

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

## Unveiling biomacromolecule interactions

NMR and optical spectroscopy studies on ligand binding to DNA and lysozyme

Lisha Wu

Department of Chemical and Biological Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2013

Unveiling biomacromolecules interactions– NMR and optical spectroscopy studies on ligand binding to DNA and lysozyme

Lisha Wu

ISBN 978-91-7385-894-6

© Lisha Wu, 2013.

Doktorsavhandlingar vid Chalmers tekniska högskola

Ny serie nr 3575

ISSN 0346-718X

Department of Chemical and Biological Engineering

Chalmers University of Technology

SE-412 96 Gothenburg

Sweden

Telephone + 46 (0)31-772 1000

Cover picture: The mushroom is a photo of the 900MHz NMR instrument in the Swedish NMR center (on the left upper panel), which provides NOESY and gCOSY spectra (on the left lower panel). The residues that display chemical shifts are marked with spheres on lysozyme structure (PDB: 1lzs, right upper panel) and the DNA-ruthenium complex (right lower panel) obtained from this thesis study.

Reproservice  
Göteborg 2013.

To my beloved family members: Baba  
mama and my fiancé.



# Unveiling biomacromolecule interactions – NMR and optical spectroscopy studies on ligand binding to DNA and lysozyme.

Lisha Wu

Department of Chemical and Biological Engineering

Chalmers University of Technology

**Abstract:** Studies on intercalation of bulky binuclear ruthenium compounds into DNA have attracted attention due to their slow dissociation rate and sequence selectivity. Previous results showed that dumb-bell shaped molecules of the type  $[\mu\text{-}(\text{bidppz})(\text{L})_4\text{Ru}_2]^{4+}$ , L=phenanthroline or bipyridine, bind rapidly on the surface of DNA prior to intercalation. Functional understanding of the intercalation mechanism may greatly gain from structural characterization of both surface bound and intercalation states. In this work the initial surface bound state has for the first time been structurally characterized by 2-D NMR in detail, showing the binuclear compound aligned into the minor groove. A structure close to the proposed following intercalated state was found by X-ray crystallography: one dppz moiety inserts into the DNA stack of a hexamer duplex through extrusion of an AT base pair while the second moiety make such complexes being dimerized by end-stacking. Optical spectroscopy results indicate that these interactions are present also in dilute solution.

Minor groove binding to DNA was also studied with the classical drug Hoechst 33258. By a combination of 1D-NMR, calorimetry and optical spectroscopy, two drug molecules were found to bind to the minor groove of a consecutive A<sub>4</sub>T<sub>4</sub> sequence, one after another, instead of a ‘sandwich’ conformation as previously proposed.

Proteins are another type of essential biomacromolecules, and in this thesis, lysozyme, an antimicrobial enzyme is studied. Lysozyme was found to undergo domain wide structure rearrangements accompanying the deprotonation of Glu 35, involving helix movements, reorientation of binding site residues and variations of the hydrogen bond pattern near the active site. Similar structural changes are observed when short polysaccharides are bound. Interestingly, surface properties such as electrostatic potential and hydrophobic patches significantly modulate the interaction of a disaccharide with intermolecular contacts remote from the catalytic site, while binding affinity near the active site shows fewer variations. Both the dynamic behavior of the  $\alpha$ -domain as well as the surface properties of all binding sites are critical factors in the design or optimization of lysozyme-based compounds for applications in food preservation and pharmaceutical usage.

**Key words:** NMR, CD, fluorescence, DNA, ruthenium compounds, Hoechst 33258, lysozyme, intercalation, structure rearrangements.



## LIST OF PUBLICATIONS

This thesis is based on the work presented in the following paper:

- I. Initial DNA Interactions of the Binuclear Threading Intercalator  $\Lambda,\Lambda\text{-}[\mu\text{-}\text{bidppz(bipy)}_4\text{Ru}_2]^{4+}$ : An NMR study with  $[\text{d}(\text{CGCGAATTCGCG})]_2$   
*Lisha Wu, Anna Reymen, Cecilia Persson, Krzysztof Kazimierczuk, Tom Brown, Per Lincoln, Bengt Nordén\*, and Martin Billeter*  
*Chemistry-A European Journal 2013, 19, 5401-5410.*
- II. Bi-dipyridophenazine diruthenium complex thread-inserts into the DNA double-helix by extrusion of AT base pairs and cross-links parallel DNA duplexes  
*D. Roeland Boer, Lisha Wu, Per Lincoln, and Miquel Coll\**.  
*Manuscript*
- III. Minor-Groove Binding Drugs: Where is the Second Hoechst 33258 Molecule?  
*Louise H. Fornander, Lisha Wu, Martin Billeter, Per Lincoln, and Bengt Nordén\**.  
*Journal of Physical Chemistry B, 2013, 117, 5820-5830.*
- IV. Lysozyme revisited by NMR: deprotonation of the catalytic Glu 35 and ligand binding are coupled to domain-wide motions  
*Lisha Wu, Vadim B. Krylov, Nikolay E. Nifantiew, Philipp Markart, Andreas Günther, Hans-Christian Siebert, Bengt Nordén\*, Martin Billeter.*  
*Submitted to Biochemistry*

# **CONTRIBUTION REPORT**

## **Paper I**

This paper is mainly my work; I carried out all the experiments and performed NMR data analysis. I calculated the initial structure and prepared the figures with exceptions of MD simulation. I participated in the discussions, and I wrote the paper.

## **Paper II**

I performed solution studies and participated in discussions. I contributed to the manuscript preparation.

## **Paper III**

I carried out the NMR measurements and analysis. I participated in the discussions. And I wrote the NMR section for the paper.

## **Paper IV**

This paper is mainly my work; I carried out all the NMR measurements, data analysis, prepared all the figures, participated in the discussions, and finally, I wrote the paper.

## Contents

1	INTRODUCTION .....	1
2	BACKGROUND .....	5
2.1	DNA .....	5
2.1.1	DNA duplex structure .....	5
2.1.2	Mismatch DNA .....	6
2.2	Ruthenium complexes and Hoechst 33258 .....	8
2.3	Lysozymes .....	11
2.3.1	Different types of lysozymes .....	11
2.3.2	Catalytic mechanisms of lysozymes related to structures .....	12
2.3.3	Biological role of lyosyzmes.....	14
3	THEORY AND METHODOLOGY .....	15
3.1	NMR Spectroscopy .....	15
3.1.1	Basic concepts: Magnetization, Boltzman equilibrium, FID, and chemical shift.....	16
3.1.2	Chemical exchange .....	21
3.1.3	Nuclear Overhauser Effects SpectroscopY (NOESY) .....	21
3.1.4	Experiments and assignments strategy.....	23
3.2	Optical measurements .....	25
3.2.1	Circular Dichroism.....	25
3.2.2	Fluorescence .....	25
3.3	Computational structure determinations .....	26
4	RESULTS .....	31
4.1	Groove binding .....	32
4.2	Binding to mismatch DNA .....	40
4.3	Intercalation into hexamer DNA .....	44
4.4	Binding studies of Hoechst 33258 to AT DNA .....	46
4.5	Correlation of human lysozyme structural changes with the deprotonation process of Glu 35.....	47
5	CONCLUDING REMARKS .....	57
6	ACKNOWLEDGEMENTS.....	59
7	REFERENCES .....	61



# 1 INTRODUCTION

DNA carries the genetic information and forms the basis for all living organisms. This genetic information is passed on from generation to generation by replication. Genes in DNA encode protein molecules, which carry out all the functions necessary for life. For instance, in the replication step, DNA polymerase, an essential enzyme, copies DNA strands. In general, polymerases are highly accurate, with an error rate of less than one mistake for every  $10^7$  nucleotide added [1]. During replication, a sliding clamp protein (Figure 1.1) binds DNA polymerase and prevents the dissociation of the polymerase from the template DNA strand. The association of a clamp protein with polymerase determines the rate of DNA synthesis [2]. When proteins are needed in a cell, the transcription occurs. In the transcription step, where the information of DNA is transferred to a messenger RNA (mRNA) molecule, the enzyme called RNA polymerase II catalyzes the formation of an mRNA molecule. The single stranded mRNA carries the genetic information which is thus translated into protein.

