magnetic field strength, chemical shift can be compared according to a nucleus x which is given by:

$$\delta_{(x)} = \frac{\omega_x - \omega_{ref}}{\omega_{ref}} \times 10^6 \tag{2.14}$$

where ω_{ref} is the standard frequency in which all other frequencies for instance, tetra-melthylsilane (TMS) which gives a sharp peak at zero ppm, are measured. ω_x is the operating frequency of the spectrometer in MH_z.

From the previous discussion about chemical shifts, one may have an impression that an NMR spectrum is only dependent on the chemical shift. This is not true because there is a magnetic interaction between the nuclei in a molecule, known as *spin-spin coupling* interaction or *J-coupling* interaction. J-couplings are used to evaluate the number of neighbouring nuclei, which depends on the number of bonds surrounding the nuclei.

In NMR, they are two coupling interactions that occur: dipolar coupling (direct interaction) and scalar coupling (indirect interaction). The dipolar coupling is not detected in solution state, because dipolar coupling relies on the direction of the internuclear vector with respect to the static field. In solution, molecules rotate fast that this effect is cancelled out and are not observed. However, *dipole-dipole* can be observed in liquid crystal and solid, and also drives relaxation.

In contrast to this, scalar coupling depends on the interaction between two nuclei through which the nuclei are indirectly connected. The coupling constant is always

represented as ^{n}J where “n” is the number of bonds involved; e.g. 1, 2, 3 bond couplings given as ^{1}J, ^{2}J, ^{3}J respectively.

In one bond coupling, for example ^{1}H-^{13}C bond, the $^{1}J_{1H\text{-}13C}$ coupling constant is usually positive and ranges from 100-250 Hz depending on the hybridization of the s-orbitals. The coupling constant increase from ($sp^3 < sp^2 < sp$) with *sp* being the highest, at 250 Hz, as it has the greatest characteristics. However, in two bond coupling also known as (geminal coupling) ^{2}J, e.g. CH_2 the coupling constants are usually negative.

The three-bond coupling called (vicinal coupling) coupling constant depends on the dihedral angle between the bonds, and are related to the Karplus equation given by:

$$^{3}J = A + B\cos\phi + C\cos^{2}\phi \qquad (2.15)$$

where A, B, and C are constants with values 7, -1, and 5, the dihedral angle ϕ ranges from 0^{o}-180^{o}, ^{3}J is the coupling from 2-12Hz show in the Karplus curve Figure 2.3.

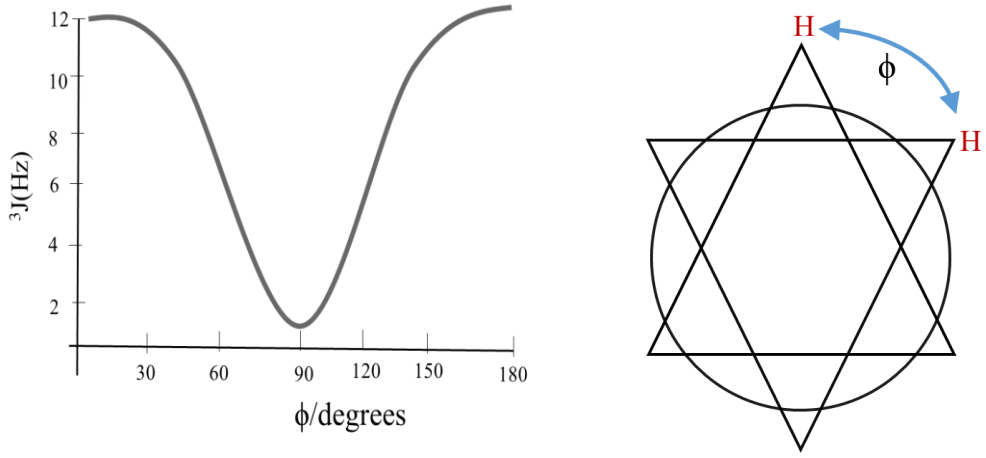


Figure 2.3. Dependence of three-bond coupling on dihedral angle and the Karplus curve for ^{3}J scalar coupling of two protons at an angle 90^{o}.

2.3. Generation of NMR Signal and Relaxation Measurement

Longitudinal magnetization at equilibrium does not depend on the time and is not detectable. Therefore, to generate an NMR signal, a radio frequency (RF) pulse is applied at the Larmor frequency to either x or y axis. This pulse allows the bulk magnetization to

rotate anticlockwise in the *x'-y'* plane. However, the degree of rotation depends on the amount of pulse applied as illustrated in the Figure 2.4.

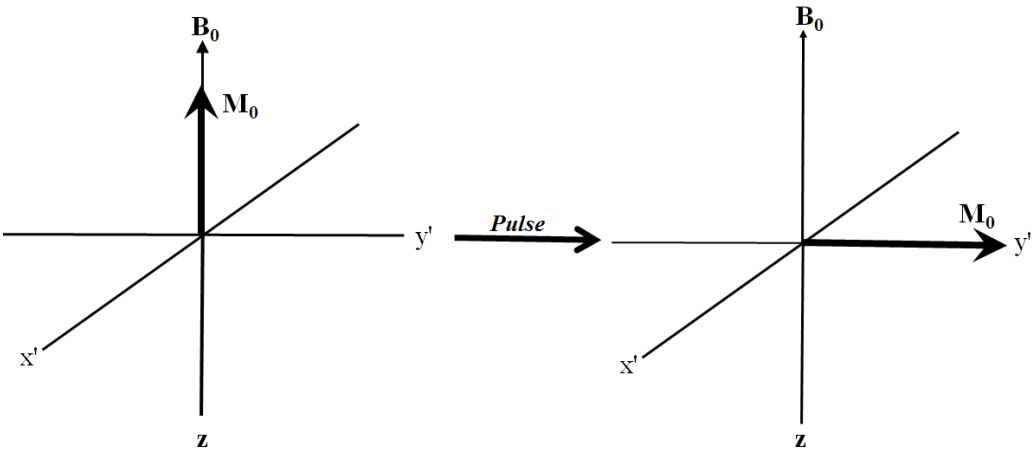


Figure 2.4. Precession of longitudinal magnetization M_z to transverse plane after a radio frequency pulse is applied.

Consequently, if the Larmor frequency is fixed, a strong pulse applied on the transverse plane, will cause the spins that are in antiphase to precess inphase with each other, and give rise to a transverse magnetization precessing at the Larmor frequency. The precession of the transverse magnetization will lead an oscillation in the magnetic field, and induces a current that is detected by the coil that was used to apply the radio frequency pulse. Overtime, when the r.f pulse is stopped, the transverse magnetization gradually returns back to it equilibrium state (M_z), such process is called relaxation.

The relaxation process that drives the spin to return back to its equilibrium state, has a timescale that is very important in the NMR experiment. Although, information such as coupling constants and chemical shifts are invaluable in characterizing the molecular structure, relaxation is related to dynamic orientation of molecules. During longitudinal relaxation, the z-magnetization at non-equilibrium (M_0= 0) Figure 2.5, slowly resolves its equilibrium value over time.

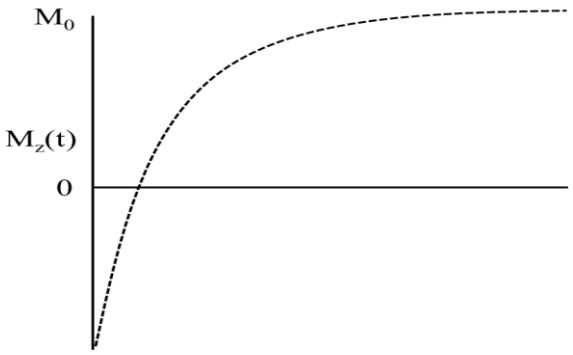


Figure 2.5. Relaxation of z-magnetization to an equilibrium value.

Since the fluctuations in the local magnetic field are caused by the thermal activated local motion, it is very convenient to study motion using relaxation measurements. In an ordinary NMR spectrum, these interactions are not detected. For instance, the interaction between a through space dipolar-dipolar couplings of a liquid cannot be observed. Whereas, relaxation measurements can reveal such process, which are useful for both structural and dynamic information of biomolecules. As mentioned earlier, the mechanism of relaxation can be divided into two: longitudinal also known as (T_1) and transverse relaxation (T_2).

Longitudinal relaxation is the process in which the magnetization along the z-axis is returned to equilibrium. The typical use for this measurement is the *inversion recovery* experiment. In which initially, a 180^o pulse is applied followed by a recovery time, at this time the relaxation occurs causing M_z to invert from $-M_0$ to the equilibrium value of M_0, after which a 90^o pulse is applied to detect the signal as shown in Figure 2.6.

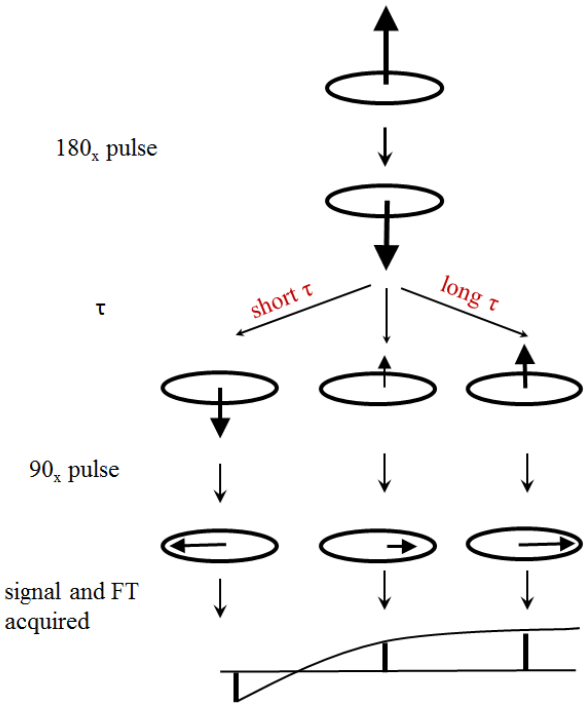


Figure 2.6. Pulse sequence of inversion Recovery method in Longitudinal relaxation experiment

The rate at which M_z returns to equilibrium is given by the first order rate constant T_1, according to the equation below:

$$M_z(t) - M_{eq} = (M_0 - M_{eq})e^{-\frac{t}{T_1}} \qquad (2.16)$$

at $t = 0$, $M_z(0) = M_0$

$t = \infty$, $M_z(\infty) = M_{eq}$

The signal intensity starts from negative value (ideally - M_{eq}) then passes zero, becoming positive at longer τ value. More so, in order to be quantitative one must wait for a duration period of 5-7 times T_1 between each scan. The intensity of each τ is also related to T_1 from this Equation:

$$M_z(t) = \left(M_0 - M_{eq}\right)\mathrm{e}^{-\frac{\mathrm{t}}{\mathrm{T}_1}} + M_{eq} \tag{2.17}$$

or

$$I(t) = (I_0 - I_{eq})\mathrm{e}^{-\frac{\mathrm{t}}{\mathrm{T}_1}} - I_{eq}$$

When a radio frequency pulse is applied on the transverse magnetization, all the spins precess inphase with each other around the Z-axis as long as the r.f is still on. When the r.f pulse is terminated, coherence is lost over time due to time dependence local field causing the transverse magnetization ($M_{x,y}$) to dissipate over time Figure 2.7.

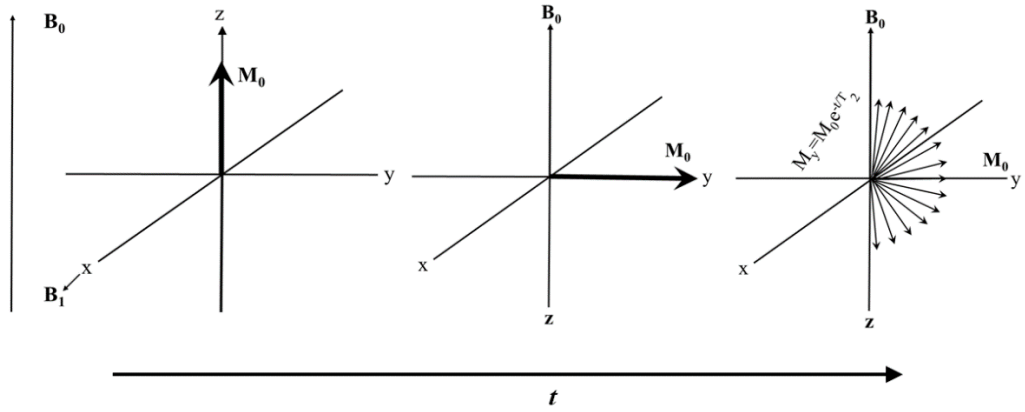


Figure 2.7. Coherence distribution at M_x and M_y plane, which leads to net transverse magnetization processing at the Larmor frequency.

The transverse components in the *x-y* plane decrease at the rate determined Transverse relaxation time. The dynamic of M_x+M_y on resonance is determined by these simplified Bloch equations:

$$\frac{dM_x}{dt} = -\frac{M_x}{T_2}\ , \frac{dM_y}{dt} = -\frac{M_y}{T_2} \tag{2.18}$$

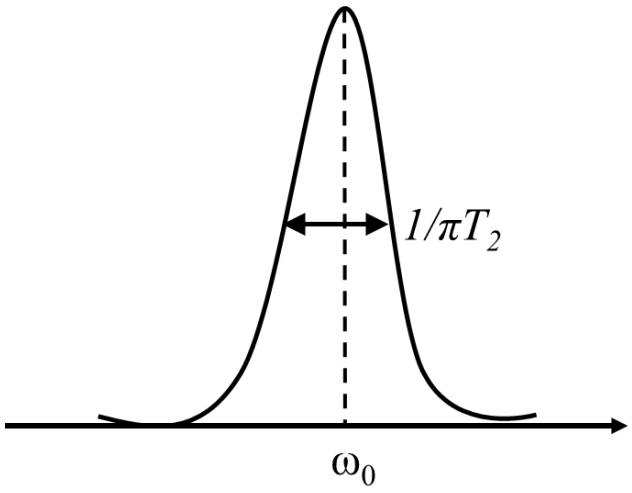
$$M_{xy}(t) = M_{eq}\mathrm{e}^{-\frac{\mathrm{t}}{\mathrm{T}_{22}}}$$

where T_2 is the transverse relaxation time constant and M_{eq} is the initial intensity of the transverse magnetization.

In transverse relaxation T_2, the transverse magnetization rotating at the Larmor frequency creates an oscillating magnetic field that induces a current in the signal-detection coil.

$$S(t) = S_{eq}\mathrm{e}(\mathrm{i}\omega_o \mathrm{t})\mathrm{e}^{-}\frac{\mathrm{t}}{\mathrm{T}_2} \quad (2.18)$$

where S_t is the signal intensity at a given time t and S_{eq} is the initial intensity. The signal intensity detected as a time function that can be converted into frequency domain using the Fourier transform (FT), which gives the final NMR spectrum. The NMR spectrum will be a plot of the signal intensity as a function of frequency that has a peak centred at the Larmor frequency ω_o (2.19) with a half-line width given by $1/\pi T_2$ Equation (2.1).



8.

Figure 2.8. Half line-width of Transverse Relaxation Magnetization.

$$\Delta v_{1/2} = \frac{1}{\pi T_2} \quad (2.20)$$

The motion of molecules results to relaxation from the fluctuation of the local field at the nucleus. For instance, if two spins are within the same molecule that interact with each other through dipolar coupling, as the molecule tumbles, the direction of the local magnetic field experienced by one spin on the other spin also tumbles. Since the dipolar

coupling between the two spins is dependent on the orientation of the molecule. Thus, relaxation times are also linked to motion of molecules.

The correlation function is used to characterise the time dependence of the motion in a molecule. If a spin experiences a local field $B_{loc}(t)$ at time τ. The correlation function is expressed as $\mathbb{G}(t, \tau)$. The time (B_t) it takes the local field to fluctuate is given by the autocorrelation function G(t) in Equation 2.20.

$$\mathbb{G}(t) = \langle B_t(t)B_t(t+\tau)\rangle \approx \langle {B^2}_t(t)\rangle exp - \left(\frac{|t|}{\tau_c}\right) \quad (2.21)$$

where B_t and $B_t(t+\tau)$ are the local field at time t, τ_c is the correlation time and it is the time it takes for a molecule to rotate at 1 radian. Correlation time is small for fast motion and large for slow motions. The autocorrelation function called the spectral density function given by:

$$J(\omega) = \frac{\tau_c}{1+\omega_0^2\tau_c^2} \quad (2.22)$$

From equation 2.14, at the extreme narrowing (fast motion):

$\omega_o\tau_c << 1$, $J(\omega_0) = 2$ (2.23)

Slow motion regime:

$$\omega_o\tau_c >> 1,\ J(\omega) = \frac{2}{\omega_0^2\tau_c} \quad (2.24)$$

Based on these assumptions, the relaxation rates $1/T_1$ and $1/T_2$ for a ^1H nucleus can be written as a function of spectral density by Equation 2.24.

$$\frac{1}{T_1} = \frac{3}{10}b^2\{J(\omega_0) + 4J(2\omega_0)\} \quad (2.25)$$

$$\frac{1}{T_2} = \frac{3}{20}b^2\{3J(0) + 5J(\omega_0) + 2J(2\omega_0)\} \quad (2.26)$$

where b is, a constant defined as:

$$b = \frac{\mu_0\gamma_1\gamma_2\hbar}{4\pi r^3} \quad (2.27)$$

where μ_0 is a physical constant, γ_2 and γ_1 are gyromagnetic ratios of two spins, r is the distance between the two spins.

The longitudinal relaxation time T_1 depends on correlation time τ_c. Correlation times are inversely related to rate of re-orientational motion. For small molecules that move fast correlation times are very short and T_1 are correspondingly long. For large molecules that move slow, τ_c is long and T_1 is short, T_1 decrease until it hit ($\tau_c = \frac{1}{\omega_0}$) where it slowly increases with τ_c, as illustrated in the curve below as shown in Figure 2.9.

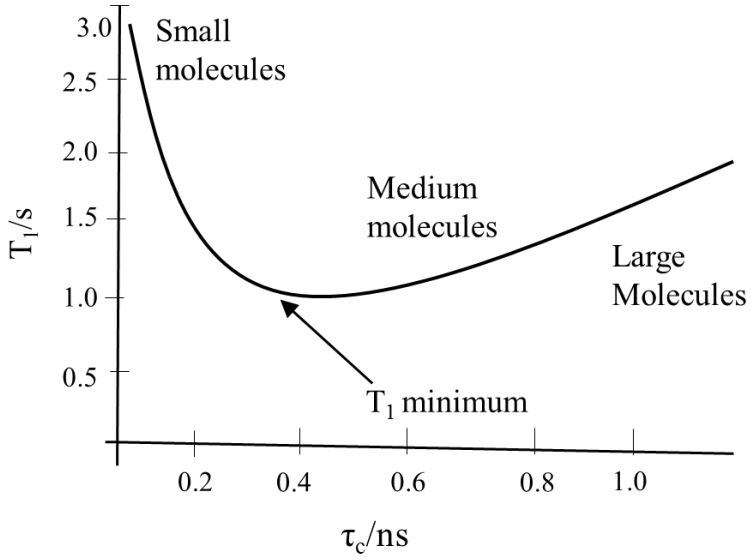


Figure 2.9. Longitudinal relaxation time constant as a function of correlation time varying field fluctuation.

From this curve, it can be observed that at room temperature, gaseous molecules have small correlation time, followed by a non-viscous liquid. When the molecule becomes bigger or in viscous liquids, the correlation times are larger. The longitudinal relaxation times also depends on temperature. This is because the random fluctuation begins around the molecule, and correlation times is temperature-dependent. Heating the sample makes the motion faster, hence, long τ_c. The T_1 relaxation time effect on temperature depends on

τ_c. Therefore, for a sample with long τ_c, heating the sample reduces T_1. For a sample with short τ_c, heating the sample increases T_1, as illustrated in Figure 2.10.

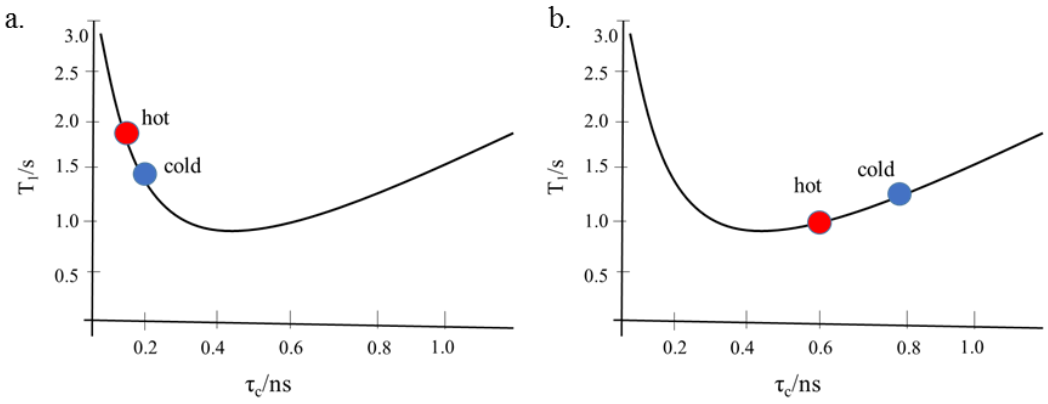


Figure 2.10. (a) Showing that at long correlation, heating the sample decreases T_1 (b) Short correlation time, heating the sample increases T_1

The mechanism that causes the motion of a molecule to randomly fluctuate, and to generate magnetic field for abundant nuclei with large gyromagnetic ratio. The *dipole-dipole* coupling interaction, drives relaxation. The dipolar interaction between spins occur when the magnetic field of one spin interacts with that of another. The relationship of this interaction is dependent on the internuclear distance *r* following approximately $1/r^6$, the spins, and the gyromagnetic ratios of the spins involved Figure 2.11. For instance, ^{1}H has larger gyromagnetic ratio as compare to ^{13}C, and will therefore have a larger magnetic moment.

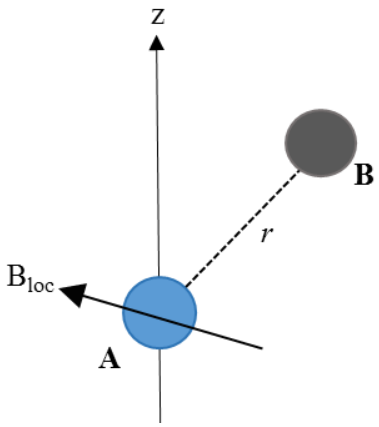


Figure 2.11. *Dipole-dipole* coupling between two spins A and B at the presence of an external magnetic field along the z-axis.

Practically, distance between the nuclei must be less than 5 Å for this pathway to be efficient. However, additional spins that have strong magnetic moment can give rise to relaxation. Comparing T_1 and T_2 with respect to correlation times, when the correlation time is short, the values of T_1 and T_2 are the same and both decrease monotonically with τc which is known as the *"extreme narrowing limit."* However, for larger molecules with larger correlation times, T_1 passes through a minimum value beyond which it then increases while T_2 continues to decrease.

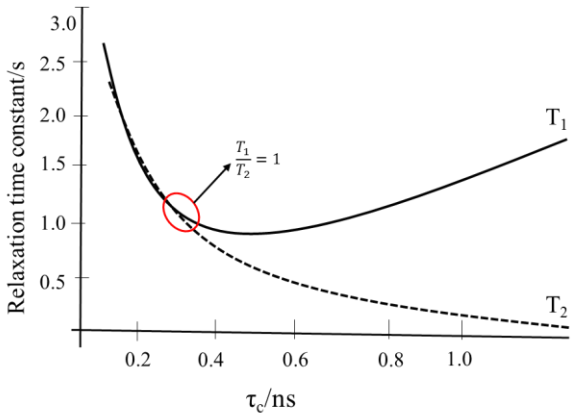


Figure 2.12. Curve of T_1 and T_2 with respect to correlation times.

Apart from the T_1 and T_2 relaxation measurement, another relaxation phenomenon is known as cross relaxation which is responsible for the Nuclear Over-Hauser Effect (NOE). The NOE is used to measure the spatial proximity of protons; and thus, depends on the *dipole-dipole* interaction. In the original NOE , low-power irradiation is applied selectively to a resonance of interest whose α and β populations are perturbed away from the Boltzmann's causing saturation Figure 2.13.

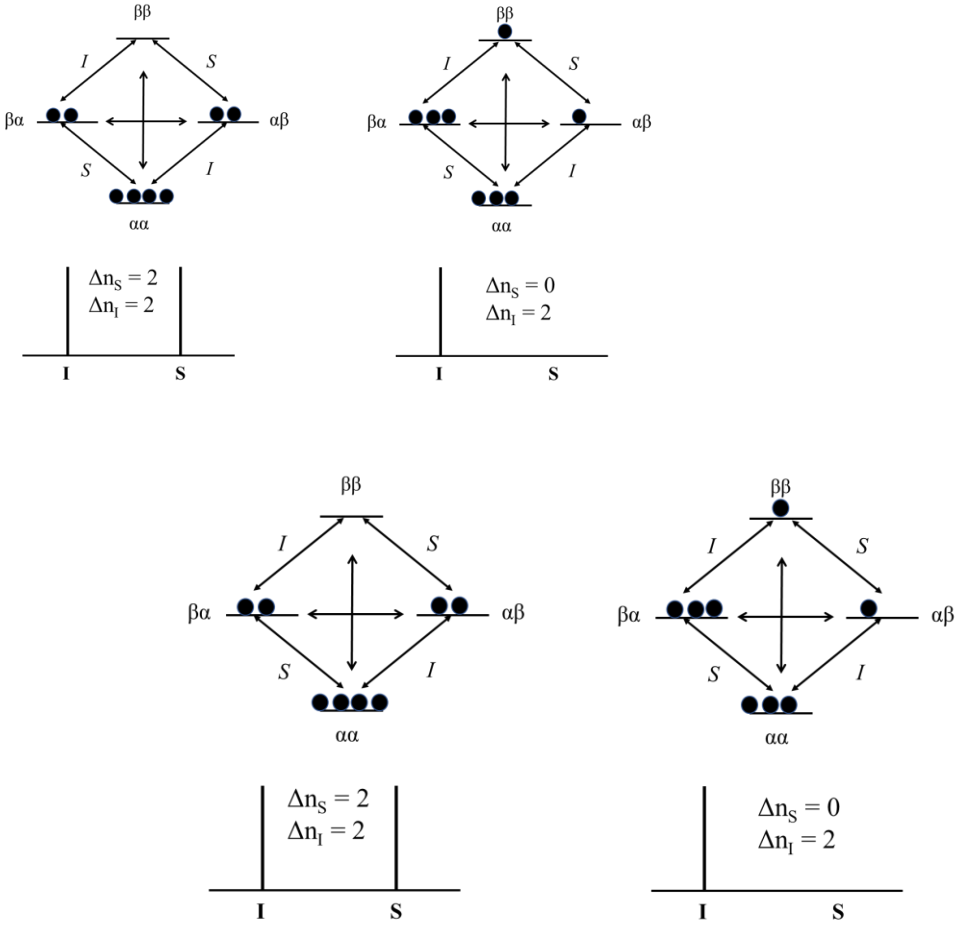


Figure 2.13. Distribution of population levels of a two-spins system **IS** consisted of a nuclear spin **I** and an electron spin **S**, which leads to inversion and saturation of signal.

The steady sate population of α and β that is achieved under the saturation signals, is a consequence of the dipolar interaction between spins I and S. The deviation change in

population The, attempt to restablish via Double Quantum Coherence (DQC), and Zero Quantum Coherence (ZQC).both resonance, where the intensity of the resonance not irradiated is enhanced.

2.4. One-Dimensional (1D) NMR Spectroscopy

2.4.1. ^{1}H NMR Spectroscopy

Proton NMR Spectroscopy which has a ½ spin is one of the most sensitive nuclei in NMR spectroscopy. This is because of their high abundance 99% and large gyromagnetic ratio, which makes them detectable even at very low sample concentrations. The combination of J-coupling and chemical shift can be used for full structural and dynamic characterisation of both large and small molecules. For an electronegative substituent such as *Br, I,* or *Cl* attached to the nucleus, the nucleus experiences substituent effect more strongly, because electron is drawn away, therefore removing electron density near the proton. Other contributions to chemical shift change are solvent and temperature. Although proton NMR is the experiment that is first observed for most NMR assignment, the long-range coupling between the protons and the overlap observed in complex molecules make it difficult for complete structural elucidation of molecules. For example, when ^{13}C is not attached to a proton (*quaternary carbon*), the signal of the proton cannot be observed in ^{1}H spectrum, which results in the study of ^{13}C NMR.

2.4.2. ^{13}C NMR Spectroscopy

As compared to ^{1}H NMR, carbon NMR spectroscopy has a low sensitivity which is due to their low gyromagnetic ratio. Their isotopes consist of ^{12}C and ^{13}C. ^{12}C has 98.9 % natural abundance but is not magnetic in nature, while ^{13}C has about 1.1% which is magnetic. Therefore, to acquire better resolution of signal, the sample must be very

concentrated. The advantage of ^{13}C NMR spectroscopy is that, it can be easily analysed because ^{13}C-^{13}C coupling are absent and ^{1}H-^{13}C are removed. ^{13}C are very useful in structural characterisation since they are less sensitive to factors such as temperature, solvent, and pH.

The ^{13}C chemical shift ranges from 0-200ppm, this large range allows the signal to be dispersed such that the overlap is less of a problem. To get a qualitative assignment of ^{13}C NMR sensitivity needs to be enhanced which can be done using *Polarization transfer pulse* sequence (INEPT and DEPT).

In the *Insensitive Nuclei Enhancement by Polarization Transfer* (INEPT), polarization is mostly transferred from ^{1}H into ^{13}C, which results in an increase of intensity of ^{13}C signal. Scalar coupling between the sensitive nuclei ^{1}H and the insensitive nuclei ^{13}C is mostly used in INEPT. As compared with INEPT experiment, another experiment that can be used to improve the sensitivity in ^{13}C NMR is known as *Distortionless enhancement by polarization transfer* (DEPT). In this experiment, like INEPT, polarization of magnetization is also transferred from sensitive nuclei ^{1}H into ^{13}C, however the signal intensity depends simply on the flip angle of the last pulse applied on the ^{1}H nuclei. This angle could be either 45^{o} also known as (DEPT 45), 90^{o} (DEPT 90), or 135^{o} (DEPT 135).

In DEPT 135 spectra, the signals of methyl and methane carbon are positive while the methylene signal is negative. In DEPT 90 spectra, only the methine (CH) carbon are observed and are positive.

2.4.3. ^{31}P NMR Spectroscopy

Phosphorus NMR spectroscopy is also a very sensitive nucleus of ½ spin and their spectrum are easy to interpret. The sensitivity of ^{31}P is due to its high natural abundance 100% and large gyromagnetic ratio. ^{31}P can be found in most organometallic compounds, metal co-ordinated complexes, and nucleic acids; and can be used to study the lipid bilayer of biological molecules.[70] ^{31}P has also been used in analyzing the dynamics and orientation of binding in protein with the lipid bilayers.

The chemical shift range of ^{31}P varies from 200 to -200 ppm and depends on the paramagnetic shielding tensor of the nuclei. The chemical shift of ^{31}P is generally referenced to phosphoric acid H_3PO_4 which absorbs at 0 ppm, are mostly observed with ^{1}H decoupled and give a sharp spectrum.

2.5. Two-Dimensional (2D) NMR Spectroscopy

One-dimensional NMR Spectroscopy experiments are routinely used in the determination of molecular structure. It suffers from the overlapping of signals obscuring the splitting patterns and signal position, making analyses next to impossible. To resolve this issue, 2D NMR spectroscopy is used which allows for the full assignment of small and large molecules. The signals in the 2D experiment are distributed over a second dimension, removing overlap. 2D NMR experiments are now rapid procedures that can produce reliable results if the acquisition parameters are properly calibrated.

2D NMR spectroscopy can be classified into two groups: *Homonuclear* experiments which shows the correlation between the same nuclei (e.g. ^{1}H- ^{1}H). And *Heteronuclear* experiment that shows the correlation between two spins of different nuclei (e.g. ^{1}H- ^{13}C). The general pulse sequence in 2D NMR experiment is categorized into four major time intervals: *preparation*, *evolution period* (t_1), *mixing time*, and the *detection time*.

During *evolution,* a 90^o pulse leads to transverse magnetization, which evolves in (t_1) and results to chemical shift in the indirect dimension.

The ability to understand how these experiment work is of great importance for characterizing the complexation in biomolecules such as DNA and peptides. In the next pages, I explained the four 2D experiments (COSY, NOESY, HSQC, and HMQC) that were used in this project.11

2.5.1 Correlation Spectroscopy (COSY)

COSY experiment is a Homonuclear 2D NMR experiment that relates the correlation between protons that are coupled to each other. This experiment requires two 90^o pulses, the first pulse creates a transverse magnetization which evolves about a time t_1 corresponding to the f_1 domain in the spectrum. Another 90^o pulse transfers coherence between the two spins, and relax at delay time t_2 which detects the f_2 domain as illustrated in Figure 2.14.

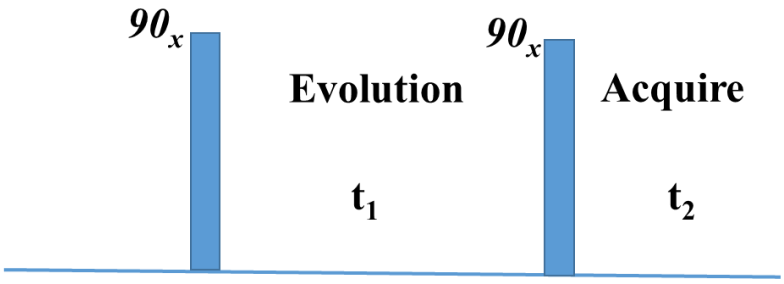


Figure 2.14. The two-dimensional ^{1}H-^{1}H COSY pulse sequence.

The spectrum shows *cross peak* and *diagonal peak*. The spin that changes the frequency during the delay time gives rise to the *cross peak,* while the part that does not change frequency gives rise to the *diagonal peak.* The *cross peak* shows the peak at both $\boldsymbol{f_1}$ and $\boldsymbol{f_2}$ frequencies that are related offset $\boldsymbol{f_1} = \Omega_A$, $\boldsymbol{f_2} = \Omega_B$, which indicates that spin at an

offset Ω_A is coupled to a spin at an offset Ω_B as shown in Figure 2.15 . The diagonal peaks also have the same offset as f_1 and f_2 .

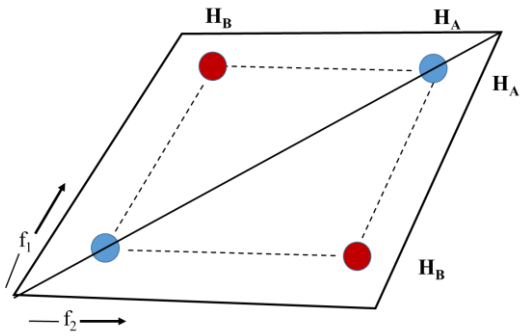


Figure 2.15. Schematic diagram of COSY spectrum for two spin system H_A, H_B. Red peaks are the cross peak; blue peaks are the diagonal peaks. Dotted line shows the correlations between H_B and H_A spins.

2.5.2. Nuclear-Overhauser Effect Spectroscopy (NOESY)

The NOESY experiment is also similar to COSY, the only differences between them is that in COSY, cross peak between the spins are acquired through a coherence transfer driven by the scalar coupling (through ***bond*** coupling) while in NOESY, cross peaks result from cross-relaxation between spin *i* and *j* (driven by dipolar coupling). The pulse sequence in NOESY is similar to COSY until the second 90^o pulse. After this pulse, a longitudinal z-magnetization that has either a positive or negative vector orientations; there is a mixing time t_m Figure 2.16, which depends mostly on the dipolar relaxation of the molecule. For instance, for small molecules the dipolar relaxation can be reduced to be less than T_1 (e.g. 0.7- 1.0), in large molecules, mixing time should be about (50-100ms). Finally, the last 90^o pulse produces a transverse magnetization and is acquired at time t_2.

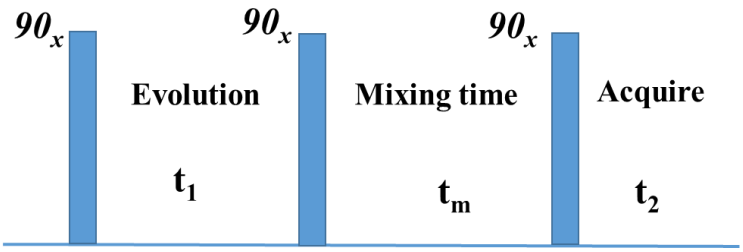


Figure 2.16. Pulse sequence for two-dimensional $^{1}H,^{1}H$ NOESY spectroscopy.

NOESY spectral shows cross peaks that are due to cross relaxation, mixing time, and diagonal peaks illustrated in Figure 2.17. Apart from the cross peak and the diagonal peak, axial peaks that arises from the magnetization that was recovered during the mixing time shows in the spectrum. The intensities of these peaks are opposite to the diagonal peak.

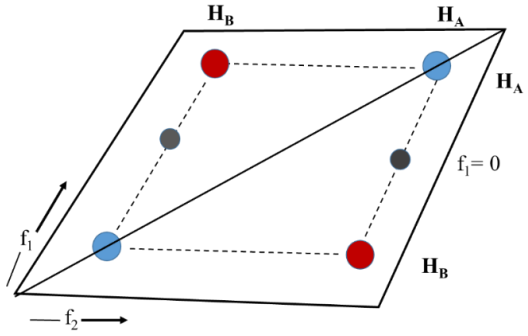


Figure 2.17. NOESY spectrum for two spin system during cross relaxation, positive peaks shown in blue (diagonal) and negative peaks in red (cross), grey axial peaks (f_1=0).

The problems that can be encountered in NOESY experiments arises from the zero quantum peaks due to protons that are scalarly coupled which gives a peak that resembles COSY. This type of artifact is mostly observed in small molecules that have sharp lines. For larger molecules, the zero-quantum effects are cancelled out during mixing time. Exchange peaks can also be observed in NOESY spectrum, if the spins are in a slow exchange regime.

2.5.3. Heteronuclear Single Quantum Coherence (HSQC)

Heteronuclear Single Quantum Coherence (HSQC) is a 2D NMR experiment used to measure the correlation between sensitive nuclei such as ^{1}H and an insensitive nucleus such as (^{13}C or ^{15}N) in which the proton is observed (*inverse method)*. The correlations between the two spins are usually about 150 Hz. The efficiency of the magnetization transfers from ^{1}H to ^{13}C in HSQC experiment depends on the size of the coupling between them.

The HSQC experiment consist of three major steps: firstly, transfer of the equilibrium magnetization from ^{1}H to ^{13}C spin; secondly, evolution of ^{13}C spin magnetization of T_1, and finally transfer of ^{13}C magnetization to ^{1}H spin and then detection. As illustrated in Figure 2.18.

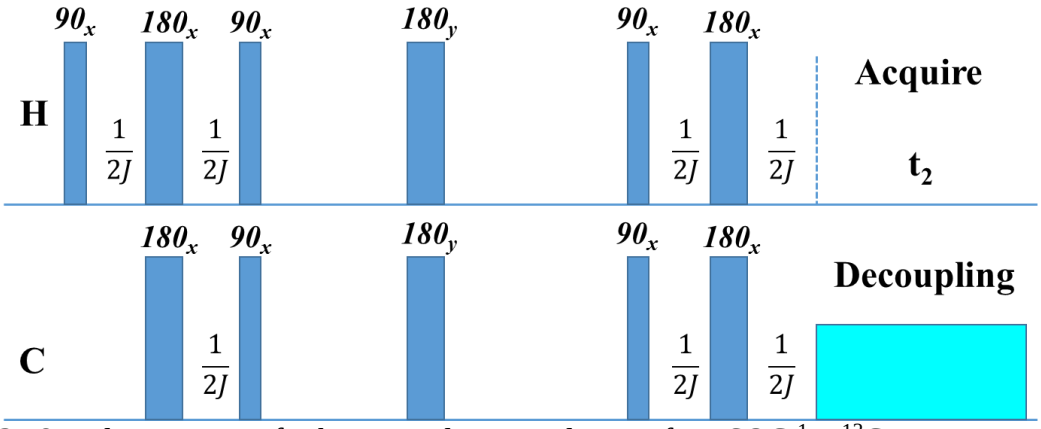


Figure 2.18. Pulse sequence for heteronuclear correlation of a HSQC ^{1}H-^{13}C experiment.

In Figure 2.18, the first set of the pulse is like the INEPT experiment in which evolution t_1 takes place in ^{13}C spin magnetization at a 180^{o} pulse on ^{1}H magnetization. This refocuses the evolution of the coupling, another 90^{o} pulse is then applied to transfer magnetization back to ^{1}H where it is detected. The intensity of the peak depends on delay period ($\frac{1}{4J}$, $\frac{1}{2J}$,..). The transfer of magnetization from ^{1}H to ^{13}C is also to improve the sensitivity of the signal, because proton has a higher Larmor frequency than carbon.

2.5.4. Heteronuclear Multiple Quantum Spectroscopy (HMQC)

Heteronuclear Multiple Quantum Coherence (HMBC) experiment as compared to HSQC experiment uses a long-range coupling between protons and ^{15}N, ^{13}C which has an advantage in observing quaternary and carbonyl carbon that can't be observed in HSQC.

As compared to HSQC pulse sequence, in HMQC experiment, the first 90° pulse allows transverse magnetization after a delay period τ, the equilibrium of ^{1}H spin becomes an anti-phase. A 180° on ^{1}H is used to created a DQC which evolved according to the ^{1}H ad ^{13}C frequencies. The ^{1}H frequency is eliminated by refocussing the pulse on the ^{1}H. And finally, the signal is acquired and observed under broadband ^{13}C-spin decoupling, illustrated in Figure 2.19.

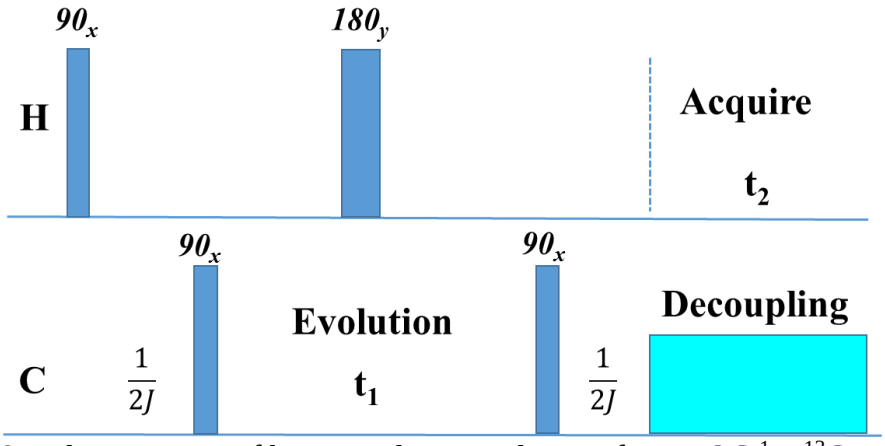


Figure 2.19. Pulse sequence of heteronuclear correlation of a HMQC ^{1}H-^{13}C experiment.

CHAPTER 3

---

3. STRUCTURAL ELUCIDATION OF SINGLE-STRANDED DNA AND DOUBLE-STRANDED DNA 5'-AGTCC-3' USING SOLUTION-STATE 1D AND 2D NMR SPECTROSCOPY

3.1. Overview

In this chapter, the structural characterization of single stranded and double stranded DNA was analysed using 1D and 2D NMR experiments. 1D -^{1}H and ^{13}C NMR experiments were employed to assign the proton and carbon resonances of the DNA. Also, 1D- ^{31}P NMR experiment was used to assign the phosphorus resonance of the DNA backbone.

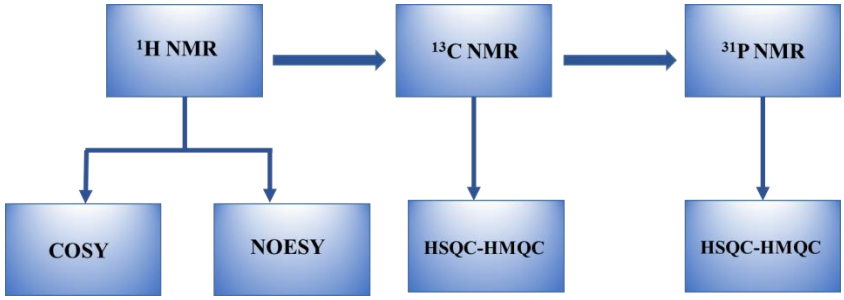


Figure 3.1. A schematic diagram showing the order of NMR experiment carried out for single stranded and double stranded DNA 5'-AGTCC-3', to assign the chemical structure and obtain full assignments.

2D NMR experiments such as COSY and NOESY was employed to understand the interaction between the bases and the sugar-phosphate backbone and also the sequential assignment of the DNA. 2D ^{31}P-^{1}H and ^{13}C-^{1}H HMBC NMR experiments were also to used to assign phosphorus, carbon and germinal protons interactions.

3.2. Experimental details

3.2.1 Materials and Methods

The single-stranded DNA 5'-AGTCC-3' and its complementary strand 5'-GGACT-3' was purchased at IDT (Ref no 136577768,136577767) and were estimated to be > 98% purity. The amino acids to be investigated are; Arginine (R), Tri-Arginine (RRR),

Nona-arginine(RRRRRRRRR) and Tat_2 (RKKRRQRRRKKRRQRRR) which were synthesis using Canpeptide, were estimated to be > 98% purity.

NMR samples of single-stranded DNA 5'-AGTCC-3' were prepared in D_2O at a concentration of 5mg/ml. For double stranded DNA, 5'- AGTCC-3' with its complementary strand 5'-GGACT-3' was prepared in 1:1 molar ratio with a total concentration of 5 mg/ml (3.4 Mm). All the CPP samples were also prepared in D_2O at concentrations of (14.35 mM) arginine, (5.14 mM) R_3, (7.02 mM) R_9, and (7.5 mM) Tat_2. 3-trimethylsilyl-1-propanesulfonate DSS was used as an internal reference for both ^1H and ^{13}C NMR measurements. Acetate buffer was used to adjust the pH to 7; by adding 114.5µl acetic acid (Sigma - 23315) into 50 ml D_2O (Sigma - 151882) and adjusting the pH with deuterium chloride solution (Sigma - 5430470) and sodium deuteroxide (Sigma - 372072).

3.2.2. One-dimensional NMR experiments

All NMR spectra presented were acquired using a temperature that ranges from (10-45^oC) and collected using Bruker Avance III HD 700 MHz spectrometer; equipped with a variable temperature unit (-150 -250^oC). The magnetic field strength is 16.5 Telsa, with corresponding larmor frequency of 700.4 MHz for ^1H, 176.12 MHz for ^{13}C, and 283.53 MHz for ^{31}P.

A typical 1D ^1H NMR spectrum for DNA molecules was recorded using a 90^o pulse width of 17.0 as 128 transients, with a relaxation delay of 4.0 s. For each transient, 128 K points were collected over a spectral window of 77 kHz for DNA. For the peptides, 16 transients were recorded, with a relaxation delay of 2 s and 56 kHz spectral window. Likewise, 1D ^{13}C NMR spectra for DNA molecules were obtained as 10,000 transients, using a 90^o pulse width of 12.0 µs, with a recycle delay of 5 s. Each transient has 256 K points over a spectral window of 416 kHz. 1D ^{31}P NMR spectra were also obtained using

a 90^{o} pulse width of 13.0 and 200 transients, with a relaxation delay of 4 s. For each transient, 32 K points were collected over a spectral window of 8.5 kHz.

3.2.3. Two-dimensional NMR experiments

A homonuclear gradient ^{1}H COSY spectrum of the DNA was obtained in magnitude mode at 25 ^{o}C as 512 increments. Each increment consists of 4 scans and collected over 2048 points, using relaxation delay of 8.0 s. Both direct and indirect dimensions have a spectral window of 58.0 kHz (8.4 ppm) and a digital resolution of 5.0 Hz and 1.1 Hz respectively.

The ^{1}H NOESY spectrum of DNA was also acquired at 25 ^{o}C as 128 increments over a spectral window of 5952.4 Hz (8.5 ppm) in both dimensions, using a relaxation delay of 8.4 s and mixing times of 0.3, 0.6, 0.9, and 1.0 s respectively. Sixteen transients were collected for each increment having 4096 data points. The direct and indirect dimension have a digital resolution of 5.8 and 93 Hz respectively, before zero-filling.

The ^{1}H-^{31}C HSQC spectrum was acquired as 256 increments, over a spectral window of 5952.4 Hz (8.5 ppm) in the direct dimension (^{1}H) and 29069.8 Hz (165.0 ppm) in the indirect dimension (^{13}C), using a relaxation delay of 1.8 s. Eighty transients were collected for each increment before zero-filling. The digital resolution for the direct and indirect dimension was 2.9 and 227 Hz respectively.

^{1}H-^{31}P HMQC spectrum was acquired in 256 increments, covering a spectral window of 9090.9 Hz (12.9 ppm) in the direct dimension and 2835.3 Hz (10.0 ppm) in the indirect dimension at a relaxation delay of 4.0 s. Linear forward prediction to 256 points were on the t_1 dimension and subsequently zero filled to a total of s8192 points. The digital resolution for the direct and indirect dimensions were at 2.2 and 22.1 Hz respectively.

3.3. Results and Discussion

3.3.1. NMR Assignment of Single-stranded DNA 5mer 5'-AGTCC-3'

The NMR assignment of the single-stranded DNA 5mer 5'-AGTCC-3' was carried out the methods discussed earlier in sections 3.2.2 and 3.2.3. The aromatic protons located on the bases are labelled H_8, H_6, H_5, H_2, CH_3 while the aliphatic protons on the sugar-phosphate backbone are labelled $H_{1'}$, $H_{2'}/H_{2''}$, $H_{3'}$, $H_{4'}$, $H_{5'}/H_{5''}$.

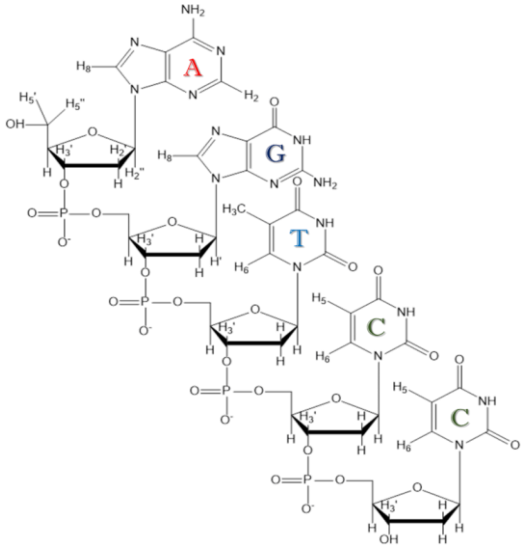


Figure 3.2. The chemical structure of single-stranded DNA 5mer 5'- $A_1G_2T_3C_4C_5$-3' showing the protons in their different chemical environments. It can also be observed from the structure that both 5' and 3' ends of adenine (A_1) and cytosine (C_5) have no linkage to a phosphate group as compared to the others.

Assignment of the ^{1}H-NMR Spectrum

The assignment for single-stranded DNA 5mer 5'-AGTCC-3' was carried out using 1D ^{1}H NMR experiments with phosphorus decoupling. Resonances corresponding to the sugar-phosphate backbone and the aromatic bases were assigned, corresponding to a total of seven proton resonances from the bases and thirty-five resonances from the sugar-phosphate backbone. The integration of isolated peaks seen in the spectrum allows the determination of the relative number of protons identified.

In the 1D ^{1}H-NMR, the aromatic protons are deshielded, and appeared at the high frequency (green) region, while the aliphatic protons mostly seen in the sugar-phosphate backbone are shielded, and appeared at the low frequency (red) region as shown in Figure 3.3.