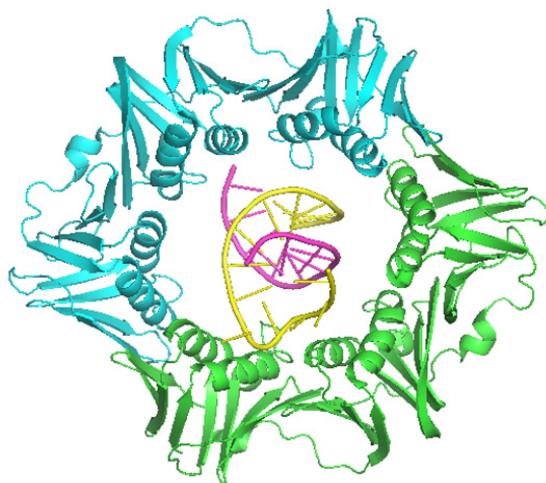


Figure 1.1. The X-ray structure of the *E. Coli* beta clamp polymerase processivity factor (green and light blue for  $\alpha$  and  $\beta$  domains) with double stranded DNA through the central pore. (PDB entry: 3bep [3]). Picture was generated in Pymol [4].

The processes of transcription and replication are triggered by binding of regulatory proteins to a particular DNA region. Synthetic/natural molecules can block the genes which contain wrong genetic information, thereby obstruct the transcription and replication processes, restricting protein synthesis, leading cell death. These types of drugs exhibit antitumor and antibiotic activities.

In general, drugs bind to DNA via covalent and non-covalent interactions: covalent binding is an irreversible type of binding, leading to complete inhibition of DNA processes and subsequent cell death [5]. The coordination complex *cis*-[Pt(NH<sub>3</sub>)<sub>2</sub>(Cl)<sub>2</sub>] makes an intra/interstrand cross-

link by substitution of the chloro groups with the nitrogens on DNA bases, and is widely used in clinical treatment of cancer since its discovery in 1965 [6]. Reversible binding of DNA by proteins and small molecules can be divided into mainly two categories: intercalation and groove binding. Intercalation is insertion of a molecule or part of a molecule in between two adjacent DNA base pairs. A large number of drugs with anti-tumor activity [7] intercalate selectively with the recognized specific regions of DNA. These intercalators share similar structural properties: planar heterocyclic groups stack between adjacent DNA base pairs. The  $\pi$ - $\pi$  stacking interactions between the drug and bases stabilize the binding. When being intercalated, DNA exhibit severe structural perturbations. An example of intercalation reagent, nogalamycin [7] is shown in Figure 1.2.

Minor groove binding drugs also possess anti-tumor activities. And they are usually crescent shaped, which complements the shape of the groove and facilitates binding by promoting van der Waals interactions. Hydrogen bond formations between the minor groove binders and the N3 of adenine and of thymine are usually present. The binding prefers to occur at the AT rich region, this is not only because of the hydrogen bond formation but also due to the increased electronegative potential. Moreover, comparing to the A/T region, the steric hindrance in the minor groove at G/C regions, presented by the C2 amino group of the guanine base makes the A/T regions attractive. Examples of minor groove binders including natural antibiotic distamycin [9-11] and synthetic DNA stain DAPI [12-14] are shown in Figure 1.3.

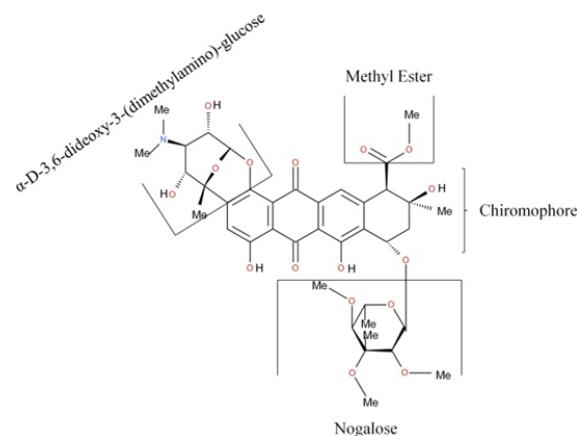


Figure 1.2. The structure of nogalamycin  $C_{39}H_{49}NO_{16}$ . This image and the following chemical structures were produced in Accelrys Draw 4.1 [15].

In this thesis, the interaction of binuclear ruthenium compounds and DNA are studied by NMR and optical spectroscopy. Ruthenium polypyridyl chemistry in the context of selectively binding to DNA has attracted scientists' attentions for many years. The most recent X-ray structures described the intercalation mode of mononuclear ruthenium compounds to DNA in detail [16-17]:

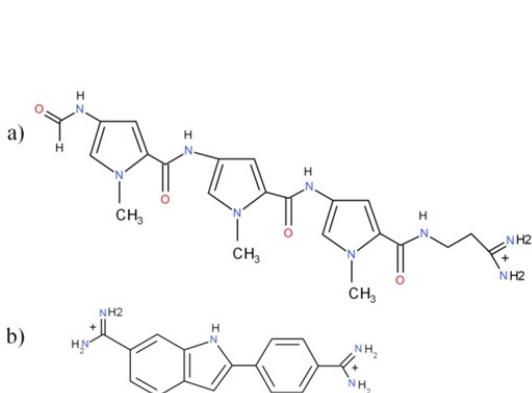


Figure 1.3. The DNA minor groove binders: a) distamycin, (b)DAPI (4',6-diamidino-2-phenylindole).

the intercalation is sequence dependent, additional binding sites were also explored. In our group, dumb-bell shaped binuclear Ru compounds had been shown to thread intercalate its bulky group into long alternating AT sequences [18]. However, up till now, no 3D structure of *binuclear* ruthenium compounds intercalation to DNA has been published. Distinct from the monomers, which slide the bridging ligand in between two base pairs [16-17], in order for the binuclear compounds to thread intercalate, structural rearrangements and base pair opening are considered as prerequisite [19]. Linear dichroism showed that the  $\Delta,\Delta$  enantiomer of  $[\mu\text{-}(\text{bidppz}(\text{phen})_4\text{Ru}_2)]^{4+}$  can thread intercalate into calf-thymus DNA from an initial binding mode to intercalation, and the rate of the intercalation processes was observed to be slow (two weeks at room temperature) [20]. In this thesis, we demonstrate the first surface bound structure of a binuclear Ru compound  $\Lambda,\Lambda\text{-}[\text{bidppz}(\text{bipy})_4\text{Ru}_2]^{4+}$  with a 12-mer Dickerson-Drew DNA  $[(\text{CGCGAATTGCG})_2]$  by NMR spectroscopy (*Paper I*): One of the Ru centers locates near the AT region, while the other part making hydrogen bonds to the guanine at the end. In a following up study, a T-T mismatched DNA was designed, in which the mismatched region creates space with four potential opening AT base pairs for the bulky centers to thread in (Figure 1.4). Spectra analysis is still ongoing; however, preliminary results indicate that one Ru compound binds to two DNA duplexes. X-ray studies of a second binuclear Ru analogue ( $\Delta,\Delta\text{-}[\mu\text{-}(\text{bidppz}(\text{phen})_4\text{Ru}_2)]^{4+}$ ) shows that the compound has intercalated into a hexamer short oligo. (*Paper II*) With one monomer inserted to one duplex, extruding AT base pairs, leaving the other one stacked into second duplex at the end. Unfortunately, solution NMR study was not possible, because of the formation of aggregates. Studies on the concentration dependence in solution indicate that the intricate hierarchy of intermolecular contacts found in the solid state is step-wise formed already in dilute solution.