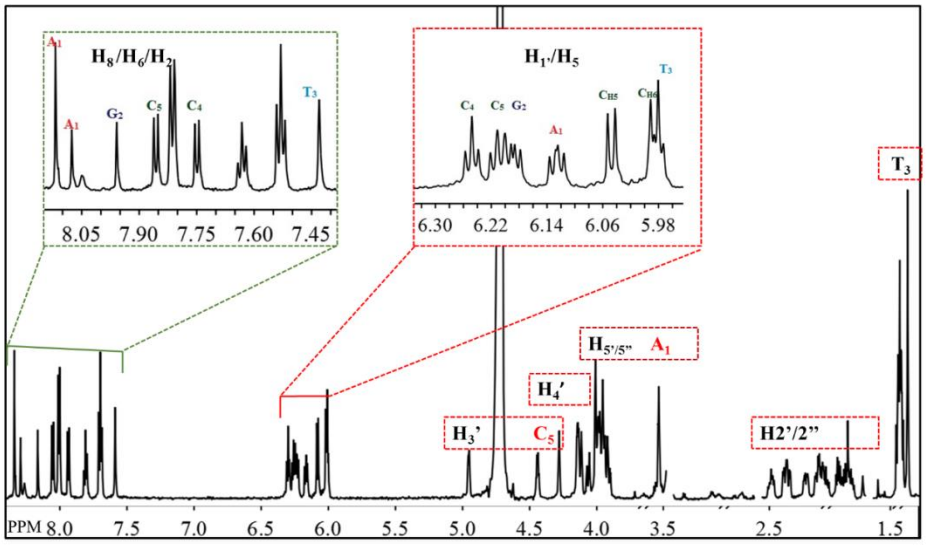


Figure 3.3. ^{1}H-NMR spectrum of single-stranded DNA 5'-AGTCC-3', measured in 99.9% D_2O at 22^{o}C. The chemical structure of DNA 5'-AGTCC-3' is shown in figure 3.2. Two different regions have been highlighted, the aromatic region (green) and the aliphatic regions (red) respectively. The different region of the spectrum has been zoomed to show some of the splitting patterns.

In addition, the protons in the H_3' of C_5 and H_5' of A_1 show a significant difference in their chemical shift (4.49, 3.50 ppm) as compared to the other bases at (4.51-5.00 ppm, 3.90-4.10 ppm) respectively. The difference in the chemical shift for A_1 and C_5 (H_3'/H_5') protons is because, oxygen were directly connected to them, whereas, the other bases were connected to phosphate groups, which leads to the higher chemical shifts observed. A similar observation can be made for the cytosine's base protons where H_5 (6.02, 6.08 ppm) appears in the aliphatic region and H_6 (7.90, 7.81 ppm) appears in the aromatic region of the spectrum, due to H_6 being close to a nitrogen atom, as compared to H_5 with a carbon atom. The sharp singlet at 1.70 ppm was assigned to the methyl group of T_3.

Significant overlap was observed for protons H_2'/H_2'' at (2.26 – 2.75 ppm) however, each proton gave a well-resolved doublet of doublets; proton H_2'/H_2'' is split into doublet ($^2J_{H2'-H2''}$ 14 Hz) due to germinal proton H_2''/H_2'. Another doublet comes from H_2'/H_2''coulpling with H_1' ($^3J_{H1'-H2'/H2''} \approx 6$ Hz) and a doublet from H_2'/H_2''coulped with H_3' which also splits into doublet due to coupling with neighbouring protons H_1'/H_3' ($^3J_{H3'-H2'/H2''} \approx 3/6$ Hz). Proton of H_1' (5.9 - 6.25 ppm) appear as triplets or doublets of doublets due to couplings with H_2'/H_2'' ($^3J_{H1'-H2'} \approx 6$ Hz), ($^3J_{H1'-H2''} \approx 9$ Hz). The cytosine's protons H_5 and H_6 also show a splitting pattern of doublets ($^3J_{H5-H6} \approx 7$ Hz) Figure 3.4 due to coupling with each other.

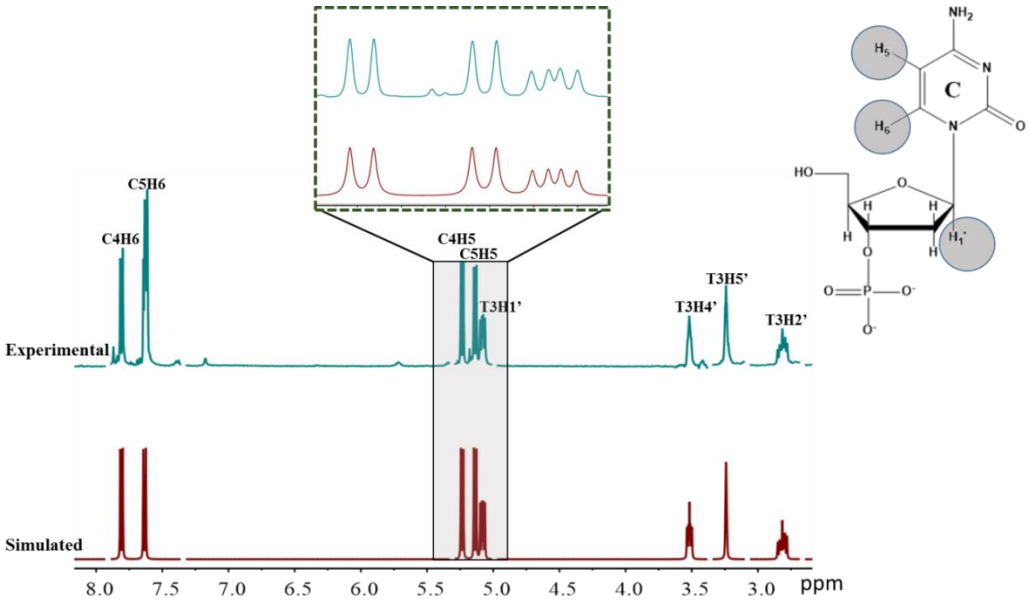


Figure 3.4. ^{1}H NMR spectrum of single-stranded DNA 5mer 5'-AGTCC-3' at 25 ^{o}C in 99.9 % D_2O. The experimental spectrum is shown in green and the simulated spectrum in red. An inset showing the H_5, H_6, H_1' of cytosine and thymine is provided to confirm the close agreement between experimental and simulated spectrum.

Overlapping of signals can be observed for H_5'/H_5'' coupling with H_4' protons. The H_3' (4.80 - 5.02 ppm) protons were not completely resolved at 22 ^{o}C because they have the same chemical with the water signal at 4.70 ppm. To increase the resolution of these signals,

^{1}H NMR at different temperature was performed at 15, 25, 35, and 45 ^{o}C Figure 3.5. Increasing the temperature resulted in decreasing chemical shift of the water signal and better peak resolution in the H_3' signals. The decrease in the chemical shift of water signal results from the fact that the hydrogen bond in the water molecule is weakened as the temperature increase, which results from an increase in the average distance of the hydrogen bond acceptor and donor.

Finally, coupling of the H_3' with phosphorus ($^{3}J_{H3'-P} \approx 6$ Hz) was observed in A_1, G_2, T_3, and C_4 bases except for C_5 base.

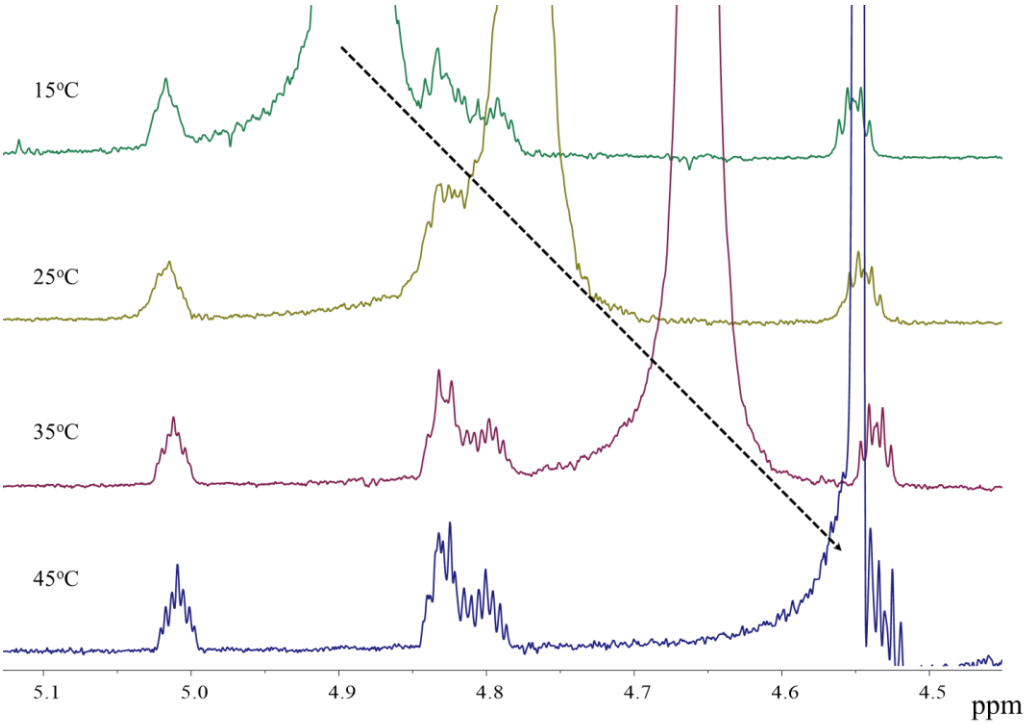


Figure 3.5. The effect of variable temperature to increase resolution. (black arrow) shows decrease in chemical shift of water signal as temperature increases and increase in resolution of H_3' signals as temperature increases.

Assignment of the COSY Spectrum

The assignment of single-stranded DNA 5mer 5'-AGTCC-3' was further analyzed using the 2D COSY experiment. This experiment allows correlation via scalar coupling of

protons that are over 2 to 3 bonds. In the aromatic region, three strong correlation were assigned between H_5-H_6 in the cytosines and CH_3-H_6 in thymine Figure 3.6.

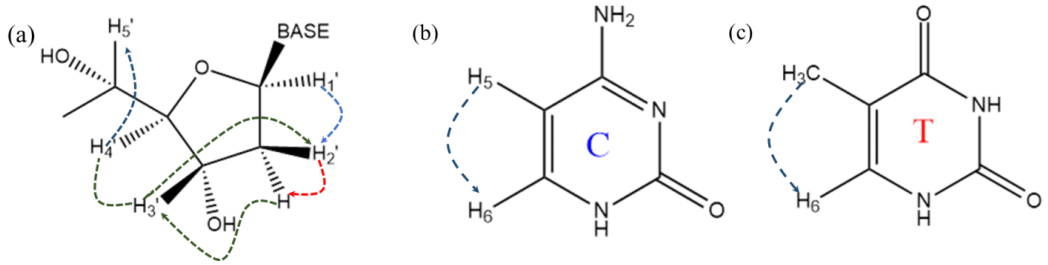


Figure 3.6. A schematic diagram showing the two bond couplings of; (a) sugar-phosphate (H_1'-H_2'/H_2'', H_2'/H_2''-H_3', H_3'-H_4', H_4'-H_5'/H_5'') (b) Cytosine's (H_5-H_6), (c) Thymine (CH_3-H_6).

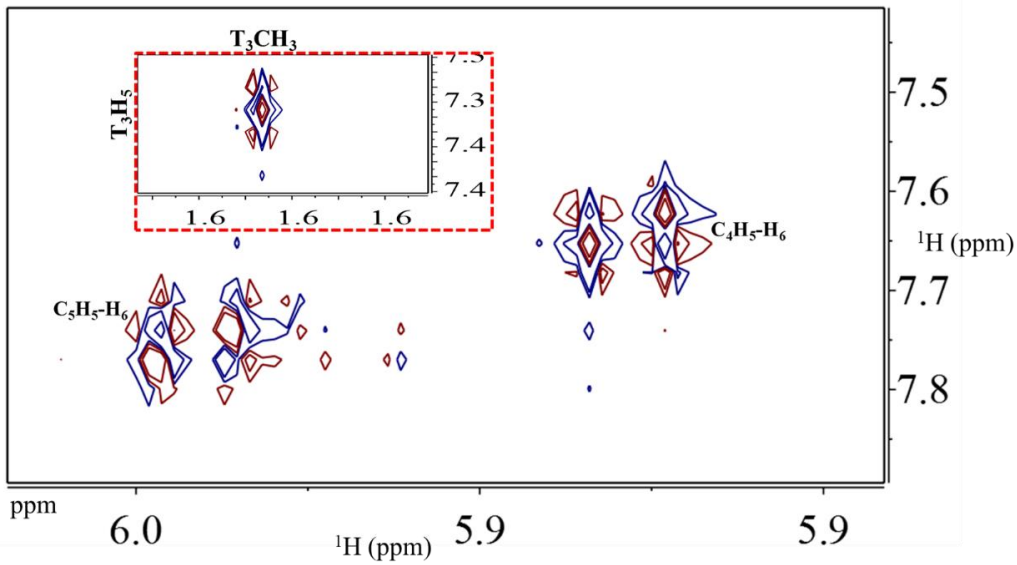


Figure 3.7. The aromatic region of COSY in 99.9% D_2O at 22°C. COSY cross peaks correspond to the H_5-H_6 of the cytosines and CH_3-H_6 of the thymine (red circle) correlations.

In the aliphatic region (Figure 3.7), a total of 10 strong cross-peaks were observed between H_1'-H_2'/H_2'' and H_3'-H_2'/H_2''. Cross-peak between geminal protons H_2'/H_2'' also appeared but are overlapped due to strong coupling between them.

a)

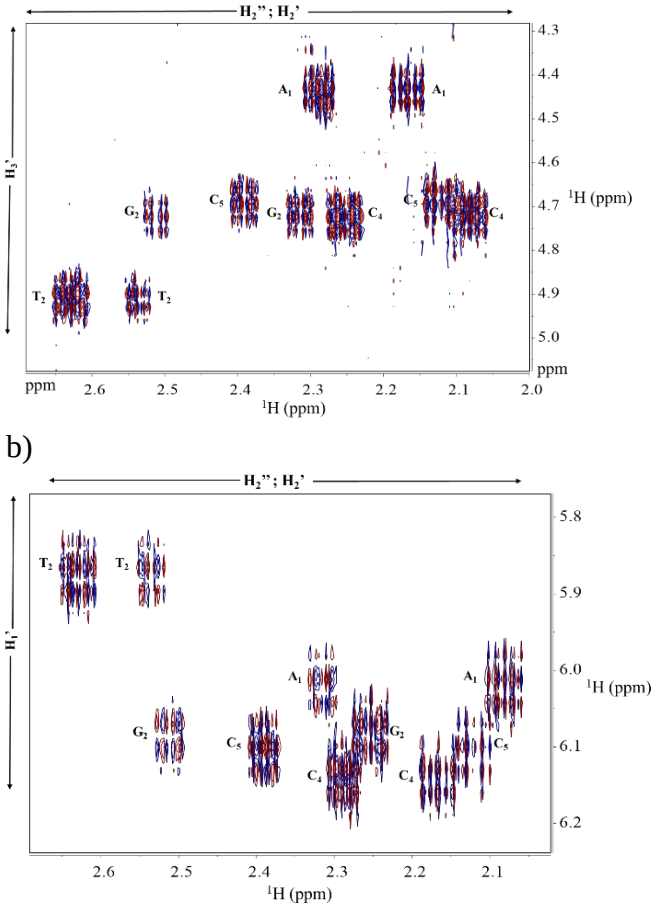


Figure 3.8. Aliphatic region of COSY in 99.9 % D_2O at 22ºC. (a) Cross-peak is correlating to H_2'/H_2''-H_1' (b) cross-peaks is correlating to H_2'/H_2''-H_3'.

Assignment of the NOESY spectrum

2D NOESY experiment was employed in order to assign protons that are close to each other in space. In the analyses of single-stranded DNA 5mer 5'-AGTCC-3' connectivity between the base and the sugar ribose phosphate were observed. Cross-peaks from the aromatic base protons $H_6/H_5/H_8$ to the aliphatic sugar-ribose phosphate protons H_1'connects the base to appropriate sugar according to the sequence of the DNA Figure 3.9.

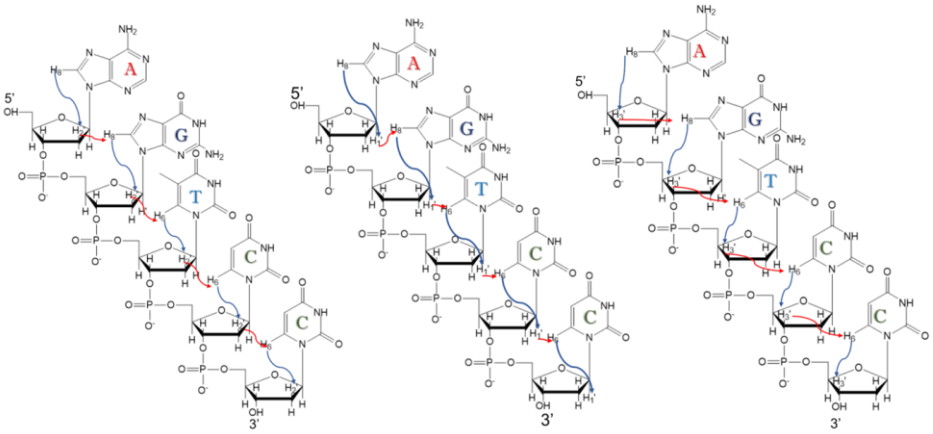


Figure 3.9. 700MHz ^{1}H-^{1}H NOESY spectrum (t_m=300 ms) of single-stranded DNA 5'-AGTCC-3' in 99.9 %. Correlation between aliphatic and aromatic protons are indicated in the boxes.

Assignment of the HSQC spectrum

To confirm the assignment of single-stranded DNA 5'-AGTCC-3', a 2D HSQC NMR experiment was obtained. The 1D-^{13}C and DEPT-135^{o} NMR spectra were first measured, the ^{13}C spectra shows 46 carbon peaks in the DNA spectrum. In DEPT 135^{o} spectrum 34 signals were seen, as the carbonyl and quaternary carbons do not appear. The CH_2 methylene carbons H_2'/H_2''and H_5'/H_5'' were inverted in the spectra while the methine carbon H_1', H_3', H_4', H_2, H_5, H_6, and H_8 as well as the methyl carbon were in phase. After assigning the DNA carbon resonance, it is very important to identify which carbons are attached directly to which protons in the DNA.

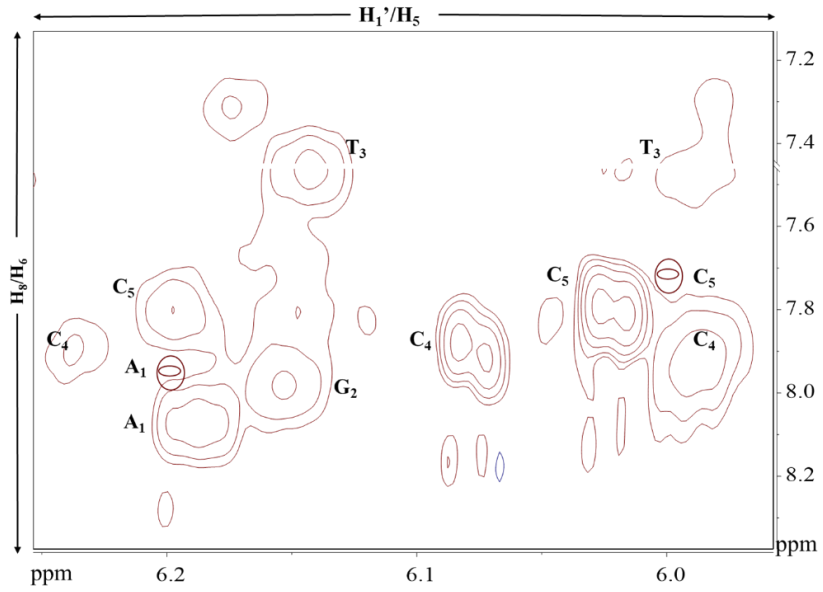


Figure 3.10. The area of the aromatic region of NOESY spectrum shown in Figure 3.9 above. Showing the sequential walk from A_1-G_2-T_3-C_4-C_5, which confirms DNA sequence.

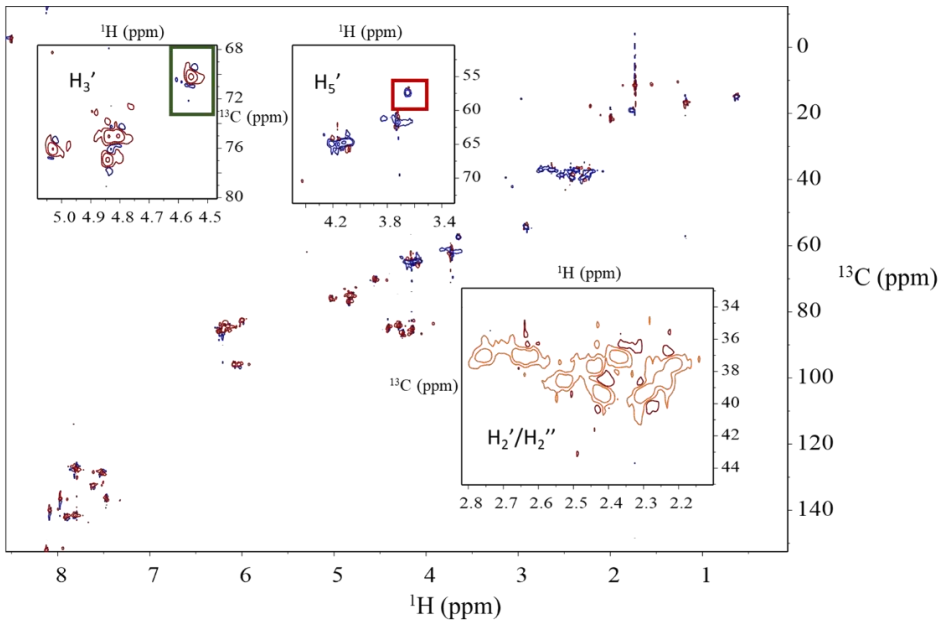


Figure 3.11. 700MHz ^{1}H-^{13}C HSQC spectrum of single-stranded DNA 5'-AGTCC-3' in 99.9 % D_2O at 22 ^{o}C. The F_2(vertical) dimension is the proton axis, and the F_1(horizontal) dimension is the carbon axis. The spectrum has been zoomed to show correlations between carbon and H_3', H_5', H_2'/H_2'', with emphasis on the A_1 5'end (red) and C_5 3'end (green) boxes.

Using a 2D HSQC ^{13}C-^{1}H NMR experiment, the assignment for carbon chemical shifts that has one bond (^{1}J) coupling to protons were identified. This experiment also allows the confirmation of both carbonyl and quaternary carbons, as they did not show any cross peaks for HSQC. The confirmation of geminal (H_2'/H_2'', H_5'/H_5'') protons in the sugar phosphate were analyzed, as the protons show the different chemical environments in the F_2 dimension and connects to the same carbon chemical shift in the F_1 dimension Figure 3.11. The spectra also shows different carbon chemical environments for both H_5 and H_6 as was previously observed from COSY spectra Figure 3.11. The 5'end of A_1 and 3' for C_5 also shows cross peaks with the protons. Finally, methyl carbon shows the lowest shift at 11.77 ppm with an H_6 aromatic proton.

Table 3.1. ^{1}H-NMR chemical shift (ppm) assignment of single-stranded DNA 5mer 5'-AGTCC-3'. Values were measured at 25 ^{o}C in 99.9% D2O, using NOESY and COSY cross peaks as described above. Standard error was calculated to be ±0.01- 0.1 ppm in all cases.

Bases	**H_1'**	**H_2'**	**H_2"**	**H_3'**	**H_4'**	**H_5'/H_5"**	**H_6/H_8**	**H_2/H_5/ CH_3**
A_1	6.19	2.62	2.37	4.84	4.28	3.72	8.11	7.97
G_2	6.14	2.43	2.22	4.82	4.25	4.13	8.08	-
T_3	5.99	2.64	2.75	5.03	4.41	4.15	7.46	1.70
C_4	6.20	2.26	2.56	4.41	4.30	4.09	7.90	6.08
C_5	6.24	2.29	2.42	4.30	4.20	4.12	7.81	6.02

Table 3.2. ^{1}H coupling constants (Hz) of the sugar-phosphate backbone and cytosine protons residues of single-stranded DNA 5mer 5'-AGTCC-3'. Values were measured at 25 ^{o}C in 99.9% D_2O. Standard error was calculated to be $\pm$0.1- 0.2 Hz in all cases.

Bases	$^3J_{1'-2'}$	$^3J_{1'-2''}$	$^3J_{2'-2''}$	$^3J_{2'-3'}$	$^3J_{2''-3'}$	$^3J_{3'-4'}$	$^3J_{3'-P}$	$^3J_{4'-5'}$	$^3J_{4'-5''}$	$^2J_{5'-5''}$	$^3J_{5-6}$
A_1	5.89	8.57	14.13	2.38	5.38	2.00	5.17	2.80	3.60	14.01	-
G_2	5.99	7.69	14.12	3.00	6.40	2.55	4.27	2.45	4.00	12.50	-
T_3	5.38	9.07	13.66	1.91	5.65	2.39	4.78	2.89	3.00	14.03	-
C_4	5.62	6.60	14.22	1.60	5.62	2.60	5.36	3.50	3.01	12.50	7.54
C_5	6.56	6.43	14.07	2.26	6.80	2.64	-	2.55	3.10	13.00	7.58

3.3.2. NMR Assignment of the Complementary DNA 5mer 5'-GGACT-3' strand

The NMR assignment of complementary DNA 5mer 5'-AGTCC-3' strand was also carried out using the experimental schemes mentioned earlier in sections 3.2.2 and 3.2.3 respectively. The sequence of the DNA bases in (section 3.2.1) is entirely different from the complementary DNA, thus it is important to assign them. Figure 3.12.

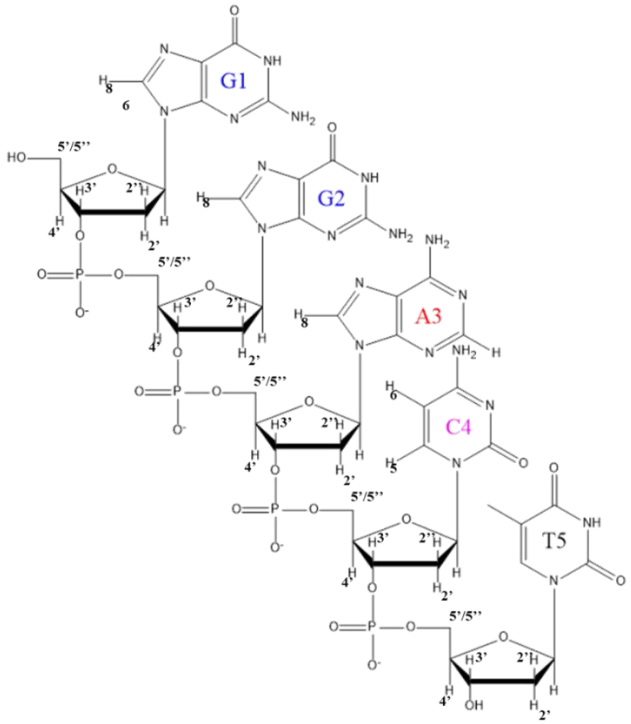


Figure 3.12. The chemical diagram of complementary DNA 5'-GGACT- 3' strand, showing the protons in their different chemical environments.

Using the same experimental procedures in section 3.3.1, the full assignments of the complementary DNA were assigned. In both DNAs, the same splitting patterns were mostly seen in the ^{1}H NMR spectrum Figure 3.13, but with slight changes in their chemical shifts which may be due to the rearrangement of the DNA sequence. The major difference in both DNAs was observed in the ^{31}P NMR spectrum which will be discussed later in this chapter.

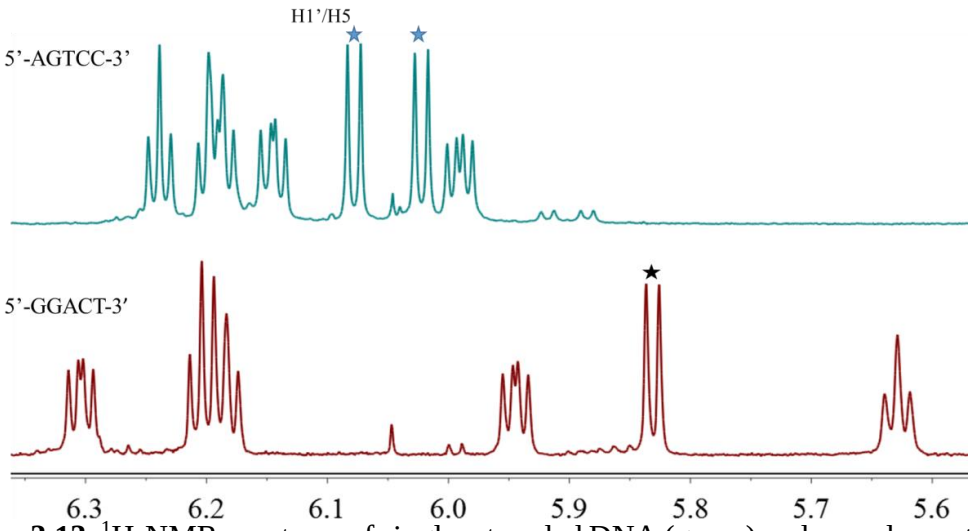


Figure 3.13. ^{1}H-NMR spectrum of single- stranded DNA (green) and complementary DNA (red), measured at 45 ^{o}C in 99.9 % D_2O. Showing the similar splitting patterns in H_1'/H_5 region, and confirming two cytosine H_5 (blue stars) signals observed in the single- stranded sequence and one cytosine H_5 signal (black star) in the complementary DNA.

Table 3.3. ^{1}H-NMR chemical shift (ppm) assignment of complementary DNA 5'-GGACT-3' strand. Values were measured at 45 ^{o}C in 99.9 % D_2O, using the same procedure discussed in section (3.2.2) above. Standard error was estimated to be $\pm$0.01- 0.1 ppm in all cases.

Bases	**H_1'**	**H_2'**	**H_2"**	**H_3'**	**H_4'**	**H_5'/H_5"**	**H_6/H_8**	**H_2/H_5/ CH_3**
G_1	6.17	2.23	2.52	4.80	4.18	3.69	8.00	-
G_2	6.19	2.22	2.51	4.92	4.92	4.03	8.06	-
A_3	6.30	2.38	2.74	4.99	4.99	4.05	8.28	7.83
C_4	5.94	2.69	2.28	4.75	4.75	4.07	7.69	5.83
T_5	5.63	2.38	2.50	4.42	4.42	4.16	7.60	1.83

3.3.3. NMR Assignment of the double-stranded DNA 5mer 5'-AGTCC-3'

Having shown the assignment of both single stranded and complementary DNA strand, further work was done on the double-stranded DNA 5'-AGTTC-3'. Double-stranded DNA 5'-AGTTC-3' was prepared in 1:1 molar ratio of both single-stranded and its complementary DNA strand following the procedure discussed in section (3.2.1). In this case, a total number of ten bases are present, where each purine (C, T) is bonded to pyrimidines (A, G) Figure 3.13.

Figure 3.14. Schematic diagram of double-stranded DNA 5'-AGTCC-3', showing the complementary DNA strand(blue) DNA and the connectivity of the bases (black).

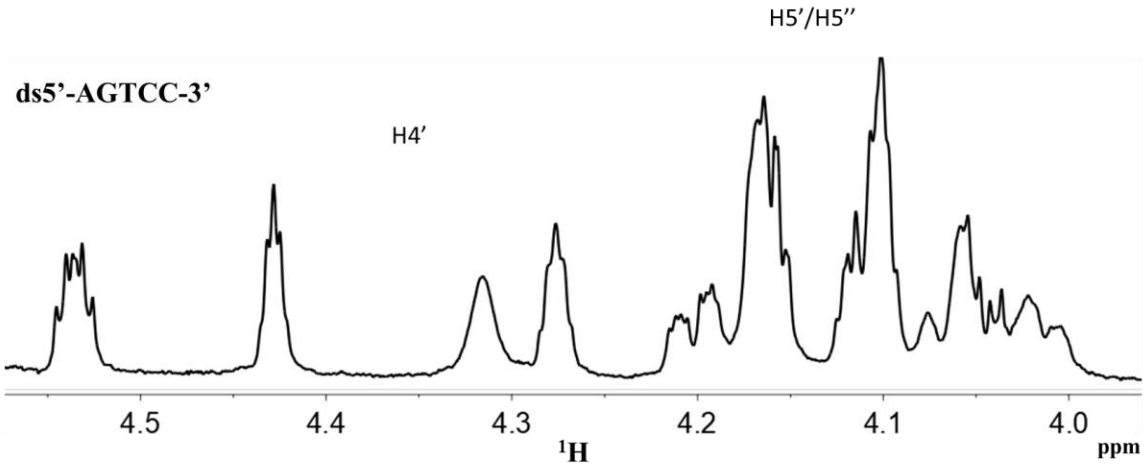


Figure 3.15. ^{1}H-NMR spectrum of double- stranded DNA in 99.9 % D_2O. Showing the H5'/H5'' region.

Table 3.4. ^{1}H-NMR chemical shift (ppm) assignment of double-stranded DNA 5'-AGTCC-3' strand. Values were measured at 35 ^{o}C in 99.9 % D_2O, using the same procedure discussed in section (3.2.2) above. Standard error was estimated to be $\pm$ 0.01-0.1 ppm in all cases.

Bases	H_1'	H_2'	H_2"	H_3'	H_4'	H_5'/H_5"	H_6/H_8	H_2/H_5/ CH_3
A_1	6.17	2.21	2.40	4.83	4.24	3.72	8.11	8.08
G_2	6.22	2.33	2.55	4.76	4.20	4.10	8.06	-
T_3	6.01	2.67	2.80	5.08	4.40	4.14	7.44	1.73
C_4	6.14	2.25	2.49	4.53	4.17	4.12	7.84	5.98
C_5	6.15	2.34	2.51	4.85	4.24	4.04	7.72	5.74
T_5	5.89	2.27	2.45	4.83	4.27	4.07	7.60	1.85
C_4	5.56	2.26	2.57	4.81	4.18	4.09	7.87	6.04
A_3	6.29	2.29	2.79	5.04	4.41	4.14	8.27	8.09
G_2	6.25	2.60	2.82	5.08	4.43	4.18	7.99	-
G_1	6.21	2.35	2.54	4.94	4.30	4.02	7.94	-

3.3.4. ^{31}P NMR Assignment of single-stranded, complementary and Double-stranded DNA

Although, ^{1}H and ^{13}C-NMR experiments have provided insight into the characterisation of DNA which is of great importance to know the proton and carbon chemical environment of the DNA. However, to study the backbone and the conformational changes in the DNA, ^{31}P NMR spectroscopy which is very sensitive to these changes was observed. A complete assignment of the ^{31}P-NMR of single-stranded DNA 5'-AGTCC-3', complementary DNA 5'-GGACT-3', and double-stranded DNA 5'-AGTCC-3' was investigated using 1D and 2D ^{1}H-^{31}P HSQC/HMQC experiments. All the assignments were carried out at 35 ^{o}C in 99. 9% D_2O.

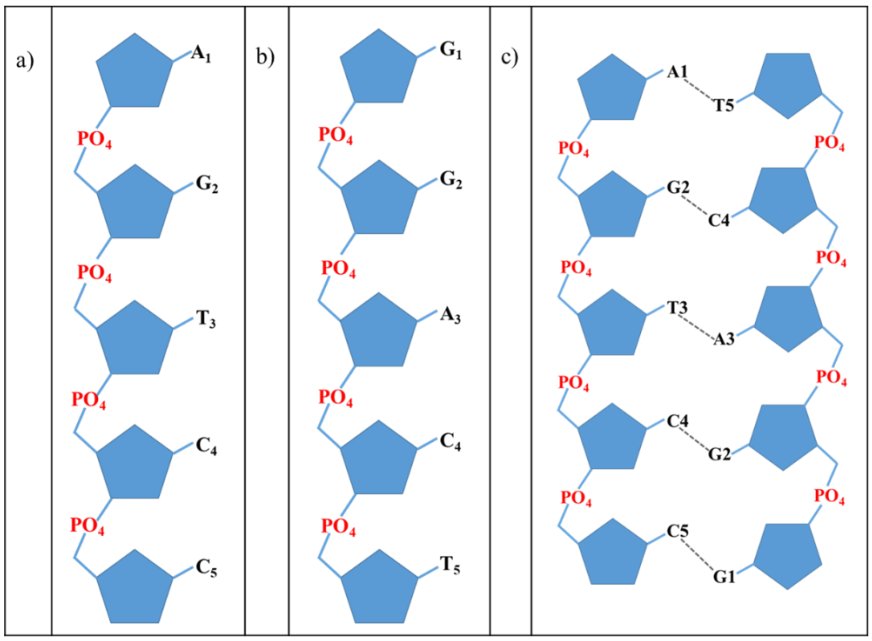


Figure 3.16. Schematic diagram showing the different chemical environment of phosphorus in the (a) Single-stranded DNA (b) Complementary DNA strand, and (c) Double-stranded DNA.

The 1D ^{31}P-NMR spectra show all the phosphate resonances in the DNA molecule. It also shows that the trend in DNA phosphate backbone corresponding to H_3'-O_3'-P-O_5'-H_5' throughout the sequence. More so, in the single-stranded and complementary DNA strand spectra; only four phosphate resonances were seen, which confirms that at the backbone, 5'and 3'ends have no phosphate attached to them. In the double-stranded DNA, although overlap of phosphate resonances were observed, integration gave eight phosphate resonances, which is expected for ten base sequence. ^{31}P chemical shift variation of -0.6 ppm were observed for all the phosphate signals of the DNAs.

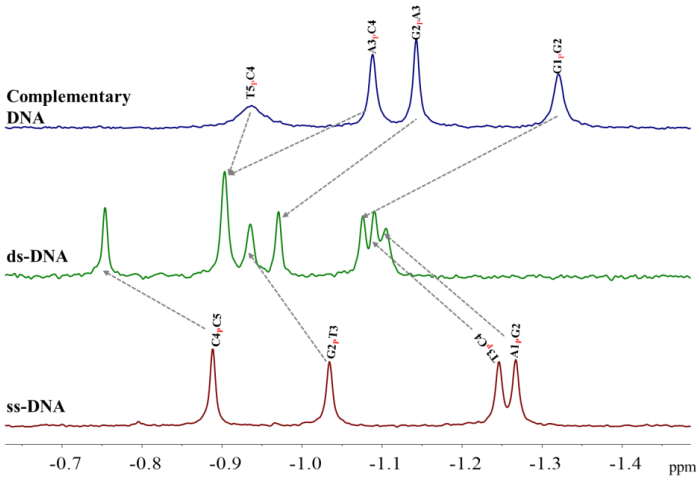


Figure 3.17. ^{31}P 1D-NMR experiment shows the stacked plot of (red) single-stranded DNA, (green) double-stranded DNA, (blue) complementary DNA strand. All experiments were measured at 35 ^{o}C in 99.9 % D_2O Arrows indicates changes in the ^{31}P resonance of ss-DNA and complementary DNA in the double-stranded DNA.

In the single stranded and complementary DNA strand as shown in Figure 3.15 above, it can be observed that the phosphate groups of the bases that are located at the beginning of the sequence ($G_{1P}G_2$ and $A_{1P}G_2$) have the lowest chemical shift of (-1.13 and -1.10 ppm) respectively. The observed chemical shift at (-1.13 and -1.10 ppm) is because phosphate group rotates freely which makes them more flexible, and it also has only one substituent as opposed to two. The bases located in the centre of the sequence ($T_{3p}C_4$, $G_{2P}T_3$, $G_{2P}A_3$, $A_{3P}C_4$) are seen at higher chemical shifts due to less motion, which is from the other sugar-phosphate group attached to them. However, the ($T_{5P}C_4$ and $C_{4P}C_5$) bases positioned at the end of the sequence have the highest chemical shift (-0.77 and -1.25 ppm) that shows even more rigidity at that region. In the double-stranded DNA spectrum, the phosphate resonance of the sugar backbone shows at the same region with a slight change of chemical shift as indicated by the arrow shown in Figure 3.17.

To further confirm the assignments, a 2D ^{31}P-^{1}H HMQC NMR experiment was performed at 35 ^{o}C in 99.9 % D_2O. In this experiment, short and long range heteronuclear

couplings were observed between ^{31}P and H_3'/H_4'/H_5'/H_5''. As shown in Figure 3.16 below, ^{31}P- H_3', ^{31}P-H_4' and ^{31}P-H_5'/H_5'' have very strong cross-peaks with each other.

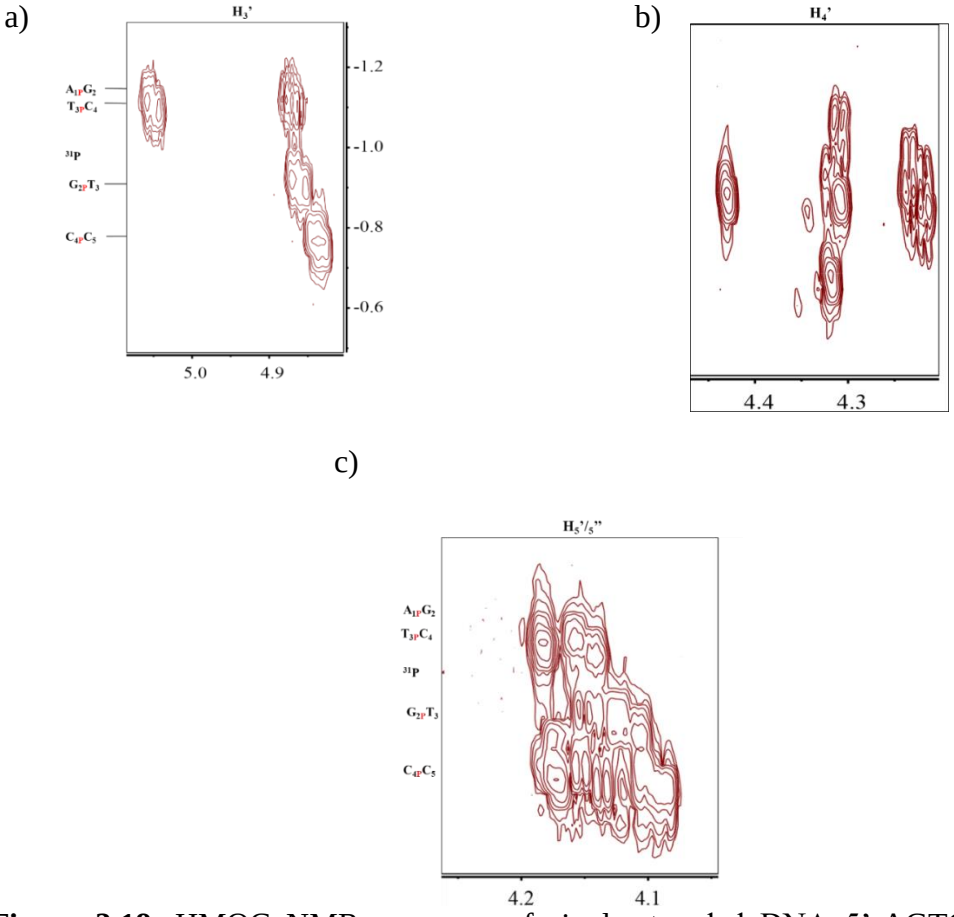


Figure 3.18. HMQC NMR spectrum of single-stranded DNA 5'-AGTCC-3'at 35 ^{o}C showing the DNA backbone ^{31}P resonances cross-peak with a) H_3' b) H_4' c) H_5'/H_5''.

Furthermore, in both Figure 3.18 (a) and Figure 3.18 (b), four and eight different cross-peaks observed confirm the assignment of the single-stranded DNA having only four phosphorus backbone, and eight phosphorus backbone in the double stranded DNA.

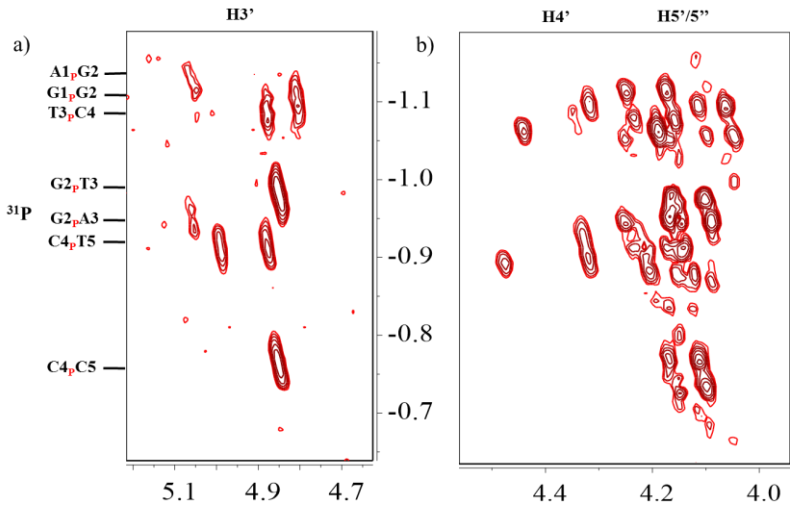


Figure 3.19. HSQC-HMQC NMR spectrum of double-stranded DNA 5'-AGTCC-3'at 35 ^{o}C showing the DNA backbone ^{31}P resonances cross-peak with a) H_3' b) H_4' and H_5'/H_5''.

Table 3.5. ^{13}P-NMR chemical shift (ppm) assignment of single-stranded DNA 5-AGTCC-3', complementary DNA strand and double-stranded DNA 5-AGTCC-3' strand. Values were measured at 35 ^{o}C in 99.9 % D_2O, using the same procedure discussed in section (3.2.2). Standard error was estimated to be $\pm$0.01- 0.1 ppm in all cases.

Bases	**A_1PG_2**	**T_3PC_4**	**G_1PG_2**	**G_2PA_3**	**G_2PT_3**	**A_3PC_4**	**T_3PC_4**	**C_4PC_5**
ssDNA	-1.27	-1.26	-1.02	-0.91				
Comple			-1.35	-1.15		-1.05	-0.89	
dsDNA	-1.12	-1.13	-1.14	-0.98	-0.95	-0.90	-0.91	-0.75

CHAPTER 4

4. THE ASSIGNMENT OF CPP AND ITS INTERACTION WITH DNA USING SOLUTION-STATE ^{1}H NMR SPECTROSCOPY

4.1. Overview

The interaction of cell-penetrating peptides (CPPs) with the lipid membrane of living cells has been known to alter the stability of the membrane bilayer.[3] CPPs achieve this by electrostatic interactions of positively charge peptides with the negatively charged nucleotide. However, questions still arise on the exact mechanism by which the peptide induces their activity, how they interact and distort the membrane bilayer to reach their target cell, and their size distribution. It is important to study the structure of the peptide and the DNA to understand the features of how they can carry on these activities. NMR spectroscopy is considered[71] the method of choice for studying the dynamic changes in DNA and peptide complexation.

In the preceding chapter, the full NMR assignment of pentamer single and double stranded DNA solution structure was assigned. In this chapter, the assignment of CPPs (R, R_3, R_9, Tat_2) and their interaction with pentamer ssDNA and dsDNA will be presented using solution-state 1D-^{1}H and 2-D HMBC NMR experiments. The results showed herein emphasize on the solution-structure of the peptide, and how the DNA interacts with different peptides. For this study, the NMR assignment of the peptides was assigned by *Samira Azimi* and will be briefly reviewed in this chapter.

4.2. Results and Discussion

4.2.1 NMR Assignment of R, R_3, R_9 and Tat_2 peptides

In general, these peptides have side chains that consists of the backbone, C-terminal, N-terminal, and the guanidinium head group. The guanidium head group has been proposed to function in hydrogen bonding interactions crucial to the structure and stability of peptide complexes.[72] These peptides also consist of α, β, γ, and δ carbons Figure 4.1. The guanidium head groups are positively charged at pH 7. Also, the amino group next to the α-carbon is positively charged, while the carboxylate group which is negatively charged at pH 7. Thus, at neutral pH the arginine has +1 charges. Equation 4.1.

$$[R]^{2+} \xrightarrow[pKa_1]{} [R]^{1+} \xrightarrow[pKa_2]{} [R]^{0} \qquad (4.1)$$

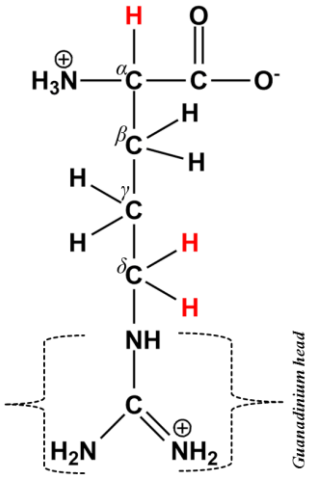


Figure 4.1. The chemical structure of arginine at pH 7, showing the protons in their different chemical environments, α-H and δ-H (red).

In this work, four different peptides were analysed; Arginine (R), Tri-arginine (R_3), Nano-arginine (R_9), and Tat_2 (RKKRRQRRRRKKRRQRRR). These CPPs were chosen due to their high number of arginine residues which was known to increase complexation.[73]

The experiments were carried out using the scheme discussed earlier in sections (3.2.1), (3.2.2), and (3.2.3) respectively.

1D and 2D NMR spectrum

The assignment of arginine, tri-arginine, nano-arginine, and Tat_2 was carried out using 1D ^{1}H NMR and 2D ^{13}C-^{1}H HMBC experiment at 22 ^{o}C, and pH ranging from 2 to10. The resonances corresponding to α-H and the δ-H were observed.

For the arginine molecule at pH 7, the signals corresponding to the α-H was seen at higher chemical shift of 3.47 ppm as compared to the δ-H at a lower chemical shift of 3.21 ppm. The difference in chemical shift is because of the inductive effect from both amine and carboxylate ion. However, the carbon connected to the δ-H is attached to only carbon and nitrogen atoms, thus making it less deshielded than the α-H and therefore found at a lower chemical shift Figure 4.2.

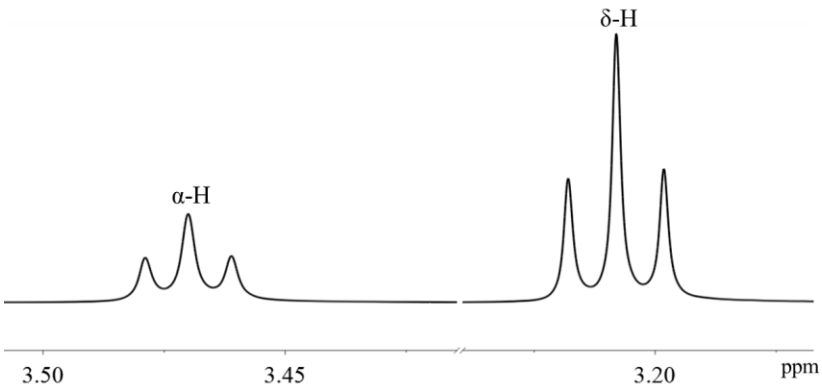


Figure 4.2. ^{1}H-NMR Arginine in acetate buffer at pH 7, measured in 99.9 % D_2O at 22 ^{o}C. The chemical structure of arginine is shown in Figure 4.1. Showing α-H, and δ-H region.

The ^{1}H-NMR experiment for the arginine molecule at pH ranging from 2 to10 at 22 ^{o}C was carried out. In the spectra, Figure 4.3 it can be observed that as the pH increases the chemical shift of the α-H decrease, this is because the carboxylate group close to the α-H

is deprotonated. It can be observed that from pH 5.2 to 9, the chemical shift almost stays the same until pH 10 where the guanadinium close to the amino group is deprotonated and thus the α-H move to lower shift.