Besides Ru-compounds, the binding properties of Hoechst 33258, a minor groove binder, are also explored in this thesis: two Hoechst 33258 (*Paper III*) molecules were shown to consecutively bind in the minor groove, one after another. The binding of the second Hoechst to  $\text{A}_4\text{T}_4$  is not cooperative and that the molecules are sitting with a small separation apart, and not in a sandwich structure as previously proposed.

Although DNA is the carrier of genetic information in a cell, however, proteins are paramount important biomacromolecules which are essential for cell functions. Proteins are long chains containing 20 different kinds of amino acids. Each cell contains thousands of different proteins and each of them has its unique function: e.g. enzymes catalyze nearly all chemical processes in cells; structural components give cells their shape and help them to move; hormones transmit signals throughout the body; antibodies recognize foreign molecules; transport molecules carry oxygen etc. Except of extensive structural studies on drug-DNA interactions, investigations on protein structural alterations, and ligand-protein interactions are also studied (*Paper IV*): lysozymes are hydrolytic enzymes, which cleave the  $\beta$ -(1,4)-glycosidic bond between N-acetylmuramic acid and N-acetylglucosamine, the major bacterial cell wall material. Its

antimicrobial activities have been used as preservatives in food and pharmaceutical [21]. The deprotonation of the catalytic residue Glu 35 is a prerequisite for its activity. We found that the deprotonation process is coupled with protein structural changes. The latter was found to be induced by ligand binding as well.

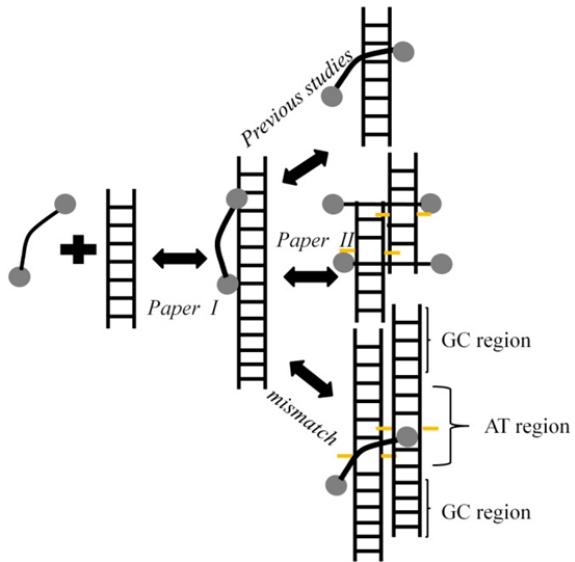


Figure 1.4. Schematic picture of overview paper I, II and the ongoing mismatch DNA project: dimeric ruthenium compounds firstly bound outside (paper I), and then inserted into a hexamer, with predictions from previous optical spectroscopy studies. Current analysis results of the mismatch DNA project indicate the compound interacts with T in the middle of the mismatch DNA. Orange lines represent flipped base pairs in the grooves.

## 2 BACKGROUND

The main theme in this thesis is structural analysis of interactions between biomacromolecules such as DNA, protein and small molecules. This chapter explains that why these biological molecules are among the highest interests in the context of ligand interactions. The motivations for selecting these small molecules are also illustrated. Summaries of some important results from earlier research are referred here.

### 2.1 DNA

#### 2.1.1 DNA duplex structure

DNA is a linear polymer with nucleotides as building blocks, which in turn are formed from a planar heterocyclic base, a sugar (deoxyribose) and phosphate acid. The most relevant conformation under physiological conditions for DNA is B-form of double stranded DNA. B-DNA is a fairly uniform right-handed double-helical structure, the two antiparallel deoxyribonucleotide chains twist around each other. Alternating phosphate and deoxyribose sugar groups constitute the backbone of DNA. In B-DNA, the helix makes a turn every 3.4nm, and the distance between two neighboring base pairs is 0.34nm and there are 10 base pairs per turn. On the deoxyribose group, one of the four bases adenine (A), guanine (G), cytosine (C) or thymine (T), distinguishes the four nucleotides from each other (Figure 2.1). The structure is stabilized by intermolecular hydrogen bonds between complementary Watson-Crick base pairs, and by hydrophobic interactions that keep the apolar surfaces of the bases away from the surrounding aqueous environment while exposing their polar edges and the phosphates to the solvent. The asymmetry of the nucleotides gives rise to two grooves with different geometric attributes (Figure 2.2). The major groove is wider and relatively shallower than the minor groove.

In B-DNA, the planar of the bases are all virtually perpendicular to the helix axis. DNA exists in other forms (A and Z) under different conditions, e.g. in a solution with higher salt concentration or with alcohol added, the DNA structure may change to an A form, which is still right-handed, but since the base-pairs are tilted, every 2.3nm makes a turn with 11 base pairs per turn. Which conformation that DNA adapts depends on the hydration level, DNA sequence, and chemical modifications of the bases, the type and concentration of metal ions *etc.*

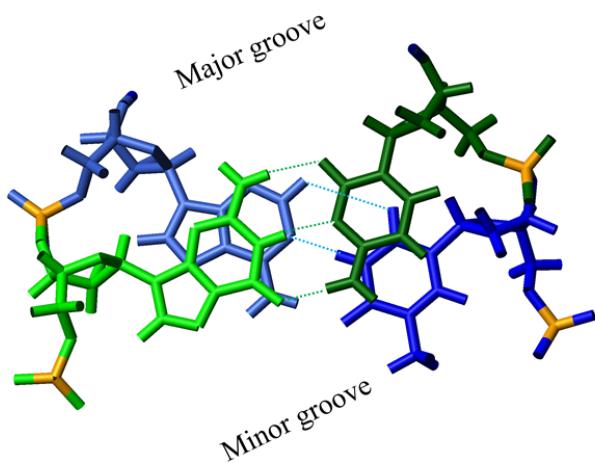


Figure 2.1. The Watson-Crick base paring scheme. Aadenin (dark blue) and thymine (light blue) are pared through two hydrogen bonds (cyan dashed lines). Three hydrogen bonds (green dashed lines) are formed between guanine (green) and cytocine (darkgreen). Image was produced in MOLMOL [22].

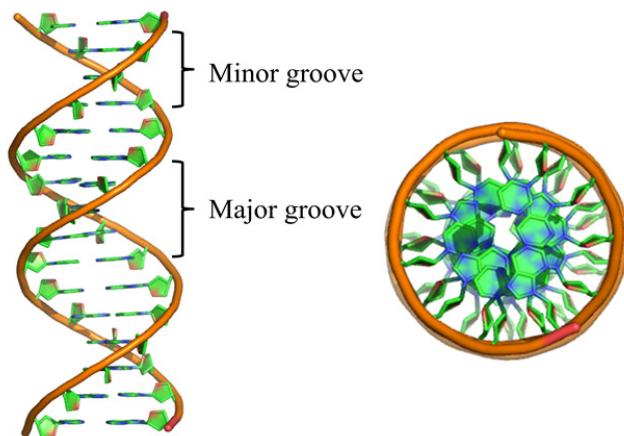


Figure 2.2. Side- (left) and top- (right) view of B-DNA, two grooves are indicated in 3D structure. DNA backbone is in orange. Image was produced in PyMOL [4].

### 2.1.2 Mismatch DNA

Although DNA is quite stable molecule, DNA damage in the cell can occur at different circumstances. Endogenous, which is induced by reactive oxygen species that result from metabolic byproducts; exogenous include e.g. ultraviolet radiation from the sun, exposure to X-ray, gamma rays and toxins, human-made mutagenic chemicals. Errors that occur during normal DNA metabolism or abnormal DNA processing reactions, such as in DNA replication,

recombination and repair are dangerous, since they will lead to mutations and potentially malfunctioning proteins, eventually to cellular dysfunction and diseases.

Structurally, mismatched base pairs form hydrogen bonds, and the double helical structure may still be preserved (Figure 2.3) [23-26]. However, mismatch containing oligonucleotides are thermodynamically less stable, and the melting temperatures decrease notably with the exception for G-containing mismatches. [27-29] Increased dynamic motion and longer base-pair opening lifetimes are also observed for mismatches in NMR [27, 29-30].

DNA mismatch repair (MMR) is an important system which recognizes and repairs the erroneous insertion, deletion and mis-incorporation of bases that can arise during DNA replication and recombination. MMR also repairs some forms of DNA damage [31]. Erroneous genes such as mismatches lead to cancer and genetic diseases. When cancer occurs, the MMR system can hardly repair the erroneous genes back to the correct ones. At this stage, small synthetic compounds can bind to and thus block the erroneous genes, finally leading to cancer cell death. One of the goals from this work is to investigate the structure of a binuclear ruthenium compound bound to a mismatched DNA (see Results).

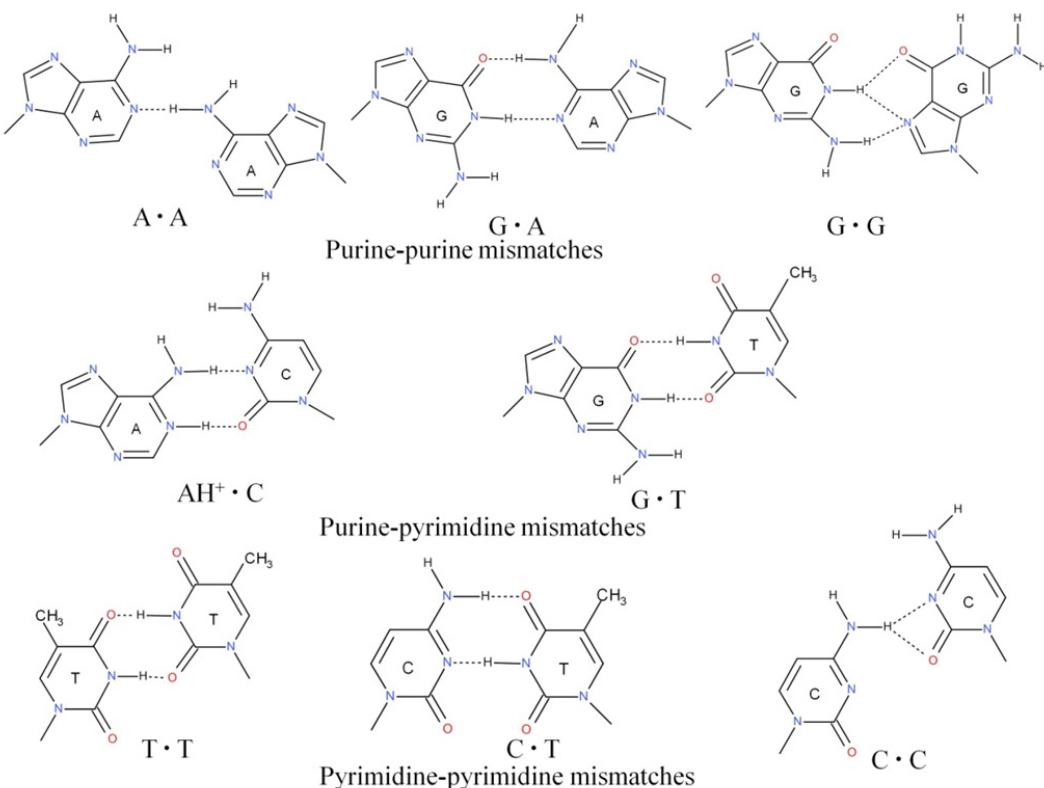


Figure 2.3. Hydrogen bonding in base mismatches [23-27].

## 2.2 Ruthenium complexes and Hoechst 33258

Ruthenium locates below iron in the periodic table. The octahedral coordination geometry around the ruthenium ion leads to two enantiomeric configurations for tris-chelates with bidentate ligands, a right-handed ( $\Delta$ ) and a left-handed ( $\Lambda$ ) ‘propeller’ shape (Figure 2.4). Due to their strong DNA binding affinities and useful photophysical properties, these types of transition metal compounds have been extensively studied. In the mid 1980’s, one of the classical Ru-compounds  $[\text{Ru}(\text{phen})_3]^{2+}$  was proposed to enantiospecifically intercalate into DNA, although this suggestion remained controversial for many years [32-38]. Different DNA-binding properties are indeed expected for  $\Delta$  and  $\Lambda$  enantiomers upon intercalation, since their ancillary (non-intercalating) ligand would fit differently into the grooves of helical DNA [35-36, 39]. Later, in the early 1990’s,  $[\text{Ru}(\text{phen})_2\text{dppz}]^{2+}$ , a ‘light-switch’ compound was synthesized by Sauvage and co-workers [40]. Comparing with  $[\text{Ru}(\text{phen})_3]^{2+}$ , the new type of compounds intercalate into DNA much more distinctly by inserting the large dipyridophenazine (dppz) ligand in-between the base pairs. The intercalation is accompanied by a significant luminescence quantum yield increase [41-42] of the metal-to-ligand charge transfer transitions. When  $[\text{Ru}(\text{phen})_2\text{dppz}]^{2+}$  is free in aqueous solution, the luminescence is quenched through solvent hydrogen bonding to the phenazine nitrogen atoms on dppz ligand, but restored again when protected from water inside the intercalation pocket [43-44]. Despite the general agreement on that dppz-compounds intercalate into DNA, the question whether they are intercalated from the major or the minor groove have been lively debated for many years [41, 44-48]. Only very recently, X-ray crystallography have revealed the intercalated 3D structures of three different dppz-compounds: ( $\Delta$ - $[\text{Ru}(\text{bpy})_2\text{dppz}]^{2+}$ ,  $\Lambda$ - $[\text{Ru}(\text{phen})_2\text{dppz}]^{2+}$  and  $\Lambda$ - $[\text{Ru}(\text{TAP})_2\text{dppz}]^{2+}$  (TAP=tetraazaphenanthrene) into various DNA sequences, all showing that the intercalation takes place from the minor groove. [16-17, 49].