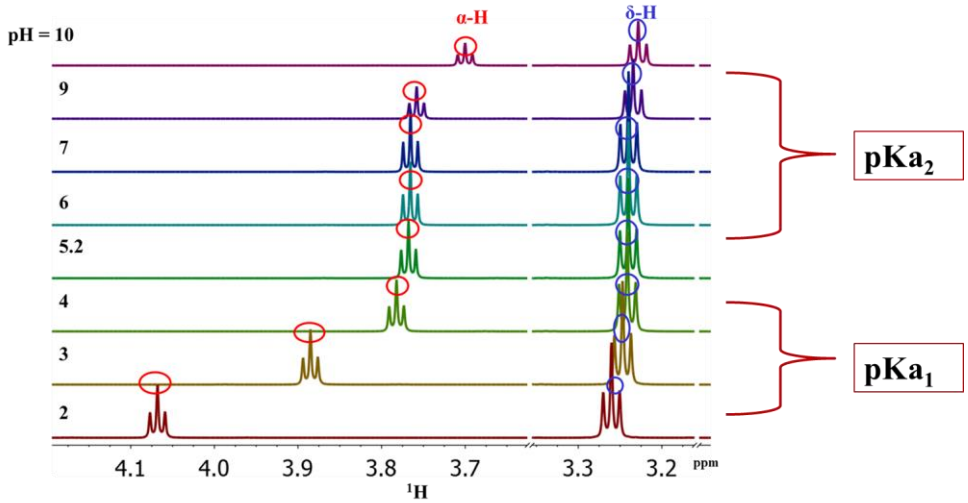


Figure 4.3. ^{1}H-NMR stacked plot of Arginine at pH 2-10, measured in 99.9 % D_2O at 22 ^{o}C. Highlighted signal shows the additional signal observe at δ-H region when dimerization occurred.

In the case of R_3, R_9, and Tat_2 peptides, there are four major sets of protons namely; α-H C-terminus (close to the carboxylate ion), α-H N-terminus (close to the ammonium ion), α-H backbone (center of the C and N terminus), guanadinium δ-H.

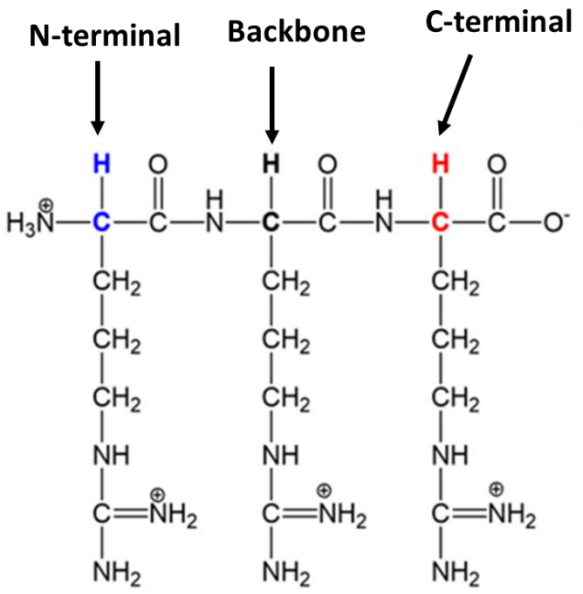


Figure 4.4. The chemical structure of tri-arginine at pH 7, showing the protons in their different chemical environments, α-H of C-terminus (bold red), N-terminus (bold-blue) and backbone (bold-black).

In R_3, R_9 and Tat_2 peptides, the α-H backbone has the highest chemical shift of 4.37 ppm as a result of the steric hindrances causes deshielding. The α-H of the C-terminus, N-terminus chemical shift are 4.25 and 4.11 ppm. The δ-H and γ-H have chemical shifts of 3.22 and 2.99 ppm Figure 4.5.

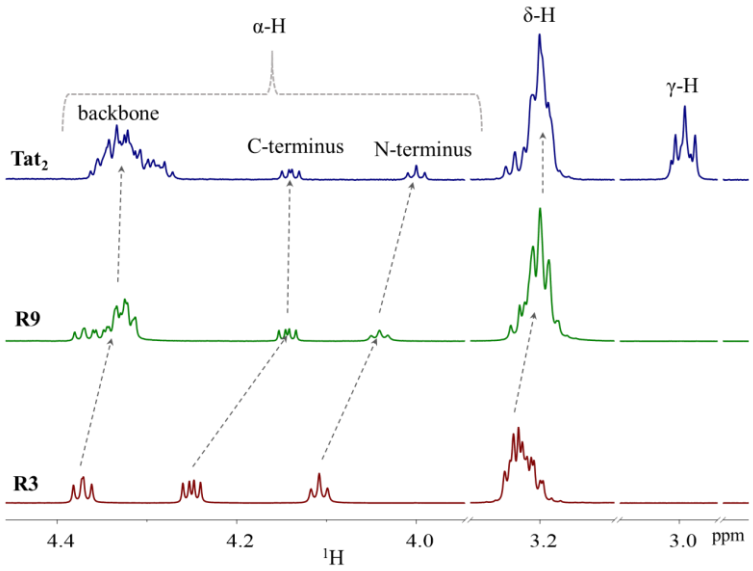


Figure 4.5. ^{1}H-NMR stacked spectrum of R_3, R_9 and Tat_2 peptide in acetate buffer at pH 7, measured in 99.9 % D_20 at 22 ^oC. Showing the α-H, δ-H and γ-H region.

The difference in the chemical shift of the α-H of N and C terminus could not only be assigned solely on their chemical environments. Therefore, to assign the protons correctly, HMBC experiment was carried out to analyze ^{13}C-^{1}H long and short-range between the carbonyl group and the α-H's at pH 6. The N-terminal carbonyl carbon (1) has a cross a peak with the α-H backbone, the α-H backbone also has a cross-peak with the backbone carbonyl carbon (2) Figure 4.6.

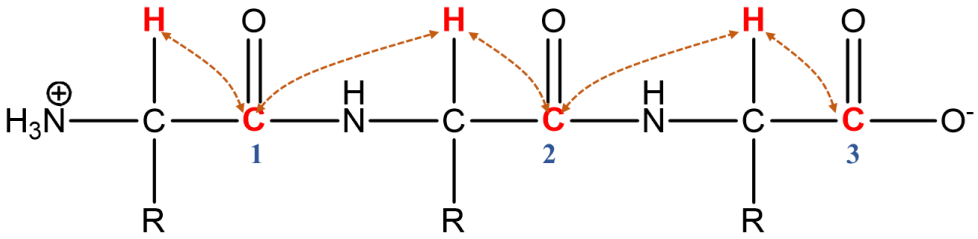


Figure 4.6. The schematic structure of tri-arginine peptide, showing the long-range coupling between the carbonyl carbon and the neighbouring connected protons (red).

The carbonyl (3) has a cross peak with α-H C-terminus, which also has a cross-peak with the amide backbone carbonyl carbon (2). However, α-H at the N-terminus only has a cross peak with N-terminal carbonyl carbon (1). These confirm that α-H at the N-terminus has a lower chemical shift at around 4.11 ppm, while the α-H at the C-terminus has a higher chemical shift at around 4.25 ppm.

From all the results stated above, it is no doubt that the pH dependence of the peptide maybe key to nanoparticle complex formation. For instance, it has been shown arginine can function as an acid or base in a number of enzymatic mechanisms.[74]

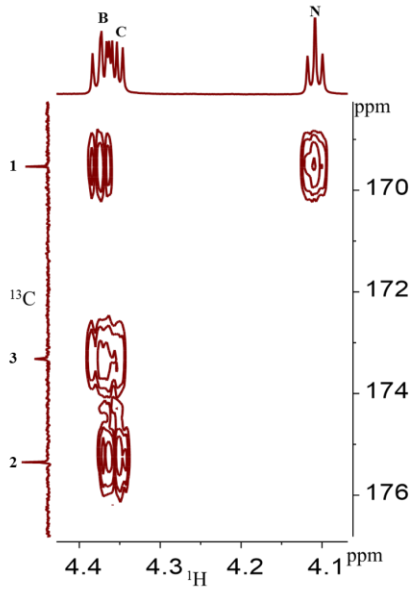


Figure 4.7. ^{13}C-^{1}H HMBC NMR spectrum of tri-arginine measured in 99 % D_2O at 22 ^{o}C. The spectrum was measured at pH 6. Cross-peaks shows long-range coupling between carbonyl carbon and proton.

4.2.2. NMR Assignment of Single-stranded DNA with Arginine Interaction

^{1}H and ^{31}P-NMR of ssDNA 5'AGTCC3' with Arginine

The ^{1}H and ^{31}P NMR of ssDNA 5'AGTCC3' with arginine was measured at pH 7, over a range of mole ratios. The experiment was carried out to observe the changes in chemical shift of peptides when it interacts with the DNA.

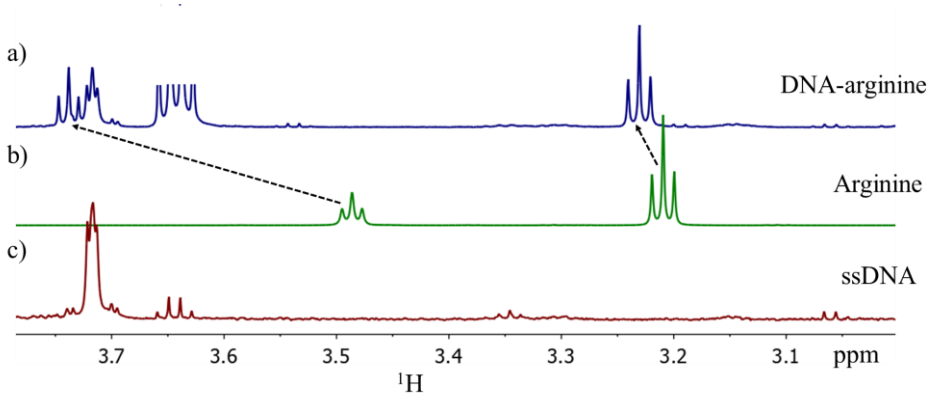


Figure 4.8. ^{1}H-NMR stacked plot of a) R- ssDNA b) free Arginine c) free ssDNA, in acetate buffer at pH 7, measured in 99.9 % D_2O at 22 ^{o}C. Arrows show the direction of change in chemical shift to α-H.

The spectra above in Figure 4.8, show that the chemical shift of the α-H in arginine increases when complexed with ssDNA at mole ratio of 1:1. To understand the change in the chemical shift of the α-H, spectra were obtained at Arginine: ssDNA mole ratios of 1:1, 2:1, 3:1, 6:1, and 10:1 at pH 7 as shown in Figure 4.9. In these spectra, only changes in the α-H and the δ-H were observed.

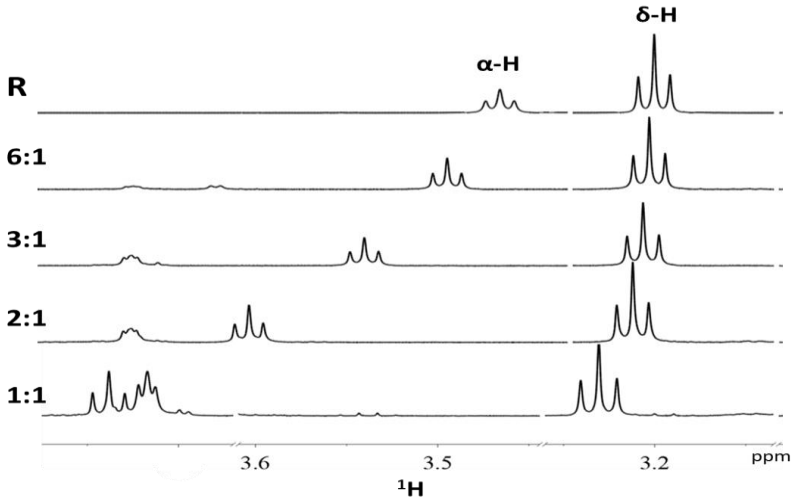


Figure 4.9. ^{1}H-NMR stacked plot of R-ssDNA in mole ratio ranging from 1:1 to 6:1 in acetate buffer at pH 7, measured in 99.9 % D_2O at 22 ºC.

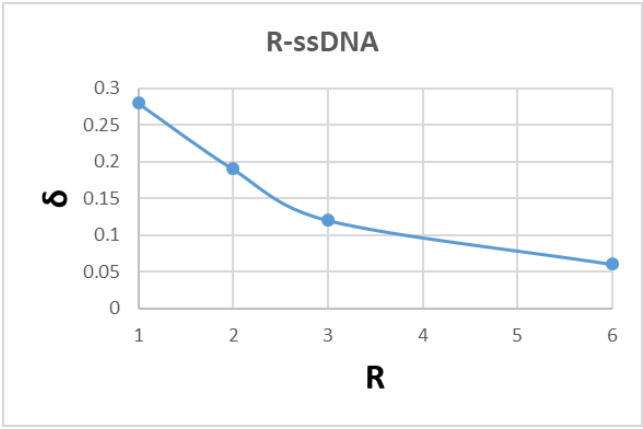


Figure 4.10. Experimental plot of the observed chemical shift against mole ratio.

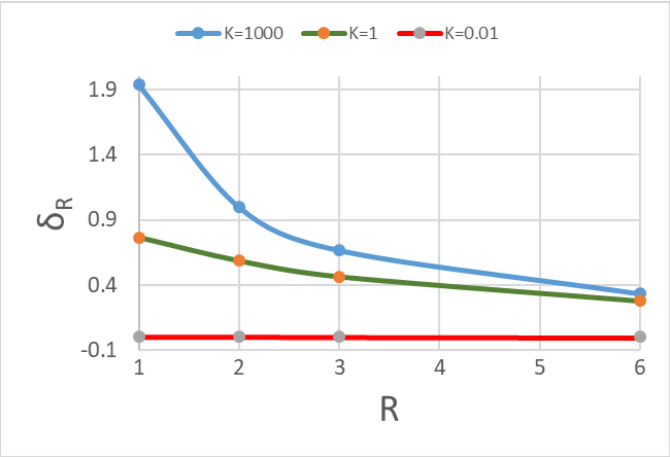


Figure 4.11. Theoretical plot of the relative chemical shift average against mole ratio.

The changes in the chemical shift of the α-H arginine compared to the complexed arginine at a mole of 1:1 is quite apparent Figure 4.9. These indicate a change in conformation of arginine upon binding to ssDNA. In Figure 4.11, the stoichiometry of R-ssDNA complexation was calculated from a quadratic equation derived from the equilibrium concentration, assuming a 1:1 molar concentration. The α-H signal in the NMR spectrum of R-ssDNA, and free arginine were observed. From the result Figure 4.09, the chemical shift observed was the average chemical shift of the free R and the complexed R amino-acid. These chemical shifts follow the same trend with the relative chemical shift in the theoretical Figure 4.10. These results assume a stoichiometry of 1:1 complexation of R-ssDNA, with a small equilibrium constant. The results also indicate that the binding of arginine to ssDNA may not be very strong, which results to excess arginine molecule upon increase in the concentration of arginine complex, and leads to decrease in the chemical shift observed.

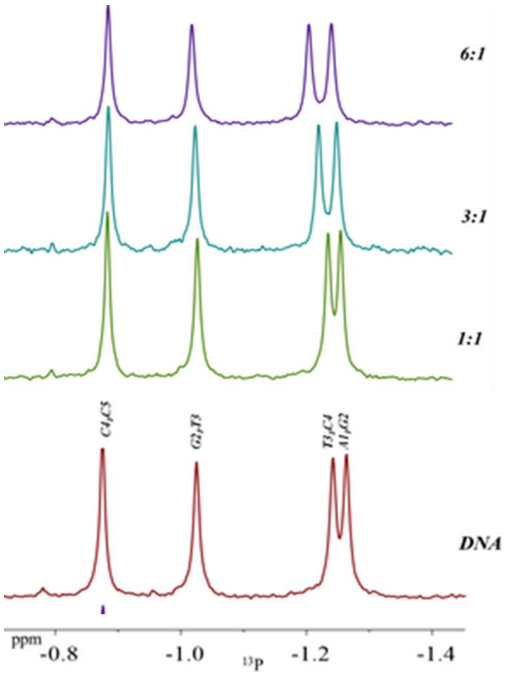


Figure 4.12. ^{31}P-NMR stacked plot of R-ssDNA at mole ratio of 1:1 to 6:1 in acetate buffer at pH 7. Showing the difference resonances of phosphorus, measured in 99.9 % D_2O at 22 ^{o}C.

^{31}P-NMR chemical shift gives information about the backbone conformation of the DNA upon binding.[77] The ^{31}P-NMR spectra show similar changes as shown previously in the ^{1}H-NMR spectra upon complexation with ssDNA. In the ^{31}P spectra, at 1:1 ratio we observe a decrease in the chemical shift of T_3pC_4, but the other signals were not affected nearly as much. However, as the concentration of arginine increases, the T_3pC_4 signal start to increase in chemical shift. These indicate that, at 1:1 ratio, the phosphate signal at T_3pC_4 have experienced a change in conformation upon binding with arginine. As the concentration of arginine increases, the conformation of the T_3pC_4 signal changed slightly.

4.2.3. NMR Assignment of Single-stranded DNA with Tri-arginine (R_3), Nano-arginine (R_9) and Tat_2 Interaction.

^{1}H and ^{31}P-NMR of ssDNA 5'-AGTCC-3' with R_3

The ^{1}H and ^{31}P NMR of ssDNA 5'AGTCC 3' with tri-arginine was measured at pH 7, and at various mole ratios of peptide-DNA. In the tri-arginine ^{1}H spectra Figure 4.13, it can be observed that at 1:1 mole ratio of complex R_3, the chemical shift of the N-terminus α-H decrease and slightly increases with increase in R_3 concentration. Similarly, the stoichiometry of R_3-ssDNA complexation was calculated from a quadratic equation derived from the equilibrium concentration, assuming a 1:1 molar concentration. The N-terminal proton signal in the NMR spectrum of R_3-ssDNA, and free R_3 were observed. In the result, Figure 4.14, the chemical shift observed was the average chemical shift of the free R_3 and the complexed R_3 peptide. These chemical shifts follow the same trend with the relative chemical shift in the theoretical Figure 4.15. These results assume a stoichiometry of 1:1 complexation of R_3-ssDNA, with a large equilibrium constant. These indicate that one molecule of R_3 binds with ssDNA. The C-terminus at the α-H and the backbone α-H chemical shift does not change, however, their linewidths became broader. These observations show that the DNA interaction with the peptide increases as the concentration of R_3 increase, as a result of the binding or conformational of R_3.

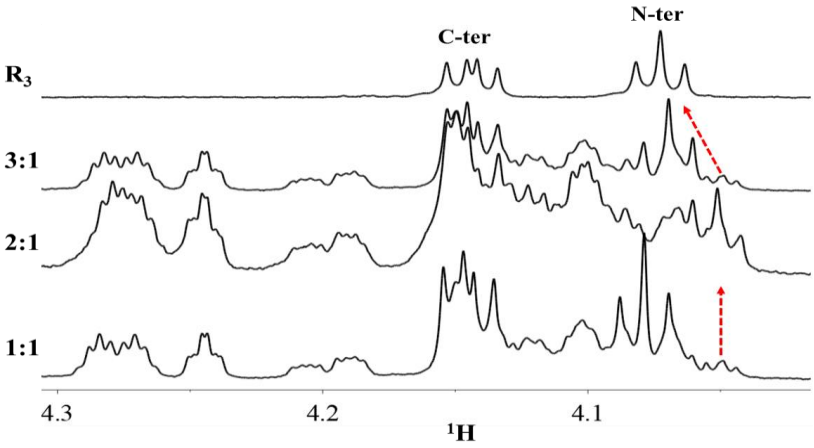


Figure 4.13. ^{1}H-NMR stacked plot of R_3-ssDNA at 1:1 to 3:1 mole ratio, in acetate buffer and at pH 7., measured in 99.9 % D_2O. Arrows show the direction of change in chemical shift corresponding N-terminus.

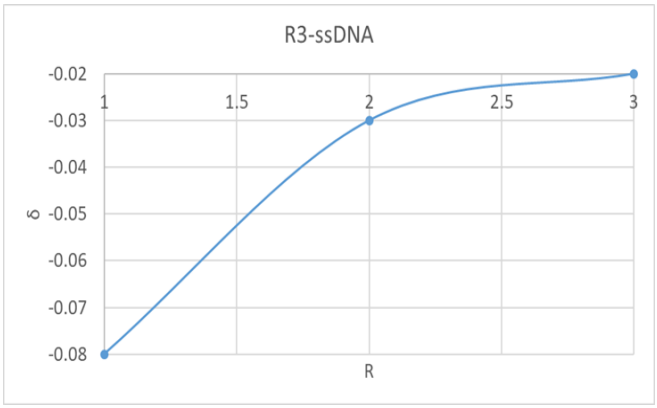


Figure 4.14. Experimental plot of the observed chemical shift against mole ratio.

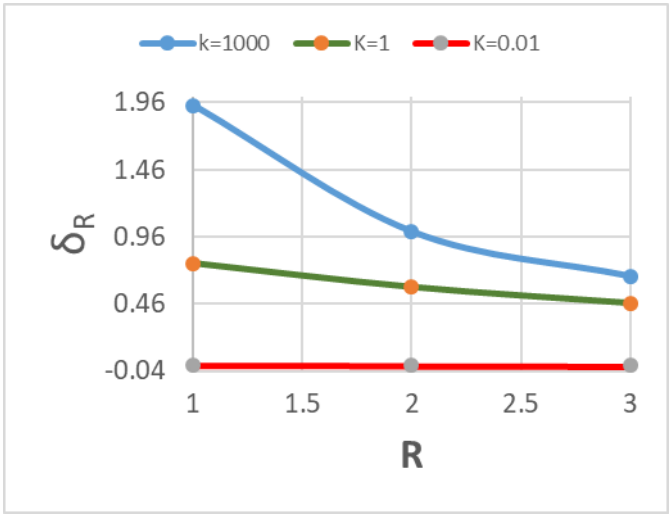


Figure 4.15. Theoretical plot of the relative chemical shift against mole ratio.

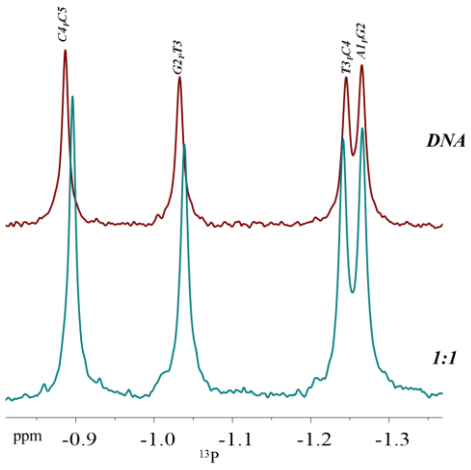


Figure 4.16. ^{31}P-NMR stacked plot of R_3-ssDNA at 1:1 ratio showing the difference resonances of phosphorus, measured in 99.9 % D_2O at 22 ^oC.

The ^{31}P-NMR spectrum at 25 ^oC was also measured for R_3-ssDNA at 1:1 molar ratio Figure 4.16. The phosphorus signals at the C_4pC_5, G_2pT_3 moved to a lower chemical shift, while the phosphorus signal of T_3pC_4 moved slightly to higher chemical shift and A_1pG_2 does not show a major change in chemical shift. These indicate that upon binding with the DNA, the conformation of C_4pC_5 and G_2pT_3 has been altered, while T_3pC_4 and A_1pG_2 conformational has not change. The above result also shows that the binding takes place at the centre of the DNA. These observations were also reported for spermine binding the Z-DNA of $d(ATGCAT)_2$ hexamer.[79] It was shown that at 1:1 mole ratio of spermine/DNA, the ^{31}P NMR chemical shift signal of the complexed moved to a lower chemical shift compared to the uncomplexed. A two-dimensional NOE studies on this complex also revealed a negative NOE contact between each spermine protons, indicating a longer correlation time τ_c (slow motion) upon binding with the DNA.

^{1}H and ^{31}P-NMR of ssDNA and dsDNA 5'-AGTCC-3' with R_9

The ^{1}H NMR of R_9-ssDNA at 1:1 ratio shows that as the concentration of R_9 increases the shift in the chemical shift of N-terminus decrease Figure 4.15. This observation was confirmed from the chemical shift of the N-terminus α-H signal at 2:1 ratio compared to 1:1 mole ratio. The signal of the N-terminus at 1:1 shows a lower chemical shift, while the C-terminus α-H does not change. These results show that as the concentration of the peptide increases the interaction between R_9 and ssDNA seems to increase, therefore more conformational changes should be experienced by the peptide. Similarly, the stoichiometry of R_9-ssDNA complexation was calculated from a cubic equation derived from equilibrium concentration, using the N-terminal proton signal in the NMR spectrum of R_9-ssDNA, and free R_9. The chemical shift observed Figure 4.18, was the average difference of the free R_9 and the complexed R_9 peptide. The chemical shift seems to follow the same trend with the relative chemical shift in the theoretical Figure 4.19. The close chemical shift range could possibly be an indication of weak equilibrium constant, assuming the complexation is a 2:1 stoichiometry.

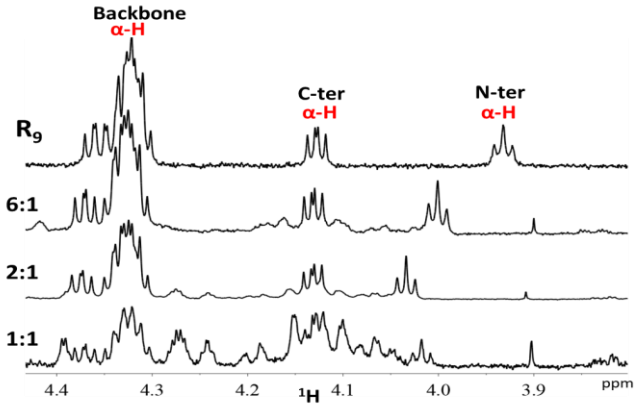


Figure 4.17. ^{1}H-NMR stacked plot of free R_9 and R_9-ssDNA at 1:1 - 6:1 mole ratio, measured in 99.9 % D_2O at 22 ^{o}C.