In the last decades, the majority of studies were focused on monomeric compounds. To increase the selectivity and decrease the dissociation rate, which is considered important for antitumor activity [50], binuclear ruthenium compounds were introduced in the early 2000’s. One of the first synthesized binuclear compounds,  $[\mu\text{-C4}(\text{cpdppz})_2(\text{phen})_4\text{Ru}_2]^{4+}$  (Figure 2.4) shows slow association and dissociation processes ascribing from threading bis-intercalation [51-53]. In the context of selectivity, its  $\Delta\Delta$  enantiomer displays higher association rate with alternating AT-DNA than with alternating GC-DNA. Later, semi-rigid binuclear ruthenium compounds like  $\Delta\Delta\text{-P}$  (Figure 2.4), in which a single bond connects the two monomers instead of a flexible linker, were found to initially bind in one of the grooves, based on a positive linear dichroism spectrum (LD) and absence of luminescence. However, after being left at room temperature for two weeks,  $\Delta\Delta\text{-P}$  was revealed to be intercalated into calf thymus DNA based on negative LD, strong luminescence, and extremely slow dissociation kinetics [18]. Comparing with  $[\mu\text{-C4}(\text{cpdppz})_2(\text{phen})_4\text{Ru}_2]^{4+}$ , the intercalation had evidently not been prevented by shortening the linker between two monomers, but the intercalation rate was dramatically decreased. Raising

temperature as well as increasing salt concentration accelerates the intercalation rate. Later, studies showed that  $(\Lambda,\Lambda)$  and  $(\Delta\Delta)$  stereoisomer of P also thread intercalate into [ploy(dAdT)]<sub>2</sub> after a surface bound state [54]. The intercalation rate not only depends on the structure of the compound [18], but also on the DNA sequence: a certain length of a stretch of AT base pairs is required for fast intercalation, a dramatic drop in the threading rate was observed when the number of consecutive base pairs (bp) with alternating AT sequence was decreased from 14 to 10 bp, and no threading seems to occur for a stretch of six AT base pairs [55].

Interaction studies of the same  $\Delta\Delta$ -P with loop-containing DNA showed the large loophole is the advantage to allow bulky Ru center to thread in, the activation barriers of threading intercalation being significantly lowered compared to normal DNA. The rate of intercalation depends on the numbers of base pairs that at the loop regions [56]. Based on this finding, studies of mismatched containing DNA are attempted. The ultimate goal is to create enough space for the bulky center to thread in.

Due to the slow threading intercalation process, the surface bound state can be regarded as a kinetically trapped intermediate. Both the surface bound and the intercalation states are characterized in this thesis work. (*paper I and II*).

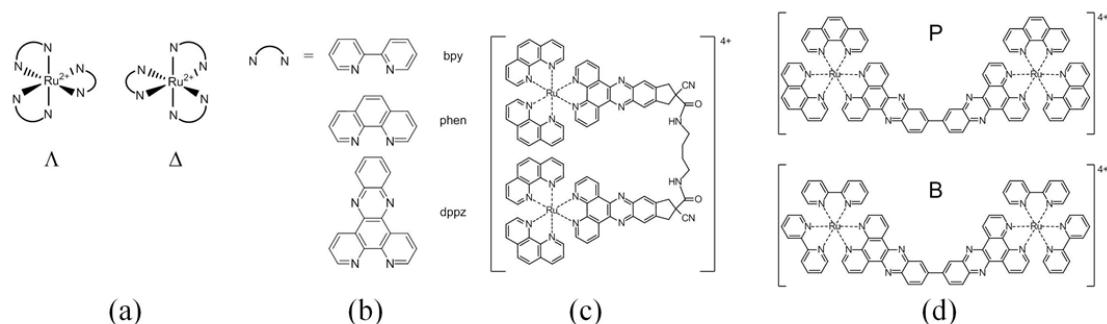


Figure 2.4. (a): the octahedral coordination of the ruthenium center gives rise to two stereoisomeric forms, the left-handed ( $\Lambda$ ) and the right-handed ( $\Delta$ ) propeller. (b): commonly used bi-dentate ligands: bpy=2,2'-bipyridine, phen=1,10-phenanthroline dppz=dipyrrolo[3,2-a:2',3'-c]phenazine. (c) and (d): structure of binuclear ruthenium compounds. Bis-intercalating  $[\mu\text{-C4}(\text{cpdppz})_2(\text{phen})_4\text{Ru}]^{4+}$  (c) and threading  $[\mu\text{-}(\text{bidppz})\text{L}_4\text{Ru}^{2+}]^{4+}$  (d), where L is either phen or bpy (d). C4(cpdppz)<sub>2</sub>=N, N'-bis(12-cyano-12,13-dihydro-11H-cyclopenta[b]dipyrrolo[3,2-h:2',3'-j]phenazine-12-carbonyl)-1,4diaminobutane. Bidppz=11'11-bis(dipyrrolo[3,2-a:2',3'-c]phenazinyl).

Besides the above mentioned studies which are focused on Ru-compounds, other synthetic molecules are also demonstrated in the context of interaction with DNA in the present work. It is known that low molecular weight small molecules can easily pass through the cellular membranes and reach the DNA. Hoechst 33258 [57], a bis(benzimidazole) (Figure 2.5) derivative showing antitumor and antibiotic properties, is such a compound although its toxicity

has prohibited a widespread medical use [57-60]. Similarly to dppz containing Ru compounds, Hoechst 33258 fluoresces brightly when bound to double-stranded DNA, but is quenched when free in buffer solution. The photo-properties of Hoechst 33258 are utilized in microscope imaging [61-63]. For example, Hoechst molecules have now been used as nuclear stains, which can bind to DNA in live or fixed cells by permeation. Studies have shown that it binds strongly in the minor groove of the DNA preferably to the sequence that is rich in adenines and thymines [64-65]. The binding preference ascribes from its crescent-shape as well as its flat hydrophobic aromatic rings which creates a snug fit between the hydrophobic sugar walls in the minor groove. Moreover, Hoechst 33258 is positively charged, matching the electronegative potential at the floor of the minor groove of AT regions [66], which is found to be more negative than GC regions. The contents of H-donors and acceptors make them form H-bonds to adenine and thymine in the minor groove [65, 66-67].

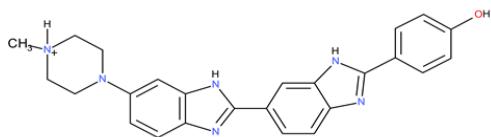


Figure 2.5. Structure of Hoechst 33258 (neutral pH).

## 2.3 Lysozymes

### 2.3.1 Different types of lysozymes

Lysozyme was discovered in 1921 by the Scottish bacteriologist Sir Alexander Fleming (1881-1955) who also discovered penicillin in 1928. Lysozyme not only serves as a model system in protein chemistry, enzymology, crystallography and molecular biology but also contributes to its antibacterial defense in animals. It is further used as a preservative in foods and pharmaceuticals [21]. Lysozyme hydrolyses the  $\beta$ -(1,4)-glycosidic bond between the alternating N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG) residues of peptidoglycan, the latter being the unique bacterial cell wall polymer (Figure 2.6) which determines the shape of the cells and provides protection against cellular turgor pressure. Lysozymes occur in all major living organisms. It is widely distributed in secretions, body fluids, and tissues of human and animal organisms; it has also been isolated from some plants, bacteria, and bacteriophages. In the eggs of birds, lysozyme serves as a defense reagent before the embryo reaches the ability to produce immunoglobulin. In animal kingdom, three major different lysozyme types have been identified, commonly designated as the c-type (chicken or conventional type), the g-type (goose-type) and more recently, also the i-type (invertebrate type). Different types of lysozymes owe to their differences in amino acid sequences, biochemical and enzymatic properties.

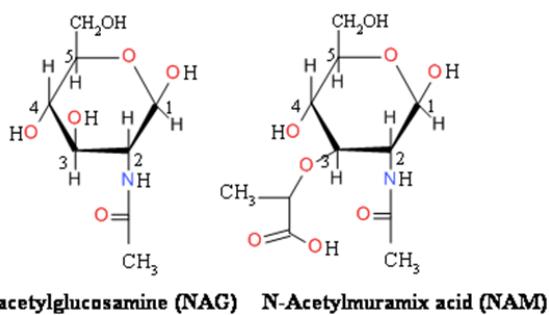


Figure 2.6. Bacterial cell wall material: NAG and NAM

C-type lysozymes are the major types of lysozymes, produced by most vertebrates, including mammals. Besides mammals, other vertebrates such as birds, fish, reptiles and amphibians produce c-type lysozymes or at least contain c-type lysozyme genes. Moreover, other c-type lysozyme homologues can be found in both the Coleoptera and Hymenoptera orders. Due to the spread of c-type lysozymes among insects, lysozyme is present in six out of twenty-seven known insect orders. The goose-type lysozyme was initially identified from egg whites of the Embden goose. Since then it has been found in several avian species including chicken, black swan, ostrich, cassowary and rhea. Initially g-type lysozyme was considered to be restricted to vertebrates. However, functional g-type genes have recently been also identified in invertebrates such as some bivalve mollusk scallops and in members of the tunicates. The third type, i-type lysozyme (*TjL*, *Tapes japonica* lysozyme) is determined to be the invertebrate type. In general,

the overall amino acid identities of these three c-, g-, and i-type lysozymes are low. Typical g-type lysozymes are significantly larger (~20-22 kDa) than c- and i-type lysozymes (~11-15 kDa). In general, c- and g-type lysozymes are basic proteins in contrast to i-type lysozymes. Some of c-type and i-type lysozymes that have evolved to function as digestive enzymes are notable exceptions.

### 2.3.2 Catalytic mechanisms of lysozymes related to structures

Although the similarity in primary structure between the three lysozyme types is limited, their three-dimensional structures show striking similarities. All three types of lysozymes possess  $\alpha$ -helices and  $\beta$ -sheet domains (Figure 2.7), as well as a large cleft for substrate binding, similar to the structure of hen egg white lysozyme (HEWL). The X-ray structure of HEWL was the first three-dimensional protein structure determined [68]. The active site of both HEWL and HL consist of six subsites A, B, C, D, E and F, which bind to six consecutive sugar residues. Four helices (A, B, C and D) are marked with different color (below). Catalytic residues Glu 35 and Asp 52 locate at binding site D.

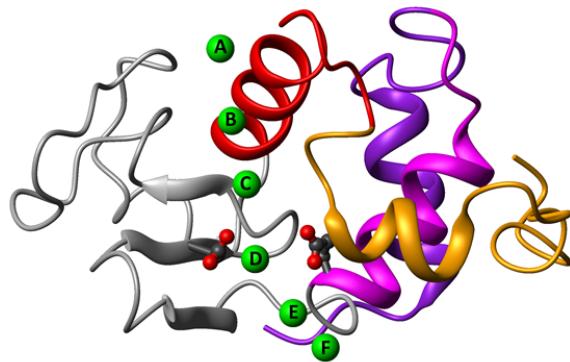


Figure 2.7. X-ray structure of HEWL (PDB entry: 2vb1 [69]). Two domains  $\alpha$  and  $\beta$  are separated by the binding cleft. Important residues Glu 35 ( $\alpha$ -domain) and Asp 52 ( $\beta$ -domain) are displayed. Six binding sites (A-F) locate in between two domains. Four helices are marked with purple (helix-A, resi. 5-14), magenta (helix-B, 25-35), red (helix-C, 89-98) and orange (helix-D, 109-115).

The crystal structure of HEWL-(NAG)<sub>3</sub> complex [70] revealed that the target C-O bond is placed in the vicinity of two potential catalytic groups, Glu 35 and Asp 52 (Figure 2.7). The mechanism of this reaction is revealed in 2001 by Vocadlo *et al* [71] in X-ray crystallography. It has been found that the  $\beta$ -(1,4)-glycosidic bond between NAM and NAG is hydrolyzed by a double displacement reaction (Figure 2.8). In the first step (A), the carboxylate group of Asp 52 acts as a nucleophile to form the glycosyl intermediate, which leads to an inversion of configuration. Meantime, Glu 35 acts as a general acid, donating a proton to the glucosidic oxygen, which then facilitates bond cleavage. Secondly, water molecule removes the enzyme carboxylate from glycosyl-enzyme intermediate, following with configuration inversion and thereby restoring the original conformation. The interactions between acetyl groups from the hexasaccharide glycan

strand with the amino acids in the long groove of HEWL are important for substrate binding. This is confirmed by the absent of enzymatic activity of HEWL resulted from interaction with deacetylated peptidoglycan [71].

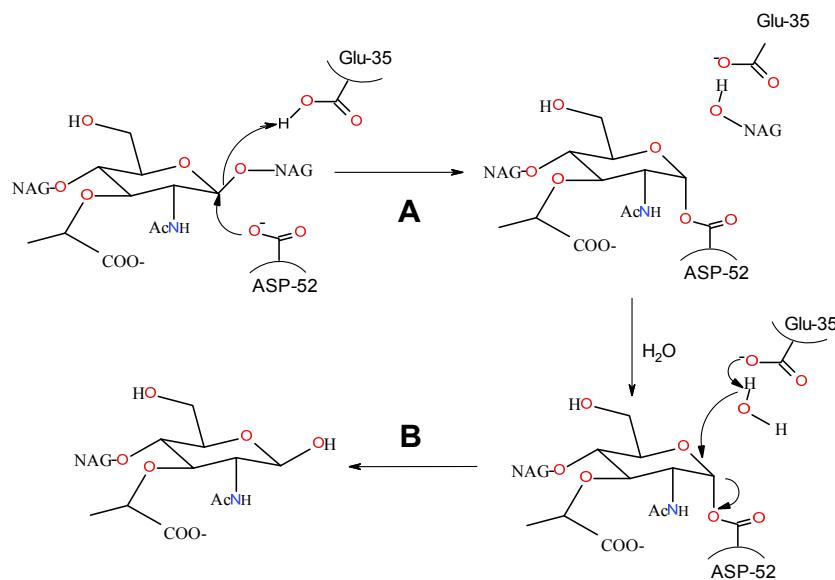


Figure 2.8. Scheme of double displacement catalytic mechanism of HEWL.

The hydrolyzation mechanism suggests that the protonation state of Glu 35 is a prerequisite for the enzyme activity; therefore, much work has been focused on the pKa measurement of Glu 35. Earlier reports have shown Glu 35 possess uncommonly high pKa, 6.8 for HL and 5.9 for HEWL [72]. The high pKa value of Glu 35 is considered not due to the hydrophobic environment but to the electrostatic interaction of ionized Glu 35 and ionized Asp 52. Nevertheless, in the same work, it also indicated that the active-site movement is accompanied with the binding of N-acetyl-chitotriose to HEWL.