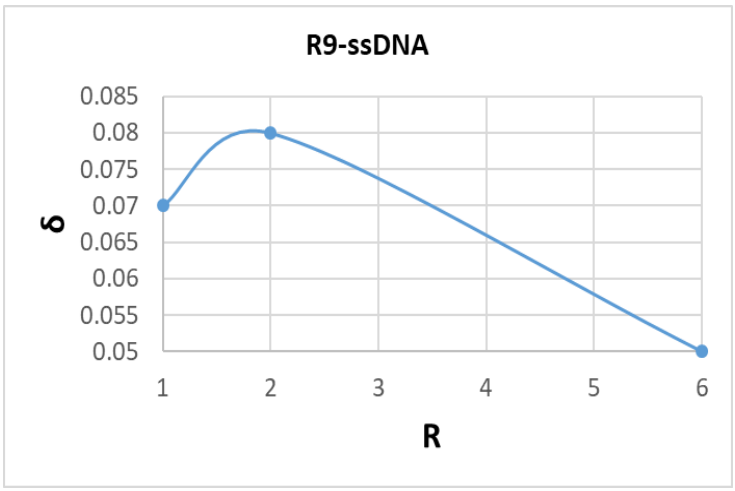


Figure 4.18. Experimental plot of the observed chemical shift against mole ratio.

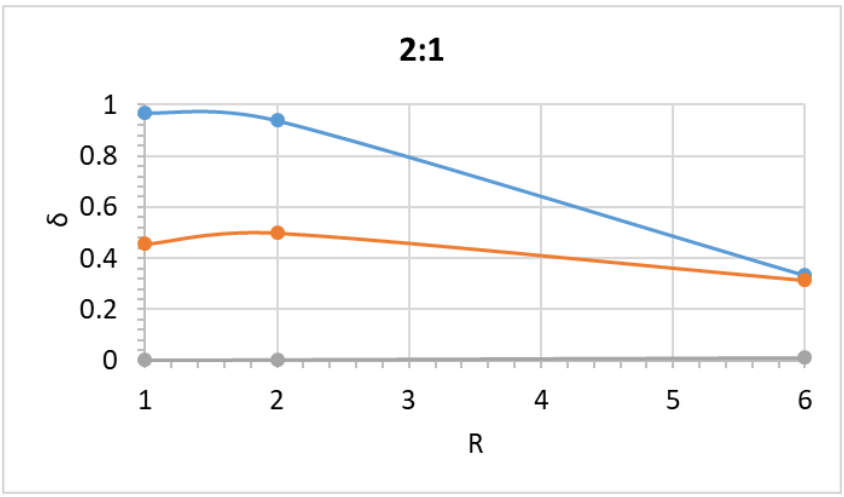


Figure 4.19. Theoretical plot of the relative chemical shift against mole ratio.

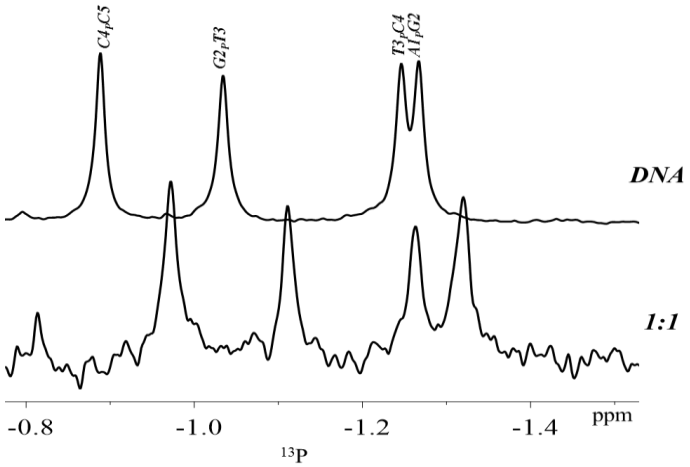


Figure 4.20. ^{31}P-NMR stacked plot of R_9-ssDNA at 1:1 mole ratio showing the difference resonances of phosphorus, measured in 99.9 % D_2O at 22 ^{o}C.

The ^{31}P NMR spectrum of R_9-ssDNA at 1:1 ratio also shows significant changes occurring to the backbone of the DNA upon complexation with R_9 Figure 4.14. As compared to the signals seen in R_3-ssDNA, the C_4pC_5 and G_2pT_3 of the R_9-ssDNA move further to a lower chemical shift. Also, the A_1pG_2 signal has also moved to a lower shift, while the T_3pC_4 signal has not changed as much. The ^{31}P indicate conformational changes to the DNA backbone upon binding to R_9 which lead to significant changes in the chemical shift.

^{1}H and ^{31}P-NMR of ssDNA and dsDNA 5'-AGTCC-3' with Tat_2

The interaction of DNA-peptide complexation was also observed with Tat_2-ssDNA and Tat_2 -dsDNA molecules at 1:1 ratio.

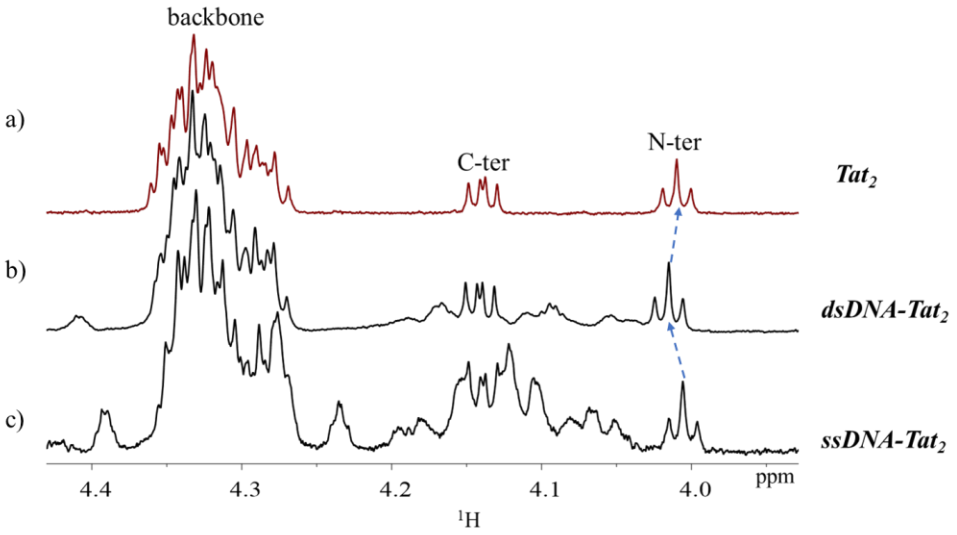


Figure 4.21. ^{1}H-NMR stacked plot of a) free Tat_2 b) Tat_2-dsDNA, 1:1 mole ratio c) Tat_2-ssDNA, 1:1 ratio, measured in 99.9 % D_2O at 22 ^{o}C. Arrow showed the direction of change in chemical shift corresponding C-terminus α-H.

The ^{1}H-NMR spectra of the ssDNA and dsDNA with Tat_2 mole ratio 1:1 spectra show an increase in the chemical shift of both ssDNA and dsDNA upon complexation with Tat_2. The dsDNA-Tat_2 complex N-terminus α-H signal is slightly higher in chemical shift than that of ssDNA-Tat_2 complex. These results were further confirmed from the ^{31}P-NMR spectroscopy of the complex. In the ^{31}P-NMR spectrum of the Tat_2-ssDNA in 1:1 Figure 4.17, the drastic change in the chemical shift of the phosphate signal and the broadening of the linewidth indicate a strong binding of Tat_2 with the phosphorus backbone of the DNA molecule. The phosphate signals of the DNA moved to a lower chemical shift upon complexation with Tat_2. Also, a comparison of the ^{31}P-NMR of the R_9-ssDNA and Tat_2-

ssDNA at 1:1 ratio was observed to see the how the DNA backbone behave with different peptides Figure 4.18.

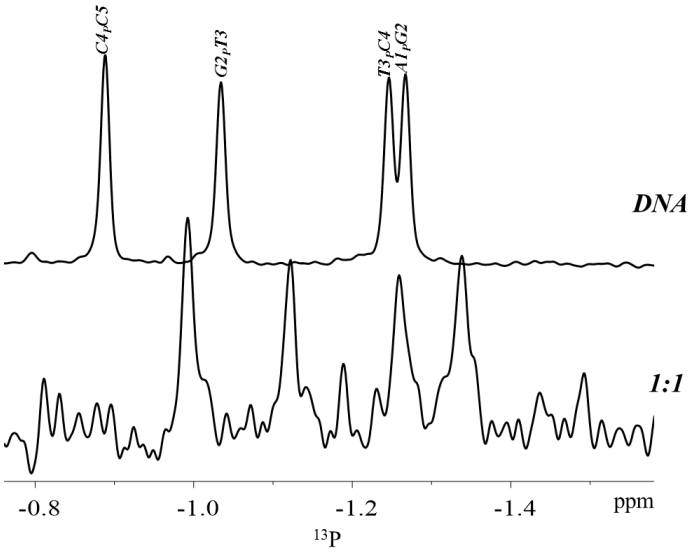


Figure 4.22. ^{31}P-NMR stacked plot of free DNA and Tat_2-ssDNA at 1:1 mole ratio showing the difference resonances of phosphorus, measured in 99.9 % D_2O at 22 ^oC.

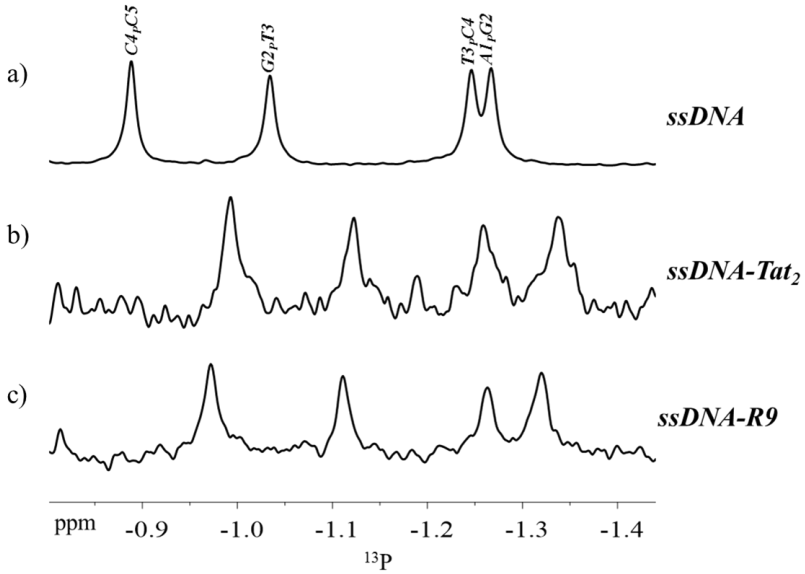


Figure 4.23. ^{13}P-NMR stacked plot of a) free ssDNA b) Tat_2-ssDNA, 1:1 mole ratio c) R_9-ssDNA, 1:1 mole ratio, measured in 99.9 % D_2O at 22 ^oC.

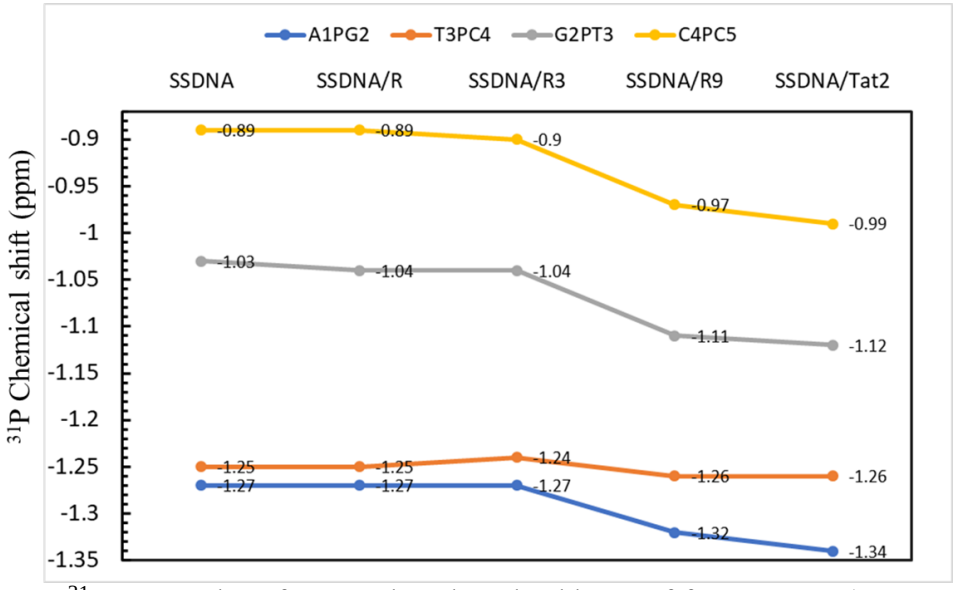


Figure 4.24. ^{31}P NMR plot of sugar phosphate backbone of free ssDNA (5'-AGTCC-3') and its interaction with R, R_3, R_9, and Tat_2, mole ratio 1:1 at pH 7.

In Figure 4.24, the ^{31}P chemical shift was plotted against the free ssDNA and the complexed R/ssDNA, R_3/ssDNA, R_9/ssDNA, and Tat_2/ssDNA. The chemical shift shows how the peptide interacts differently with the sugar phosphate backbone. The results show that in the R/ssDNA complex, only the G_2pT_3 sugar backbone is affected by the binding with a difference of -0.02 ppm compared to the free ssDNA while the other sugar backbone remains the same. In the R_3/ssDNA complex, the C_4pC_5 and G_2pT_3 backbone was also shifted while the other backbone remains the same. However, in the R_9/ssDNA and Tat_2/ssDNA complex, the phosphorus chemical shift has significantly shifted compared to the free ssDNA. Surprisingly, the T_3pC_4 phosphate backbone was not affected by the complexation. This implies that the peptides have no specific interaction with the T_3pC_4 backbone. This is unexpected because, if the negatively charged phosphate group was directly interacting with the positively charged N-terminus of the peptide, then the chemical shift of the phosphorus signal should decrease. This observation implies that the peptide might be binding to other regions in the DNA, thereby causing slight conformational

changes to the sugar phosphate backbone. This is contrary to the reported studies of spermine with Z-DNA complexation. It was observed that spermine binds with the phosphate backbone but no specific interaction with the base region.[79] Therefore it may be interesting to study at the base region of the DNA.

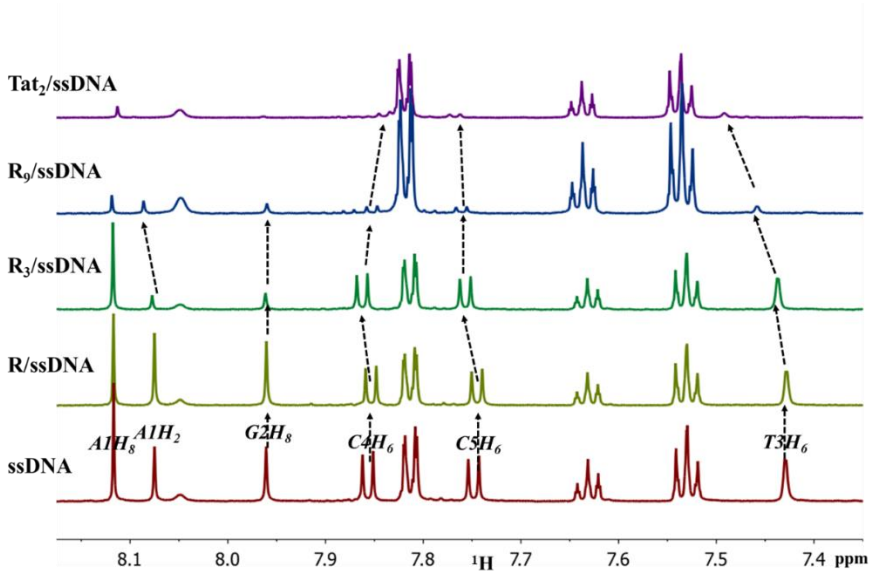


Figure 4.25. ^{1}H-NMR stacked plot of free ssDNA and complexed R/ssDNA, R_3/ssDNA, R_9/ssDNA, and Tat_2/ssDNA at 1:1 mole ratio, measured at pH 7. Arrows show the direction of change in chemical shift corresponding H_8, H_6, H_5 base protons.

In Figures 4.23 ,4.24, and 4.25, the ^{1}H NMR spectra of the free ssDNA with complexed R, R_3, R_9, and Tat_2 were studied. Changes in the chemical shift of the purine and pyrimidine base and sugar regions were observed upon complexation. In the R/ssDNA spectra, no changes were noticed in the chemical shift of the base protons H_8, H_6, H_5, and CH_3 and the sugar prorton H_1'. This implies that arginine has a weak interaction with all the bases in the DNA. The spectra of the R3/ssDNA however show a slight change (0.3-0.5ppm) in the chemical shift of the bases in the complexed ssDNA compared to the free ssDNA. This result suggests that although R_3 maybe interacting with the bases in the DNA, it does not neccesarily result in direct contact with the bases.

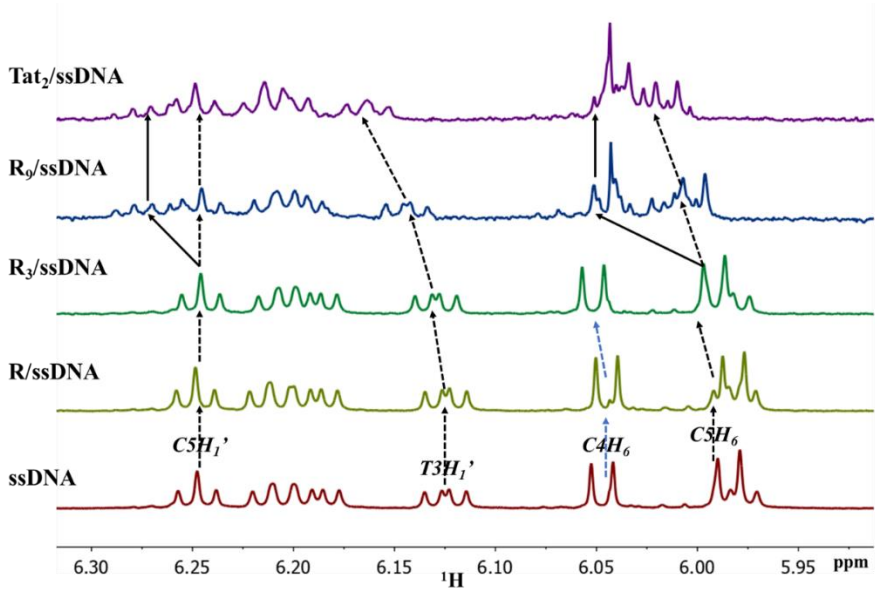


Figure 4.26. ^{1}H-NMR stacked plot of free ssDNA and complexed R/ssDNA, R_3/ssDNA, R_9/ssDNA, and Tat_2/ssDNA at 1:1 mole ratio, measured at pH 7. Arrows show the direction of change in chemical shift corresponding H_1' sugar and H_5 base protons.

In contrast to this, the R_9/ssDNA and Tat_2/ ssDNA spectra shows significant changes in the chemical shift of the complexed protons in the bases and the sugar backbone. Comparing the bound ssDNA to the free ssDNA in this case, it was observed that the linewidth of the C_4H_6, C_5H_6, T_3H_6, A_1H_8 and A_1H_2 resonances have decrease and broadened. More so, the sugar proton H_1' close to the bases for C_5H_1' have two comformations Figure 4.26 (indicated with the thick straight lines) resulting from an extra C_5H_1' signal. Similarly, the C_4H_6 and C_5H_6 resonances also shows the same comformation. These conformations together with the chemical shift changes can be attributed to a strong interaction of R_9 and Tat_2 with the DNA, which may be as a result of a hydrogen bonding from the DNA base binding with the side chains of the peptides. A study has shown that the interaction of pentalysine and T_7 DNA was driven by an electrostatic interaction between them[82]. The results show the effect of pH on the complexation of $(Lys)_5$-T7 DNA, similar behavoiur was also reported in oligolysine binding DNA[83] and in promine-DNA complexes[84]. The increase in the pH of the complex reduces the net charge on the lysine

due to the deprotonation of the ε and α-NH_2 groups, and therefore binding consant was reduced. It has also been reported that the weak binding of some repressor to DNA may play a major role in the process of locating an operator sequence by the repressor.[85]

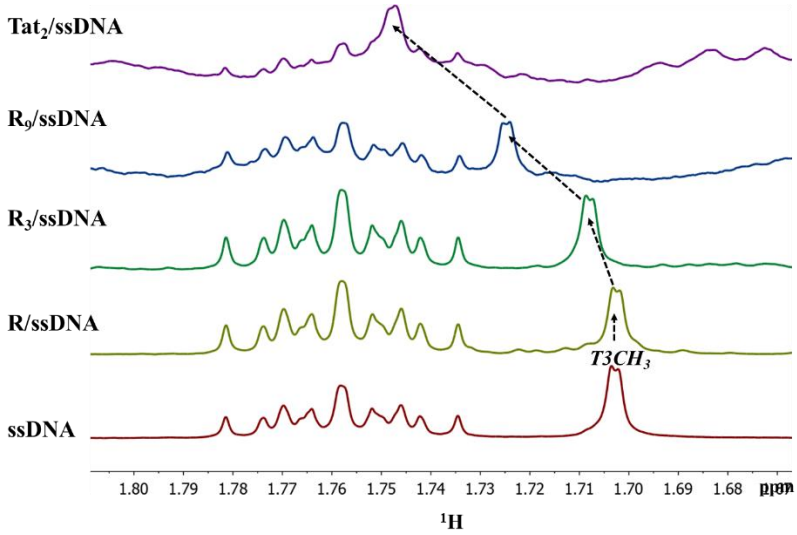


Figure 4.27. ^{1}H-NMR stacked plot of free ssDNA and complexed R/ssDNA, R_3/ssDNA, R_9/ssDNA, and Tat_2/ssDNA at 1:1 mole ratio, measured at pH 7. Arrows show the direction of change in chemical shift corresponding methyl base proton.

In Figure 4.26, 4.27, 4.28, and 4.29, the ^{1}H NMR spectra of the free dsDNA with complexed R, R_3, R_9, and Tat_2 were studied. Changes in the chemical shift of the purine and pyrimidine base and the sugar regions were observed upon complexation. The complexed R/ssDNA spectra shows no major changes in the chemical shift of the base protons H_8, H_6, H_5, and CH_3, and sugar H_1'proton as compared to the free dsDNA spectra. This observation maybe be due to the small size of arginine amino acid, rotating fast in solution with the dsDNA causing no mjor interaction to take place. In the R_3/ssDNA complex, a slight increase in the chemical shift of C_5H_6 and G_5H_8 were observed but no change in the C_5H_5 signal. This change suggests a sort of hydrogen bonding at the end of the dsDNA strand, where the side chains in R_3 is interacting with the base pair of C_5 and G_5.

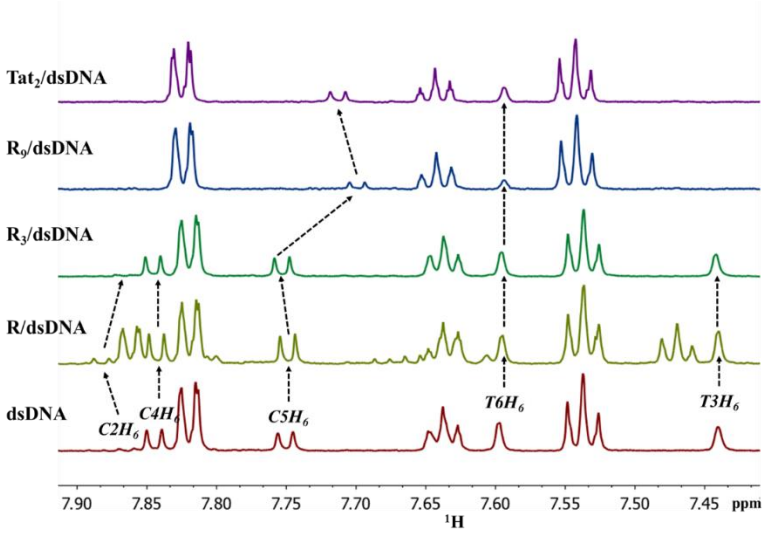


Figure 4.28. ^{1}H-NMR stacked plot of free dsDNA and complexed R/dsDNA, R_3/dsDNA, R_9/dsDNA, and Tat_2/dsDNA at 1:1 mole ratio, measured at pH 7. Arrows show the direction of change in chemical shift corresponding H_6 protons.

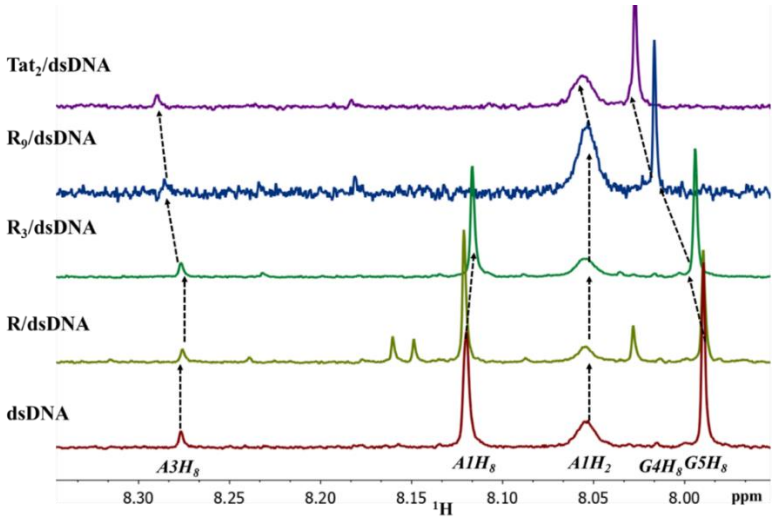


Figure 4.29. ^{1}H-NMR stacked plot of free ssDNA and complexed R/dsDNA, R_3/dsDNA, R_9/dsDNA, and Tat_2/dsDNA at 1:1 mole ratio, measured at pH 7. Arrows show the direction of change in chemical shift corresponding H_8, H_2 base proton.