Human lysozyme (HL) is constructed by 130 amino acids, and belongs to c-type class. Structure of HL is in close similarity to that of HEWL, HL also consists of four helices A (residues 5-14), B (25-35), C (90-99) and D (110-116) in  $\alpha$  domain, and three-stranded antiparallel sheets (42-46, 51-56, 59-61) in the  $\beta$ -domain. It is widely distributed in different organs, such as in liver, articular cartilage, plasma, saliva, tears and milk [73]. The normal concentration in the blood ranges from 4 to 13mg/L and it can be only traced from the urine of healthy subjects. Large amount of lysozyme (500mg) are produced daily, but its lifetime is very short. Approximately 75 percent of the total is eliminated within 1h via the kidneys [74]. HL, along with HEWL has been studied in the context of protein structure and function, including the mechanism of protein folding and protein stability [75]. Comparisons between the structure of lysozyme-(NAG)<sub>6</sub> complex [76] and the structure of lysozyme alone, which was determined by Artymiuk and Black [77], revealed significant shifts of the backbone atom at region Asp102-Asp120 on helix-D and the proceeding loop towards residues Cys 30-Thr 70. This ‘closing’ construction of the

active site cleft leads to tight adhesion with polysaccharides [76]. Interestingly, later NMR dynamic measurements have shown that the motion of the same region at active site (Val100-Arg115) in human lysozyme is different from that of HEWL upon binding to (NAG)<sub>3</sub>, and these differences may be involved in the improvement of the enzymatic activity [78]. Moreover, in the same paper, it was determined that binding substrate stabilizes the motions of the residues at the binding sites, while other residues exhibit higher order of flexibilities. This is because that the binding results in an entropy penalty, and this penalty is compensated by other residues motilities. Similar movement of the active site was also shown in a temperature study without ligand binding at 35° and 4° by NMR spectroscopy, structural comparison shows that low and high temperature can induce the active cleft ‘closing’ and ‘opening’ forms respectively [79-80]. The structural movements were also observed while pressure alterations [81]. The correlation between the deprotonation of Glu 35 and the structural changes of helices of human lysozyme is studied by extensive 1D and 2D NMR titration experiments in this thesis.

### 2.3.3 Biological role of lyosyzmes

As mentioned above, lysozymes are wide spread in animal kingdom. Multiple types of lysozymes are produced in most organisms, and same types of lysozyme may even possess different functions. The antibacterial activity of lysozyme is well established and appeared to be environment dependent: human lysozymes suppress directly *E.coli* growth in saliva [82] but not in breast milk, this indicates that the antibacterial activity of lysozyme maybe important in some tissues and secretions but not in others; however, human lysozyme from transgenic goats lowered the *E.coli* population in the gut of young pigs fed with this goat milk [83-84]. The role of lysozyme in the antibacterial activity of airway secretion was confirmed by the restored antibacterial activity of the cationic polypeptide-depleted nasal fluid [85]. Moreover, lysozymes in birds, fish and reptiles are mostly specialized to protect eggs (with activity also found in the gill, skin and gastrointestinal tract in fish). In insects, an important role of lysozyme is protection during the vulnerable metamorphosis stage. Lysozyme is a safe and commonly used protein in some foods [21], pharmacological and therapeutic applications: it is used in food industry as a natural food preservative in cheese, wine and beer. Lysozyme is also used in oral health care products such as toothpaste, mouthwash, and chewing gum [86].

Except of the well-known antibacterial activity, some i-type lysozymes also possess isopeptidase activity [87]. For example, they hydrolyze isopeptide bonds formed between fibrin molecules in the final steps of blood coagulation by specific transglutaminase enzymes. Since isopeptidases can reverse the critical final step in blood coagulation, lots of attentions have been paid in the context of inhibition of thrombus formation. It is assumed that both lysozyme and isopeptidase activities are found in all i-type lysozymes. Other biological functions such as anticancer [88], antiHIV [89] are also documented.

### 3 THEORY AND METHODOLOGY

Interactions between biomacromolecules (e. g. DNA and proteins) and synthetic compounds are investigated in this thesis. In the context of DNA-ligand binding, the intercalation mode of dimeric ruthenium compounds to DNA has been suggested from earlier studies by optical spectroscopy measurements. To be able to understand and characterize the mechanism of this unusual interaction processes, 3D structural studies are essential. Two-dimensional NMR spectroscopy was utilized. Besides NMR, other methods, e.g. circular dichroism, fluorescence and absorption also provided valuable results. This chapter describes the principles of the above mentioned techniques and also explains how to interpret results in the following chapter.

#### 3.1 NMR Spectroscopy

Back in 1924, W. Pauli proposed the theoretical basis for NMR (Nuclei Magnetic Resonance) spectroscopy. He found that certain atomic nuclei have properties of spin and magnetic moment. When nuclei are exposed to magnetic field, their energy levels will be split. From then till now, NMR has been proved to be a remarkably powerful tool. Large ranges of experimental schemes were developed to extend NMR applications to characterize complex molecules. In the context of genetic engineering, multidimensional NMR enables the observations of stable isotope labeled signals in living microorganisms.

NMR spectroscopy is based on the absorption by nuclear spins of electromagnetic radiation in the ratio-frequency region; macromolecular NMR typically uses a range from 100MHz to 1GHz. In contrast to UV and IR absorption spectroscopy, which involve outer-shell atomic electrons, NMR arises from the magnetic properties of atomic nuclei. As mentioned above, when placed in an intense magnetic field, the energy of some nuclei split into different energy states, allowing absorption and emission of electromagnetic radiation to occur.

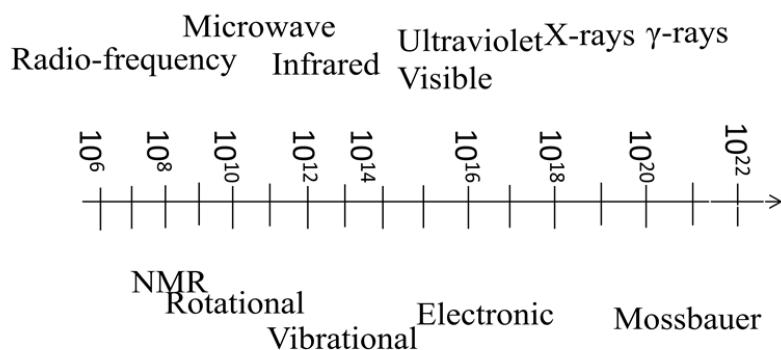


Figure 3.1. The electromagnetic spectrum from  $\gamma$ -rays ( $\nu=10^{22}$  Hz) to radio-frequency radiation ( $10^4$  Hz).

Figure 3.1 shows the electromagnetic spectrum from the radio-frequency region ( $10^4$  Hz) through microwaves, IR, visible, UV and X-rays to  $\gamma$ -rays ( $10^{22}$  Hz). The frequency that NMR utilizes is at the lower end of the spectrum, this is because of the very small energy splitting of nuclei, even placed in strong magnetic fields.

### 3.1.1 Basic concepts: Magnetization, Boltzman equilibrium, FID, and chemical shifts.

Atoms consist of nuclei surrounded by electrons. The nucleus of a hydrogen atom, a single proton, has the quantum-mechanical property of spin=1/2; in a magnetic field this results in two spin states, spin up and spin down. In biomolecular NMR, only spin 1/2 nuclei are considered. The energy difference between these two spin states is characterized by  $\Delta E = \hbar\gamma B_0$ , where  $\gamma$  is the gyromagnetic ratio, a proportionality constant characteristic of the isotope being examined. And  $\hbar$  is the Plancks constant;  $B_0$  is the strength of a magnetic field. The energy between these two spin states determines the population of these states. For any system of energy levels at thermal equilibrium, there will always be more particles in the lower energy state than in the upper energy state according to the Boltzmann distribution.