Similar changes were observed for R_9/ssDNA and Tat_2/DNA in this case, the binding of R_9 and Tat_2 causes major and significant changes to the conformation of bases and sugar pucker except for the G_2H_8, C_2H_5, and C_2H_6 . The strong binding of the R_9 and Tat_2 shows a linewidth broadening at the T_3H_6, C_4H_1', T_3H_1', and C_4H_5 protons. As

compared to the ssDNA complex with R, R_3, R_9, and Tat_2, it is strongly evident that complexation of the peptide with the dsDNA is much stronger than with ssDNA. This maybe due to the larger size of the dsDNA 10 bases compared to ssDNA 5 bases, or the specific binding of the peptide to the dsDNA helix which results in more peptide binding.

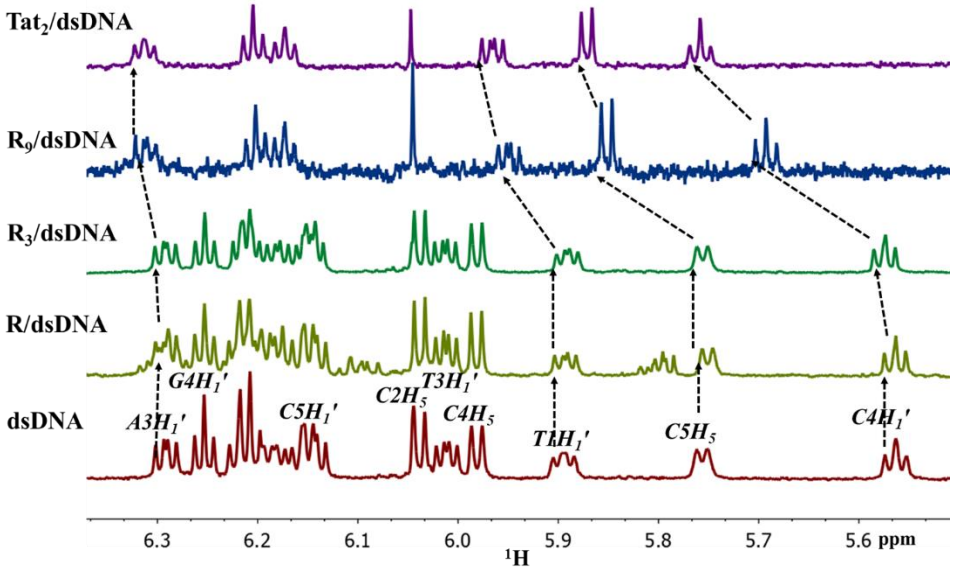


Figure 4.30. ^{1}H-NMR stacked plot of free ssDNA and complexed R/dsDNA, R_3/dsDNA, R_9/dsDNA, and Tat_2/dsDNA at 1:1 mole ratio, measured at pH 7. Arrows show the direction of change in chemical shift corresponding H_1' sugar and H_5 base proton.

The binding affinity of T4 coded gene 32 protein with oligonucleotide and polynucleotide was reported.[86-87] The results shows that the binding of the protien with oligonucelotide was independent of the base composition or sugar type, and binding took place at the end of the oligonucleotide and the DNA is not free to move at this site. However, the binding of the polynucleotide with the protein cause conformational changes that results in the pulling apart of the cluster of the postively charged residues on the protein. [88] From this study, it can be deduced that the size of the DNA plays a major role in the complex formation of peptide-DNA.

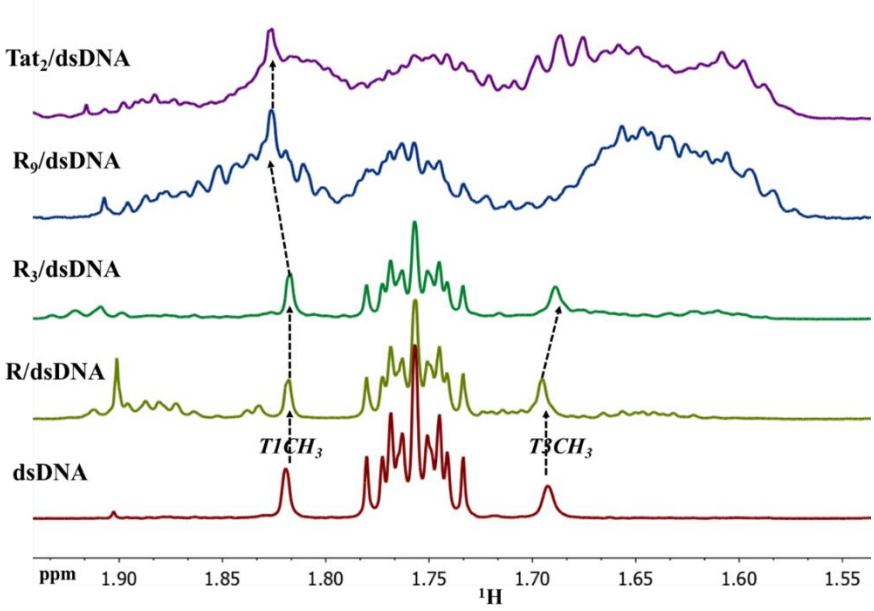


Figure 4.31. ^{1}H-NMR stacked plot of free ssDNA and complexed R/dsDNA, R_3/dsDNA, R_9/dsDNA, and Tat_2/dsDNA at 1:1 mole ratio, measured at pH 7. Arrows show the direction of change in chemical shift corresponding methyl base proton

CHAPTER 5

5. DYNAMIC STUDY OF CELL PENETRATING PEPTIDE AND DNA COMPLEXATION USING ^{1}H AND ^{31}P NMR RELAXATION MEASUREMENTS

5.1. Overview

In the preceding chapter, the behaviour of peptide-DNA complexation was observed using both 1D-^{1}H and ^{31}P NMR spectroscopy. The results showed changes in the chemical shifts and linewidth of resonances in specific regions of the DNA and peptide spectra. These findings were important for observing changes that occur within the chemical and physical environments both of DNA and peptide in their free and complexed states. However, only few major structural changes were observed from the analysis. Therefore, more experimental work is needed to observe the dynamics of complexation over a broad range of timescale at an atomic level.

In this chapter, we provide detailed ^{1}H and ^{31}P T_1 and T_2 relaxation analysis. Relaxation at specific sites in the DNA and peptide molecules will depend on the dynamic that drives the force of interaction between the DNA and peptide. The results acquired from the relaxation measurements will not only reflect on the different behaviour of the biomolecule, but rather the dynamic changes. In the DNA, the phosphorus backbone will be observed also the α-H of the C-terminus, N-terminus, backbone, and guanidinium regions in the peptides will also be observed.

5.2. Results and Discussion

5.2.1. Relaxation time measurements of free peptides at different temperatures

T_1 relaxation measurement of Tri-arginine (R_3) peptide

The longitudinal relaxation time constant (T_1) for the α-H backbone, C-terminus, N-terminus and δ-H (Figure 4.4) resonances in the ^{1}H NMR spectrum of tri-arginine (R_3) is shown in Table 5.1.

Table 5.1. Longitudinal relaxation (T_1) time measurement for the ^{1}H NMR peaks of un-complexed R_3 peptide at varying temperatures.

Regions	δ(ppm)	15°C	25°C	35°C	45°C
α-H Backbone	4.37	1.73	1.75	1.99	2.34
α-H C-terminus	4.25	1.66	1.77	2.00	2.51
α-H N-terminus	4.11	1.58	1.69	1.85	2.02
δ-H	3.22	0.68	0.76	0.85	0.92

Four regions in the ^{1}H spectrum of the peptide was carefully selected; the α-H of the backbone, C-terminus, N-terminus, and guanidinium δ-H. These selections were based on the possible role they play in the study of complexation with the DNA molecule. The four sets of longitudinal relaxation time constants are depicted in Table 5.1.

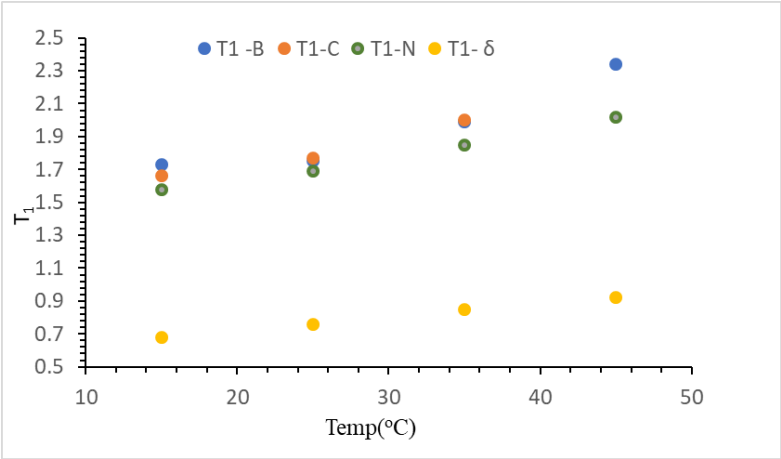


Figure 5.1. The stack plot of the T_1 against the temperature of the α-H backbone, C-terminus, N-terminal and guanadinium δ-H of the R_3 molecule.

From Figure 5.1 above, the four regions of the peptide T_1 relaxation time constants increase as the temperature increases, indicating faster motion within the peptide molecule. The δ-H T_1 relaxation time constant is shorter than the others, indicating higher conformational flexibility in the region as compared to elsewhere within the peptide.

T_1 relaxation measurement of Tri-arginine (R_9) peptide

The longitudinal relaxation time constant (T_1) for the backbone, C-terminus, N-terminus, and guanadinium δ-H peaks in the ^{1}H NMR spectrum of Nano-arginine (R_9) peptides

Table 5.2. Longitudinal relaxation (T_1) time measurement for the ^{1}H NMR peaks of free R_9 peptide at varying temperatures.

Regions	δ(ppm)	15°C	25°C	35°C	45°C
α-H Backbone	4.37	1.80	1.89	1.92	2.14
α-H C-terminus	4.25	2.84	2.54	2.42	2.39
α-H N-terminus	4.11	2.13	2.51	2.74	2.87
δ-H	3.22	0.82	0.77	0.75	0.62

The relaxation time constants T_1 observed in the R_9 peptide is quite different from the R_3 peptide. In this case, the T_1 time constant of the C-terminus α-H and δ-H decreases with increasing temperature, indicating reduced motion. However, in the α-H backbone and α-H N-terminus regions, the increase in T_1 relaxation time constants with increase in temperature indicates higher degree of motion Figure 5.2. This indicates reduced motion in the inner domain of the peptides.

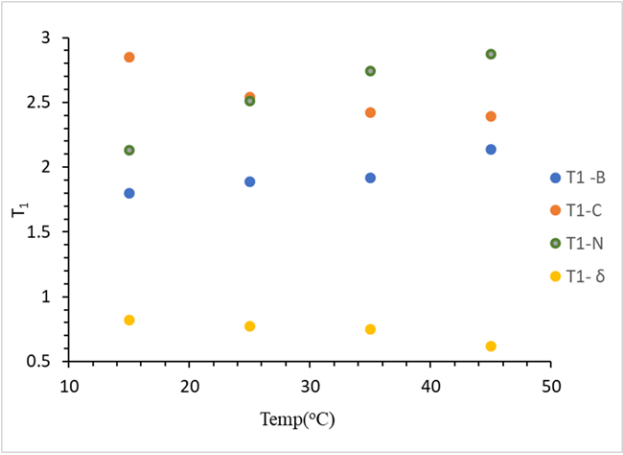


Figure 5.2. The stack plot of the T_1 against the temperature of the α-H backbone, C-terminus, N-terminal and δ-H of the R_9 molecule.

T_1 relaxation measurement of Tat_2 peptide

The longitudinal relaxation time constant (T_1) for the backbone, C-terminus, N-terminus, and δ-H peaks in the ^{1}H NMR spectrum of Nano-arginine (Tat_2) peptide.

Table 5.3. Longitudinal relaxation (T_1) time measurement for the ^{1}H NMR peaks of free Tat_2 peptide at varying temperatures.

Regions	δ(ppm)	15°C	25°C	35°C	45°C
α-H Backbone	4.37	2.26	2.38	2.47	2.81
α-H C-terminus	4.25	2.82	3.94	4.35	4.52
α-H N-terminus	4.11	2.35	2.54	3.56	4.10
δ-H	3.22	0.90	1.17	1.41	1.82

In the Tat_2 peptide, increasing temperature leads to increase in the T_1 relaxation time for all four regions, indicating high motion within the molecule. However, comparing the T_1 relaxation rate constants of R_3 and R_9 with Tat_2 peptide, it can be observed that Tat_2 peptide shows a longer T_1 relaxation rate constants as compared to R_3 and R_9 peptides Figure 5.3. In R_3 and R_9 peptides, the δ-H proton shows the shorter T_1 relaxation constant. This results show that s the peptide length becomes longer, the equilibrium constant become bigger, and the motion becomes rigid.

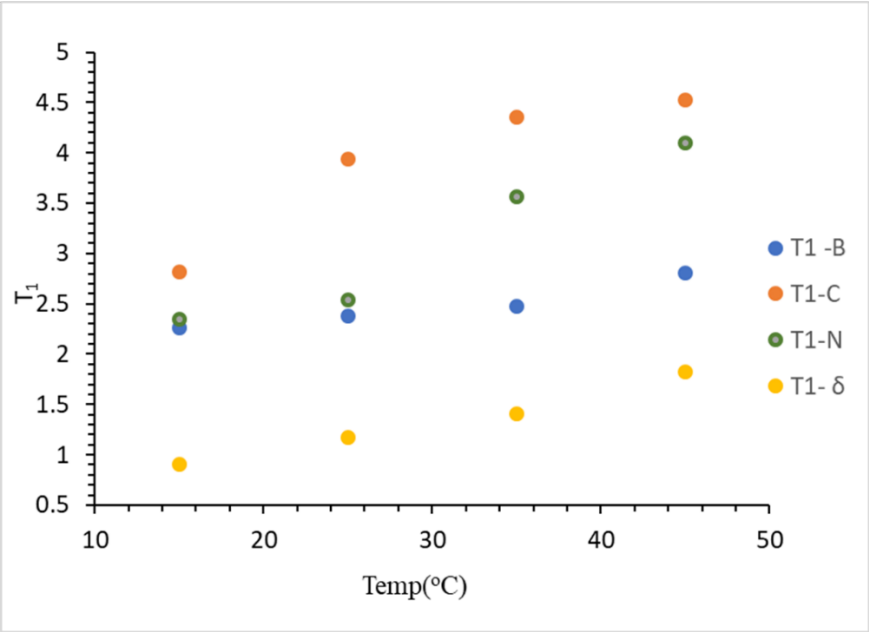


Figure 5.3. The stack plot of the T_1 against temperature of the α-H backbone, C-terminus, N-terminal and δ-H of the Tat_2 molecule.

5.2.2. Relaxation time measurements of free DNA at different temperatures

The dynamic behaviour of the single stranded DNA 5mer 5'-AGTCC-3' was investigated by analyzing the relaxation time of selected phosphorus ^{31}P nuclei at the backbone of the DNA molecule. The phosphorus backbone of the DNA was chosen based on its dynamic effect in the determining the binding site of peptide-DNA complexation. The dynamics of the DNA molecule was observed using T_1 relaxation measurements

The longitudinal relaxation rate constant $T_1(s)$ of single stranded DNA 5mer 5'-AGTC-3' is shown in Table 5.4.

Table 5.4. Longitudinal relaxation T_1(s) time measurement for the ^{31}P NMR peaks of free DNA at varying temperature

Regions	15°C	25°C	35°C	45°C
$A_{1p}G_2$	0.39	0.46	0.52	0.57
$G_{2\,p}T_3$	0.49	0.45	0.37	0.36
$T_{3p}C_4$	0.50	0.43	0.40	0.37
$C_{4p}C_5$	0.36	0.40	0.44	0.46

The A_1pG_2 and C_4pC_5 Figure 3.14 T_1's increase with increasing temperature, while the T_1's of T_3pC_4 and C_4pC_5 decreases. The differences observed in the T_1 relaxation constants within the phosphorus backbone implies that A_1pG_2 and C_4pC_5 have shorter correlation time τ_c, thus are more flexible as they are located at far ends of the DNA strand. In contrast, the T_3pC_4 and C_4pC_5 backbone that are located on the centre of the DNA strand, thus having longer correlation time τ_c, hence, resulting in longer T_1's that decrease with increasing temperature.

5.2.3. Relaxation time constant measurements of un-complexed peptides

In biomolecules, when the molecular motion occurs at the same rate with the Larmor frequency ω_o, T_1 passes through a minimum value. When the molecular motion is slower than the Larmor frequency, T_1 time constant increases while continue T_2 decreasing. At this point, the ratio of T_1/T_2 provide useful an overall estimate of the relative time scale motion in molecules, i.e. if the molecular motion is reduced, T_2 decreases and T_1 increases,

thus T_1/T_2 increases dramatically, this can occur when complexation takes place within a molecule.

Table 5.5. Longitudinal relaxation and transverse relaxation time $\boldsymbol{T_1/T_2}$ ratio measurement for the ^{1}H NMR peaks of free peptide

Residues	α-H Backbone	α-H C-terminus	α-H N-terminus	α-H	δ-H
R	-	-	-	1.35	1.21
R_3	1.59	1.05	1.15	-	1.13
R_9	1.65	1.27	1.21	-	1.19
Tat_2	1.76	1.39	1.28	-	1.25

The dynamic behaviour of the peptides as shown above indicate that, the molecules are still in the fast regime where the correlation time τ_c is about (0.001-0.1ns) because, the T_1/T_2 time ratio is just slightly greater than 1. Although, the trends in the T_1/T_2 time ratio of the R_3, R_9, and Tat_2 backbone region shows an increase in the time constant as the peptide becomes larger, the mobility of the peptide is becoming less flexible, as τ_c becomes larger. It can also be observed that the relaxation constant of α-H and δ-H in arginine (R) is also in fast motion regime where the τ_c is short. For instance, a ^{15}N relaxation rate constant of a N-labelled hen lysozyme a 28-amino acid peptide was reported. The measurement shows an increase in motion for the C-terminus chain of the arginine residue (0.03 - 0.5), where 5 of 6 tryptophan residues located at the middle core of the chain have high parameter of ($\geq$ 0.8).[89]In conclusion, it can be observed that the value of T_1/T_2 ratio relaxation constant differs for different types of peptides. Therefore, as the size of the peptide becomes longer, the relaxation rate constant also becomes bigger and the correlation time τ_c becomes longer.

5.2.4. Relaxation time constant measurements of peptide-bound DNA

The longitudinal and transverse relaxation time constants measurement ^{1}H T_1 /T_2 ratio have been analysed for R-bound ssDNA, R_3-bound ssDNA, R_9-bound ssDNA and Tat_2-bound ssDNA, R_9-bound dsDNA, and Tat_2-bound dsDNA peptides as shown in Table 5.5. Also, longitudinal and transverse relaxation time constant measurement ^{31}P T_1 /T_2 ratio have been analysed for R-bound ssDNA, R_3-bound ssDNA, R_9-bound ssDNA, and Tat_2-bound ssDNA

Table 5.6. Longitudinal relaxation and transverse relaxation time T_1 /T_2 ratio measurement for the ^{1}H NMR peaks of peptide-bound DNA

Residues	Backbone	C-terminus	N-terminus	α-H	δ-H
R-ssDNA	-	-	-	1.47	1.69
R_3-ssDNA	2.42	1.98	2.83	-	1.54
R_9-ssDNA	3.24	3.09	4.93	-	2.02
R_9-dsDNA	3.98	4.23	5.56	-	2.33
Tat_2-ssDNA	3.45	3.11	4.38	-	3.14
Tat_2-dsDNA	4.12	4.78	4.97	-	3.67

The ^{1}H T_1 /T_2 ratio relaxation measurements show that the binding of ssDNA to free peptide caused significant changes in the peptide's dynamic as shown in Table 5.6. The difference in the parameters between free and complexed arginine shows that, arginine molecule is a little affected by the binding to ssDNA. In R_3-bound ssDNA, significant changes in the peptide backbone and α-H N and C-terminus relaxation parameters for free R_3 was initially 1.59-1.05, and has increased to 2.83-1.98 when complexed to ssDNA. These observations suggest significant binding to the ssDNA, because the time ratio is at a

regime where the correlation time is becoming longer, and therefore less flexibility. However, the T_1 /T_2 ratio of the δ-H in free and complexed state shows higher motion, indicating that the guanidium group may not be directly involved in binding.

Furthermore, comparing R_9-ssDNA in their free and complexed states, in this case, show an increase in the ratio of the peptide α-H backbone, C-terminus, N-terminus, and δ-H in the complexed. To some extent this result suggests that, the ssDNA binds much more strongly to R_9 residues. Similar changes were also observed for Tat_2-bound ssDNA, where the relaxation parameters are even more dramatic at the α-H backbone, C-terminus, N-terminus, and δ-H of the peptide regions. Surprisingly, the relaxation constants of R_9-bound ssDNA and Tat_2-bound ssDNA are nearly identical.

Lastly, the dynamic properties of R_9-bound dsDNA and Tat_2-bound dsDNA in free peptide and complexed was observed. In this case, it was observed that the T_1/T_2 ratio relaxation constant of the backbone, C-terminus, N-terminus, and δ-H has increased drastically, suggesting strong binding and is consistent with longer τ_c of the ssDNA.

Table 5.7. Longitudinal relaxation and transverse relaxation time T_1 /T_2 ratio measurement for the ^{31}P NMR peaks of peptide-bound DNA

Residues	**A_1pG_2**	**G_2pT_3**	**T_3pC_4**	**C_4pC_5**
ssDNA	1.71	1.45	1.48	1.14
Arg-DNA	1.21	2.37	1.56	1.05
TriArg-DNA	2.35	2.87	2.05	2.04
R9-DNA	3.27	3.08	3.13	3.00
Tat_2-DNA	3.37	4.60	5.25	3.69

The dynamic behaviour of free DNA with complexed peptide was also observed using ^{31}P T_1/T_2 relaxation measurement. In this analysis, the changes in the T_1/T_2 ratio of the phosphate backbone in the DNA was carefully observed for complexation. From Table 5.7, R-bound ssDNA ratio shows a small increase in T_1/T_2 ratio of G_2pT_3 region of the DNA suggesting a strong binding with the ssDNA at that region compared to the other sites. However, in R_3-bound ssDNA T_1/T_2 ratio, all the phosphorus backbones experience significant dynamic changes, suggesting stronger binding with ssDNA, and a higher stoichiometry.

Furthermore, as observed for the ^{1}H T_1/T_2 ratio parameter in the R_9-bound ssDNA and Tat_2-bound ssDNA, the ^{31}P relaxation measurement of the two complexes also show similar behaviour of complexation in the observed DNA backbone. The ^{31}P T_1/T_2 ratio of both complexes shows strong binding with the ssDNA backbone, suggesting that complexation is taking place within the molecules.

CHAPTER 6

6. CONCLUSION AND FUTURE DIRECTIONS

The structural analysis of CPP: nucleic acids complexes have been characterized through solution-state nuclear magnetic resonance (NMR). Although, information has been gained in regards to aggregation behaviour of CPP-nucleic acid polyplexes, little has been done in the way of their structural dynamic behavior. In this thesis, one of the primary objectives was to understand the dynamic details of the CPP- nucleic acid polyplexes, and relate the observations to the mechanism of their function. Of the various techniques employed for CPP- nucleic acid polyplexes, NMR spectroscopy is one of the most valuable, not only providing the quantitative and qualitative structural characterization, but also dynamic information.

The first aspect of this thesis focused primarily on the use of solution-state NMR analysis of 5mer single and double stranded DNA. Complete ^{1}H, ^{13}C and ^{31}P 1-D and 2-D NMR techniques were performed in D_2O to verify the full structural assignment of the compounds in solution. The analysis helps to confirm the relative assignment of all the proton, carbon and phosphorus within the spectra. The overall assignment provided a greater understanding of the structural features of the DNA that gave rise to their unique binding properties.

In the second part of this thesis, the complexation of CPP-DNA behavior was investigated using ^{1}H and ^{31}P solution sate NMR spectroscopy. The complexation behaviour was monitored between single stranded DNA 5mer 5'-AGTCC-3' and arginine, tri-arginine, nano-arginine, and Tat_2 using a NMR titration experiment. The titration was restricted to the analyzes of N-terminus α-H of the peptides. The ^{1}H NMR spectra of R-

ssDNA shows rapid decrease in the chemical of the α-H. From the titration curve Figure 4.10, the result seems to show a 1:1 stoichiometry complex of R-ssDNA with equilibrium constant greater than one. The R_3-ssDNA ^{1}H NMR titration curve also seems to be a stoichiometry of 1:1 complex, although with a small equilibrium constant. However, R_9-ssDNA complex appears to be a 2:1 stoichiometry complex. This is not surprising because R_9 peptide has more positively charge amino group that can bind to the ssDNA, and as such will have more binding affinity as compared to arginine and tri-arginine. The ^{31}P NMR spectroscopy was also used to observe complexation behaviour at the backbone of the DNA molecules. In this analysis, it was observed that as the length of the peptide complexed with the DNA became longer the linewidth of the ^{31}P spectra became broader, particularly for R_9-ssDNA and Tat_2-ssDNA spectra. It was also observed that the ^{31}P chemical shift difference of 0.2ppm at the backbone of the ssDNA upon complexation with R_9 and Tat_2, indicate similar behaviours of these two peptides. This observation was also reported by *Jordan pepper* et al. It was reported that the values known as "jerk" parameter that describes the change in the acceleration of growth upon complexation of R_9 and Tat_2 with oligonucleotide was 13.39nm/hr^3 and13.39nm/hr^3 for R_9 and Tat_2 respectively.[81] The complexation behavior of these CPPs and DNA shows that the length of the peptide is dependent on the complexation of CPP-nucleic acids.

In the third part of this thesis, a time-scale relaxation NMR measurement was analysed to investigate the dynamic behaviour of CPP-DNA complexes. This NMR technique offers a great deal of insight into dynamic and structural orientational motion of the complexes. Using T_1 and T_2 relaxation measurements, more flexibility was observed in the interaction of DNA with R and R_3, although the mobility is somewhat reduced relative

to the free DNA. The relaxation parameter of the R_9 and Tat_2 upon complexation with ssDNA and dsDNA indicates relatively less mobility.

Finally, the complexation behaviour of CPP-DNA was investigated using ^{1}H NMR spectroscopy. In this case, complexation was observed at the base region of the ssDNA and dsDNA upon interaction with R, R_3, R_9, and Tat_2. The results show a strong interaction of R, R_3, R_9, and Tat_2 with both ssDNA and dsDNA at 1:1 mole ratio. The chemical shift of the proton close to the nitrogen group (acceptor and donors) in the DNA has changed drastically upon interaction with the peptide. These might indicate a sort of hydrogen bonding interaction taking place between the side chains of the peptide and the major and minor groove of the DNA. For instance, in the ^{1}H NMR spectra of the complexation between R_9/dsDNA and Tat_2/dsDNA, the base signals were broadened because of stronger interactions. This observation was also reported for spermine binding Z-DNA form of $d(CGCCGCG)_2$, where the complexation of spermine with the DNA at mole ratio 1:1 results in the broadening of the ^{1}H NMR chemical shift of the NH and NH_2 signals of the DNA.[79]

In the case of the R_9 and Tat_2 complexation with ssDNA and dsDNA, pH must be investigated further. This is because the structure of the complex in various protonated conditions may contribute to the transfection efficiency. For instance, as shown in Chapter 3 the ^{1}H chemical shift of arginine changes drastically as the pH decreases. This indicates that lowering the pH below 7 might increase complexation, thereby inducing the internalization mechanism before the translocation of CPP-cargo complexes. It was shown that at pH 5 the size of the complexes formed between polyhistidine and DNA increases the electrostatic interaction between them, and was useful for the transfer of DNA to cells for the purpose of gene therapy.[90]

In the case of the size of the polypeptides, it appears that longer chains of arginine rich polypeptides will be viable option when considering the type of CPP to be used for cellular uptake. This was evident in Chapter 4, in which the interaction of the small peptide and amino acid (R and R_3) show a 1:1 stoichiometry complexes which may not be strong enough to deliver DNA to cell. However, R_9 peptide shows a strong binding of 2:1 stoichiometry with large equilibrium constant. with the ssDNA behaviour but also stronger interaction with the DNA. It may also be helpful to increase the dsDNA strand up to 22mer, this will result in less tumbling of the DNA helix, thereby allowing stable interaction with polypeptide that can be easily observed using NMR techniques.

While the studies presented within this thesis maybe the starting point into the NMR structural dynamic behavior of CPP-nucleic acid complexations, there remain several confirmatory studies which must be carried out. The future work will likely take the shape of taking advantage of the length of the peptide and DNA (DNA or RNA, single stranded versus double stranded) specialized for delivery in plant cells. This is because one of the results in this thesis shows a change in equilibrium constant, and stoichiometry ratio as the length of the peptide increases. Also for the DNAs, we saw changes in the signals of the base protons for ssDNA and dsDNA.

Secondly, the actual mode of binding has to be further investigated. One of the results we found in this work is that, the N-terminal proton signals of the peptide is a good indicator of how the peptide behaves upon binding with the DNA. Another mode of binding has been shown for spermine a polyamine peptide. In these studies, it was shown that the binding of spermine and DNA occur through hydrogen bonding interaction at the base region of the DNA. A similar behaviour was observed at the base region of the ssDNA and dsDNA ^{1}H NMR spectra. The changes in the chemical shift of the complex peptide-DNA

compared to the free DNA seems to be an indication of hydrogen bonding, although further experiments need to be performed to confirm it.

More so, the effect of temperature on CPP-nucleic acid aggregation behavior has not been studied in this thesis, which is an important factor to consider. Studying how the complexations are affected by temperature, the role of entropy can be determined. 1-D ^{1}H NMR and relaxation measurements can be used for this purpose and can reveal how the interactions between CPP- nucleic acids change with temperature.

REFERENCES

1. Chugh, A.; Amundsen, E.; Eudes, F., Translocation of cell-penetrating peptides and delivery of their cargoes in triticale microspores. *Plant Cell Reports* **2009,** *28* (5), 801-810.

2. Ryser, H. J. P.; Shen, W.-C.; Merk, F. B., Membrane transport of macromolecules: New carrier functions of proteins and poly(amino acids). *Life Sciences* **1978,** *22* (13), 1253-1260.

3. Evans, W. H.; Graham, J. M., *Membrane structure and function*. IRL Press at Oxford University Press: Oxford, England;New York;, 1989.

4. Armaleo, D.; Ye, G. N.; Klein, T. M.; Shark, K. B.; Sanford, J. C.; Johnston, S. A., Biolistic nuclear transformation of Saccharomyces cerevisiae and other fungi. *Current genetics* **1990,** *17* (2), 97-103.

5. Chugh, A.; Eudes, F.; Shim, Y.-S., Cell-penetrating peptides: Nanocarrier for macromolecule delivery in living cells. *IUBMB life* **2010,** *62* (3), 183-193.

6. Nielsen, P. E.; Egholm, M.; Buchardt, O., Sequence-specific transcription arrest by peptide nucleic acid bound to the DNA template strand. *Gene* **1994,** *149* (1), 139-145.

7. Said Hassane, F.; Saleh, A. F.; Abes, R.; Gait, M. J.; Lebleu, B., Cell penetrating peptides: overview and applications to the delivery of oligonucleotides. *Cellular and Molecular Life Sciences* **2010,** *67* (5), 715-726.

8. Rudolph, C.; Plank, C.; Lausier, J.; Schillinger, U.; MÃ¼ller, R. H.; Rosenecker, J., Oligomers of the Arginine-rich Motif of the HIV-1 TAT Protein Are Capable of Transferring Plasmid DNA into Cells. *Journal of Biological Chemistry* **2003,** *278* (13), 11411-11418.

9. Gelvin, S. B., Agrobacterium-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool. *Microbiology and Molecular Biology Reviews* **2003,** *67* (1), 16-37.

10. Pitzschke, A., Agrobacterium infection and plant defense—transformation success hangs by a thread. *Frontiers in Plant Science* **2013,** *4*, 519.

11. Tzfira, T.; Citovsky, V.; MyiLibrary, *Nuclear import and export in plants and animals*. Landes Bioscience/Eurekah.com: Georgetown, Tex;New York, N.Y;, 2005.

12. Gelvin, S. B., Agrobacterium VirE2 Proteins Can Form a Complex with T Strands in the Plant Cytoplasm. *Journal of Bacteriology* **1998,** *180* (16), 4300-4302.

13. Shrawat, A. K.; Lörz, H., Agrobacterium-mediated transformation of cereals: a promising approach crossing barriers. *Plant Biotechnology Journal* **2006,** *4* (6), 575-603.

14. Klinner, U.; Schäfer, B., Genetic aspects of targeted insertion mutagenesis in yeasts. *FEMS Microbiology Reviews* **2004,** *28* (2), 201-223.

15. Sparks, C. A.; Jones, H. D., Genetic transformation of wheat via particle bombardment. *Methods in molecular biology (Clifton, N.J.) U6 - ctx_ver=Z39.88-2004&ctx_enc=info%3Aofi%2Fenc%3AUTF-8&rfr_id=info%3Asid%2Fsummon.serialssolutions.com&rft_val_fmt=info%3Aofi%2Ffmt%3Akev%3Amtx%3Ajournal&rft.genre=article&rft.atitle=Genetic+transformation+of+wheat+via+particle+bombardment&rft.jtitle=Methods+in+molecular+biology+%28Clifton%2C+N.J.%29&rft.au=Sparks%2C+Caroline+A&rft.au=Jones%2C+Huw+D&rft.date=2014-01-01&rft.eissn=1940-6029&rft.volume=1099&rft.spage=201&rft_id=info%3Apmid%2F24243206&rft_id=info%3Apmid%2F24243206&rft.externalDocID=24243206&paramdict=en-US U7 - Journal Article* **2014,** *1099*, 201.

16. Sanford, J. C., BIOLISTIC PLANT TRANSFORMATION. *PHYSIOLOGIA PLANTARUM* **1990,** *79* (1), 206-209.

17. van Meer, G.; Feigenson, G. W.; Voelker, D. R., Membrane lipids: where they are and how they behave. *Nature reviews molecular cell biology* **2008,** *9* (2), 112-124.

18. Alves, I. D.; Goasdoué, N.; Correia, I.; Aubry, S.; Galanth, C.; Sagan, S.; Lavielle, S.; Chassaing, G., Membrane interaction and perturbation mechanisms induced by two cationic cell penetrating peptides with distinct charge distribution. *BBA - General Subjects* **2008,** *1780* (7), 948-959.

19. Bechara, C.; Sagan, S., Cell-penetrating peptides: 20years later, where do we stand? *FEBS Letters 587* (12), 1693-1702.

20. Frankel, A. D.; Pabo, C. O., Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **1988,** *55* (6), 1189-1193.

21. Billeter, M.; Qian, Y. Q.; Otting, G.; Muller, M.; Gehring, W.; Wuthrich, K., DETERMINATION OF THE NUCLEAR-MAGNETIC-RESONANCE SOLUTION STRUCTURE OF AN ANTENNAPEDIA HOMEODOMAIN-DNA COMPLEX. *JOURNAL OF MOLECULAR BIOLOGY* **1993,** *234* (4), 1084-1094.

22. Deshayes, S.; Morris, M.; Heitz, F.; Divita, G., Delivery of proteins and nucleic acids using a non-covalent peptide-based strategy. *Advanced Drug Delivery Reviews* **2008,** *60* (4), 537-547.

23. Torchilin, V., Intracellular delivery of protein and peptide therapeutics. *Drug Discovery Today: Technologies* **2008,** *5* (2–3), e95-e103.

24. Munyendo, W. L. L.; Lv, H.; Benza-Ingoula, H.; Baraza, L. D.; Zhou, J., Cell Penetrating Peptides in the Delivery of Biopharmaceuticals. *Biomolecules* **2012,** *2* (2), 187-202.

25. Gottschalk, S.; Sparrow, J. T.; Hauer, J.; Mims, M. P.; Leland, F. E.; Woo, S. L. C.; Smith, L. C., A novel DNA-peptide complex for efficient gene transfer and expression in mammalian cells. *GENE THERAPY* **1996,** *3* (5), 448-457.

26. Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A., The third helix of the Antennapedia homeodomain translocates through biological membranes. *Journal of Biological Chemistry* **1994,** *269* (14), 10444.

27. Das, A. T.; Harwig, A.; Berkhout, B., The HIV-1 Tat Protein Has a Versatile Role in Activating Viral Transcription. *Journal of Virology* **2011,** *85* (18), 9506-9516.

28. Herce, H. D.; Garcia, A. E., Cell Penetrating Peptides: How Do They Do It? *Journal of Biological Physics* **2007,** *33* (5), 345-356.

29. Pooga, M.; Soomets, U.; Hallbrink, M.; Valkna, A.; Saar, K.; Rezaei, K.; Kahl, U.; Hao, J. X.; Xu, X. J.; Wiesenfeld-Hallin, Z.; Hokfelt, T.; Bartfai, A.; Langel, U., Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo. *NATURE BIOTECHNOLOGY* **1998,** *16* (9), 857-861.

30. Pooga, M.; Hallbrink, M.; Zorko, M.; Langel, U., Cell penetration by transportan. *FASEB JOURNAL* **1998,** *12* (1), 67-77.

31. Green, M.; Ishino, M.; Loewenstein, P. M., Mutational analysis of HIV-1 Tat minimal domain peptides: Identification of trans-dominant mutants that suppress HIV-LTR-driven gene expression. *Cell* **1989,** *58* (1), 215-223.

32. Su, Y. C.; Waring, A. J.; Ruchala, P.; Hong, M., Membrane-Bound Dynamic Structure of an Arginine-Rich Cell-Penetrating Peptide, the Protein Transduction Domain of HIV TAT, from Solid-State NMR. *BIOCHEMISTRY* **2010,** *49* (29), 6009-6020.

33. Milletti, F., Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discovery Today* **2012,** *17* (15–16), 850-860.

34. Duchardt, F.; Fotin-Mleczek, M.; Schwarz, H.; Fischer, R.; Brock, R. E., A comprehensive model for the cellular uptake of cationic cell-penetrating peptides. *Traffic* **2007,** *8* (7), 848-866.

35. Koehler Leman, J.; Ulmschneider, M. B.; Gray, J. J., Computational modeling of membrane proteins. *Proteins: Structure, Function, and Bioinformatics* **2015,** *83* (1), 1-24.

36. Su, Y.; Li, S.; Hong, M., Cationic membrane peptides: atomic-level insight of structure–activity relationships from solid-state NMR. *Amino Acids* **2013,** *44* (3), 821-833.

37. McDermott, A., Structure and Dynamics of Membrane Proteins by Magic Angle Spinning Solid-State NMR. *ANNUAL REVIEW OF BIOPHYSICS* **2009,** *38* (1), 385-403.

38. Hammill, J. T.; Miyake-Stoner, S.; Hazen, J. L.; Jackson, J. C.; Mehl, R. A., Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization. *Nature protocols* **2007,** *2* (10), 2601-2607.

39. Crowley, P. B.; Kyne, C.; Monteith, W. B., Simple and inexpensive incorporation of 19F-Tryptophan for protein NMR spectroscopyElectronic supplementary information (ESI) available: Overlaid 1H-15N HSQC spectra of GB1 and GB1 containing 5-fluorotryptophan. See DOI: 10.1039/c2cc35347d. **2012,** *48* (86), 1681-1683.

40. Crowley, P. B.; Kyne, C.; Monteith, W. B., Simple and inexpensive incorporation of 19F-tryptophan for protein NMR spectroscopy. *Chemical communications (Cambridge, England) U6 - ctx_ver=Z39.88-2004&ctx_enc=info%3Aofi%2Fenc%3AUTF-8&rfr_id=info:sid/summon.serialssolutions.com&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&rft.genre=article&rft.atitle=Simple+and+inexpensive+incorporation+of+19F-tryptophan+for+protein+NMR+spectroscopy&rft.jtitle=Chemical+communications+%28Cambridge%2C+England%29&rft.au=Crowley%2C+Peter+B&rft.au=Kyne%2C+Ciara&rft.au=Monteith%2C+William+B&rft.date=2012-11-07&rft.eissn=1364-548X&rft.volume=48&rft.issue=86&rft.spage=10681&rft_id=info:pmid/23000821&rft.externalDocID=23000821&paramdict=en-US U7 - Journal Article* **2012,** *48* (86), 10681.

41. Hippel, P. H. v.; Berg, O. G., Facilitated target location in biological systems. *Journal of Biological Chemistry* **1989,** *264* (2), 675-678.

42. Freisinger, E.; Fernandes, A.; Grollman, A. P.; Kisker, C., Crystallographic Characterization of an Exocyclic DNA Adduct: 3,N4-etheno-2′-deoxycytidine in the Dodecamer 5′-CGCGAATTεCGCG-3′. *Journal of Molecular Biology* **2003,** *329* (4), 685-697.

43. Cullinan, D.; Johnson, F.; de los Santos, C., Solution structure of an 11-mer duplex containing the 3,N4-ethenocytosine adduct opposite 2′-deoxycytidine: implications for the recognition of exocyclic lesions by DNA glycosylases1. *Journal of Molecular Biology* **2000,** *296* (3), 851-861.

44. Columbus, L.; Hubbell, W. L., A new spin on protein dynamics. *Trends in Biochemical Sciences* **2002,** *27* (6), 288-295.

45. Hennig, J.; Wang, I.; Sonntag, M.; Gabel, F.; Sattler, M., Combining NMR and small angle X-ray and neutron scattering in the structural analysis of a ternary protein-RNA complex. *Journal of Biomolecular NMR* **2013,** *56* (1), 17-30.

46. Xiao, S. Y.; Zhu, H.; Wang, L.; Liang, H. J., DNA conformational flexibility study using phosphate backbone neutralization model. *SOFT MATTER* **2014,** *10* (7), 1045-1055.

47. Iwahara, J.; Wojciak, J. M.; Clubb, R. T., Improved NMR spectra of a protein–DNA complex through rational mutagenesis and the application of a sensitivity optimized isotope-filtered NOESY experiment. *Journal of Biomolecular NMR* **2001,** *19* (3), 231-241.

48. van der Werf, R. M.; Tessari, M.; Wijmenga, S. S., Nucleic acid helix structure determination from NMR proton chemical shifts. *Journal of Biomolecular NMR* **2013,** *56* (2), 95-112.

49. Schmiedeke, T. M.; Stöckl, F. W.; Weber, R.; Sugisaki, Y.; Batsford, S. R.; Vogt, A., Histones have high affinity for the glomerular basement membrane. Relevance for immune complex formation in lupus nephritis. *The Journal of Experimental Medicine* **1989,** *169* (6), 1879.

50. Maiti, P. K.; Bagchi, B., Structure and dynamics of DNA-dendrimer complexation: Role of counterions, water, and base pair sequence. *NANO LETTERS* **2006,** *6* (11), 2478-2485.

51. Rothbard, J. B.; Jessop, T. C.; Lewis, R. S.; Murray, B. A.; Wender, P. A., Role of membrane potential and hydrogen bonding in the mechanism of translocation of guanidinium-rich peptides into cells. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY* **2004,** *126* (31), 9506-9507.

52. Zaro, J. L.; Shen, W.-C., Cytosolic delivery of a p16-peptide oligoarginine conjugate for inhibiting proliferation of MCF7 cells. *Journal of Controlled Release* **2005,** *108* (2), 409-417.

53. Boisguérin, P.; Deshayes, S.; Gait, M. J.; O'Donovan, L.; Godfrey, C.; Betts, C. A.; Wood, M. J. A.; Lebleu, B., Delivery of therapeutic oligonucleotides with cell penetrating peptides. *Advanced Drug Delivery Reviews* **2015,** *87*, 52-67.

54. Morris, M. C.; Vidal, P.; Chaloin, L.; Heitz, F.; Divita, G., A new peptide vector for efficient delivery of oligonucleotides into mammalian cells. *NUCLEIC ACIDS RESEARCH* **1997,** *25* (14), 2730-2736.

55. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991,** *254* (5037), 1497-1500.

56. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E., Peptide nucleic acids (PNAs) containing thymine monomers derived from chiral amino acids: Hybridization and solubility properties of D-lysine PNA. *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION* **1996,** *35* (17), 1939-1942.

57. Autiero, I.; Saviano, M.; Langella, E., Conformational studies of chiral D-Lys-PNA and achiral PNA system in binding with DNA or RNA through a molecular dynamics approach. *EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY* **2015,** *91*, 109-117.

58. Tedeschi, T.; Sforza, S.; Dossena, A.; Corradini, R.; Marchelli, R., Lysine-based peptide nucleic acids (PNAs) with strong chiral constraint: Control of helix handedness and DNA binding by chirality. *Chirality* **2005,** *17* (S1), S196-S204.

59. Smith, H. C., *RNA and DNA editing: molecular mechanisms and their integration into biological systems*. Wiley-Interscience: Hoboken, N.J, 2008.

60. Kim, S.; Zhu, L.; Yu, B.; Chen, Y., Letter to the Editor: Assignment of 1H, 13C and 15N resonances of the ARID domain of P270. *Journal of Biomolecular NMR* **2003,** *27* (3), 277-278.

61. Winkler, F. K.; Banner, D. W.; Oefner, C.; Tsernoglou, D.; Brown, R. S.; Heathman, S. P.; Bryan, R. K.; Martin, P. D.; Petratos, K.; Wilson, K. S., THE CRYSTAL-STRUCTURE OF ECORV ENDONUCLEASE AND OF ITS COMPLEXES WITH COGNATE AND NON-COGNATE DNA FRAGMENTS. *EMBO JOURNAL* **1993,** *12* (5), 1781-1795.

62. Iwahara, J.; Zweckstetter, M.; Clore, G. M., NMR Structural and Kinetic Characterization of a Homeodomain Diffusing and Hopping on Nonspecific DNA. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** *103* (41), 15062-15067.

63. K.T, D.; A.S, B.; J.R, W., Base Flexibility in HIV-2 TAR RNA Mapped by Solution 15N, 13C NMR Relaxation. *Journal of Molecular Biology* **2002,** *317* (2), 263-278.

64. Keeler, J., *Understanding NMR spectroscopy*. John Wiley and Sons: Chichester, U.K, 2010; Vol. 2nd.

65. Jacobsen, N. E., *NMR spectroscopy explained: simplified theory, applications and examples for organic chemistry and structural biology*. Wiley-Interscience: Hoboken, N.J, 2007.

66. Levitt, M. H., *Spin dynamics: basics of nuclear magnetic resonance*. John Wiley & Sons: Chichester, England;Mississauga, ON;, 2008; Vol. 2nd.

67. Hore, P. J., *Nuclear magnetic resonance*. Oxford University Press: Oxford;New York;, 1995; Vol. 32.

68. Günther, H., *NMR spectroscopy: basic principles, concepts, and applications in chemistry*. Wiley-VCH: Weinheim, 2013; Vol. 3rd, completely rev. and updat.

69. Zerbe, O.; Jurt, S., *Applied NMR spectroscopy for chemists and life scientists*. Wiley-VCH: Weinheim, 2014.

70. Egelman, E., *Comprehensive biophysics*. Academic: United States, 2012.

71. Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M., Protein structure determination from NMR chemical shifts. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (23), 9615-9620.

72. McMahon, B. H.; Fabian, M.; Tomson, F.; Causgrove, T. P.; Bailey, J. A.; Rein, F. N.; Dyer, R. B.; Palmer, G.; Gennis, R. B.; Woodruff, W. H., FTIR studies of internal proton transfer reactions linked to inter-heme electron transfer in bovine cytochrome c oxidase. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2004,** *1655* (Supplement C), 321-331.

73. El-Andaloussi, S.; Järver, P.; Johansson, Henrik J.; Langel, Ü., Cargo-dependent cytotoxicity and delivery efficacy of cell-penetrating peptides: a comparative study. *The Biochemical Journal* **2007,** *407* (Pt 2), 285-292.

74. Holliday, G. L.; Mitchell, J. B. O.; Thornton, J. M., Understanding the Functional Roles of Amino Acid Residues in Enzyme Catalysis. *Journal of Molecular Biology* **2009,** *390* (3), 560-577.

75. Artz, P. G.; Valentine, K. G.; Opella, S. J.; Lu, P., Lac repressor-operator interaction: N-terminal peptide backbone 1H and 15N chemical shifts upon complex formation with DNA. *Journal of molecular recognition : JMR* **1996,** *9* (1), 13.

76. Fitch, C. A.; Platzer, G.; Okon, M.; Garcia-Moreno E, B.; McIntosh, L. P., Arginine: Its pKa value revisited. *Protein Science* **2015,** *24* (5), 752-761.

77. Castagné, C.; Murphy, E. C.; Gronenborn, A. M.; Delepierre, M., 31P NMR analysis of the DNA conformation induced by protein binding SRY/DNA complexes. *European journal of biochemistry / FEBS* **2000,** *267* (4), 1223-1229.

78. Vorlíčková, M.; Kypr, J.; Sklenář, V.; Klug, A., Salt-induced conformational transition of poly[d(A-T)]·poly[d(A-T)]. *Journal of Molecular Biology* **1983,** *166* (1), 85-92.

79. Banville, D. L.; Feuerstein, B. G.; Shafer, R. H., 1H and 31P nuclear magnetic resonance studies of spermine binding to the Z-DNA form of d(m5CGm5CGm5CG)2: Evidence for decreased spermine mobility. *Journal of Molecular Biology* **1991,** *219* (4), 585-590.

80. Millili, P. G.; Selekman, J. A.; Blocker, K. M.; Johnson, D. A.; Naik, U. P.; Sullivan, M. O., Structural and functional consequences of poly(ethylene glycol) inclusion on DNA condensation for gene delivery. *Microscopy Research and Technique* **2010,** *73* (9), 866-877.

81. Pepper, J. T. Characterization of cell-penetrating peptide complexation and interaction with plant cells. ProQuest Dissertations Publishing, 2015.

82. Lohman, T. M., Pentalysine-deoxyribonucleic acid interactions: a model for the general effects of ion concentrations on the interactions of proteins with nucleic acids. *Biochemistry (Easton)* **1980,** *19* (15), 3522-3530.

83. Record, M. T.; Lohman, T. M.; Haseth, P. d., Ion effects on ligand-nucleic acid interactions. *Journal of Molecular Biology* **1976,** *107* (2), 145-158.

84. Porschke, D., Nature of protamine-DNA complexes: A special type of ligand binding co-operativity. *Journal of Molecular Biology* **1991,** *222* (2), 423-433.

85. Winter, Diffusion-driven mechanisms of protein translocation on nucleic acids. 2. The Escherichia coli lac repressor-operator interaction: equilibrium measurements. *Biochemistry (Easton)* **1981,** *20* (24), 6948.

86. Kowalczykowski, S. C.; Lonberg, N.; Newport, J. W.; von Hippel, P. H., Interactions of bacteriophage T4-coded gene 32 protein with nucleic acids: I. Characterization of the binding interactions. *Journal of Molecular Biology* **1981,** *145* (1), 75-104.

87. Jensen, D. E.; Kelly, R. C.; Hippel, P. H. v., DNA "melting" proteins. II. Effects of bacteriophage T4 gene 32-protein binding on the conformation and stability of nucleic acid structures. *Journal of Biological Chemistry* **1976,** *251* (22), 7215.

88. Shokri, L.; Rouzina, I.; Williams, M. C., Interaction of bacteriophage T4 and T7 single-stranded DNA-binding proteins with DNA. *Physical Biology* **2009,** *6* (2), 025002.

89. Buck, M., Structural Determinants of Protein Dynamics: Analysis of 15N NMR Relaxation Measurements for Main-Chain and Side-Chain Nuclei of Hen Egg White Lysozyme. *Biochemistry (Easton)* **1995,** *34* (12), 4041-4055.

90. Putnam, D.; Zelikin, A. N.; Izumrudov, V. A.; Langer, R., Polyhistidine–PEG:DNA nanocomposites for gene delivery. *Biomaterials* **2003,** *24* (24), 4425-4433.