When samples are placed in side of the static magnetic field  $B_0$ , individual nucleus experiences different local magnetic field  $B$  ( $B$  is the local magnetic field experienced by an individual nucleus, discussed later in the chemical shift section). The strength of  $B$  is dependent on the local environment (see below, chemical shift). The spins which experience the same strength of  $B$  are confined to cones, both parallel and antiparallel to  $B_0$ , with a fixed angle around the field direction. Due to the Boltzmann equilibrium, nuclei are distributed on the two cones with the random location on each cone, all spins for one type of nuclei add up to a vector along the magnetic field direction (Figure 3.2 A); this vector is called magnetization  $M$ . By applying short pulses of a radio frequent field ( $B_1$ ) perpendicular to the static field, the magnetization can be tilted away from the equilibrium state by a certain angle. This  $B_1$  pulse is normally applied in microseconds, this brief pulse covers a range of frequencies. The power of the pulse must be sufficient to cause all sets of nuclei that are submerged into the same local magnetic field  $B$ , precessing at frequencies within the SW (spectral width, the range that the pulse should cover) to become phase coherent, each set of nuclei that experience the same local  $B$  tips their  $M$  vectors toward the  $x,y$  plane. After a  $90^\circ$  short pulse, the magnetization vectors then start to precess about the field axis with the so-called Larmor frequency ( $\omega=\gamma B$ ). The rotating magnetization vectors (Figure 3.2B) generate currents in the receiver coils along the  $x$ - or  $y$ -axis according to Faraday induction: A current is induced in the wire by the movement of the magnetic field near the wire. This current is sampled at discrete times and digitized. The  $x$  and  $y$  components of this magnetization are written as:  $M_x(t) = M\cos\omega t$  and  $M_y(t) = M\sin\omega t$ . This signal will eventually decay due to the action of relaxation. After the  $B_1$  is turned off, not only do the populations of the two states revert to the Boltzmann distribution, but also the individual nuclear magnetic

moments begin to lose their phase coherence and return to a random arrangement around the z axis.

The NMR signal is generated by monitoring the induced ac (alternating current) signal decay in the receiver coil as a function of time; it is called the modulated free induction decay (FID). This FID signal in time-domain is turned into a spectrum (frequency-domain) by a mathematical method called Fourier transformation (Figure 3.2 C). As a consequence of the way the Fourier transformation works, it is also convenient to regard  $M_x(t)$  and  $M_y(t)$  as the real and imaginary parts of a complex signal  $M(t)$ . A complex FID yields a complex spectrum. For the final spectrum, only the real points are shown.

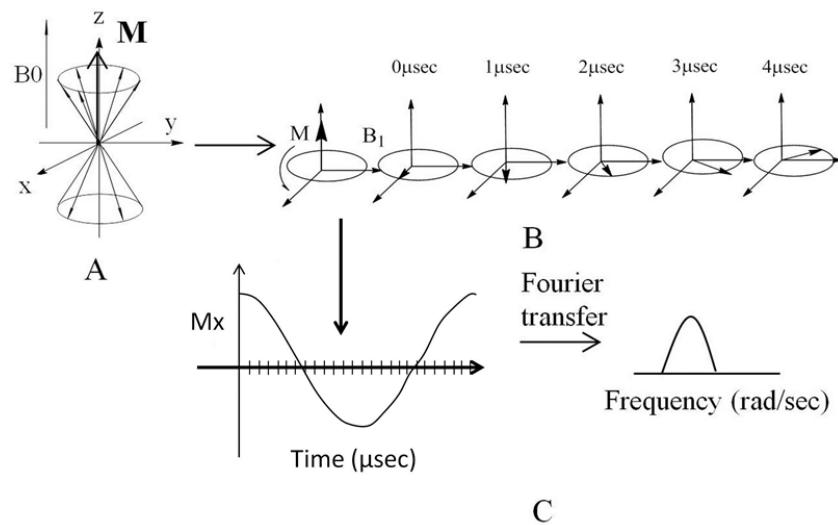


Figure 3.2. A: two magnetization cones indicate two spin states: spin up and spin down, small arrows are individual spins. The big  $\mathbf{M}$  is the magnetizations, which is the sum of the spins. B: magnetization precessing: arrows indicate the magnetization  $\mathbf{M}$ . C: Fourier transformation turned the signal (cosine function) from time domain to frequency domain.

The absorption frequency of a nuclear spin depends on the magnetic field strength at the nucleus. This magnetic field usually differs slightly from the applied field,  $B_0$ , this is because of the shielding effect by electrons of nearby electronegative atoms. Another modification of the local magnetic field is due to delocalized  $\pi$ -electrons of nearby aromatic rings. The field  $B_0$  induces a ring current, which in turn generates an additional magnetic field that is opposite to the external applied magnetic field. Since the chemical shift is a relative measure of Larmor frequency, and as mentioned above, the Larmor frequency is proportional to local magnetic field. In turn, the surrounding magnetic field effects the actual chemical shift of nuclei. An example for this is the ring current effect: placement of a proton 4 Å directly above the center of an aromatic ring (e. g. phenylalanine ring in Figure 3.3) will decrease the chemical shift by approximately 0.75 ppm. However, the local field strength for the nuclei is decreased to  $B_0 - B'$ . In contrast, if the atom is

placed in the plane of the ring at the same distance from the center, its chemical shift will be increased by approximately 0.33 ppm [90], and the local field  $B$  will be increased by  $B'$ . For aromatic protons, because of their close proximity to the center of the ring, experience an increase of approximately 2 ppm due to ring current effects.

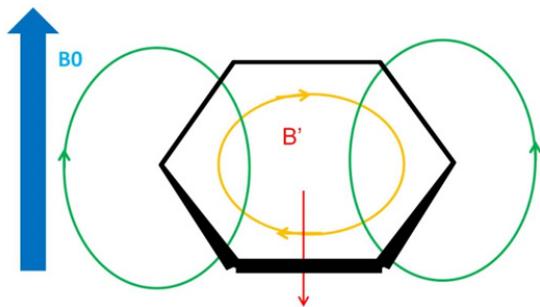


Figure 3.3. Diagram of an aromatic ring current effect: orange and green show the directions of the ring current inside and beside the ring respectively. The blue ( $B_0$ ) and red ( $B'$ ) indicate the directions of the external magnetic field, and the induced local magnetic field respectively.  $B'$  is proportional to  $B_0$  and typically  $10^4$ - $10^5$  times smaller.

The typical chemical shifts for protons of nucleic acids and proteins at room temperature have previously been reported [91]. Tables 3.1 and 3.2 list the chemical shifts of DNA and protein, respectively. Figure 3.4 summarizes the most representative values for DNA and protein.

Table 3.1 Typical chemical shift ranges for proton resonances in NMR spectra of nucleic acids. Proton chemical shifts are for Watson-Crick base pairs in B-DNA. Statistics from BMRB (Biological Magnetic Resonance Data Bank).

Type	Expected chemical shift	Type	Expected chemical shift
T (CH <sub>3</sub> )	1.00-2.00ppm	A,G,C,T (CH)	6.50-8.20ppm
Sugar 2'	2.00-3.00ppm	C (NH <sub>2</sub> )	8.30-8.50ppm
Sugar 5' terminal	3.70ppm	G (NH)	12.50-13.00ppm
Sugar 5';4'	4.00-4.50ppm	T (NH)	13.50-14.00ppm
Sugar 3'	4.50-5.20ppm		
Sugar 1'	5.30-6.20ppm		

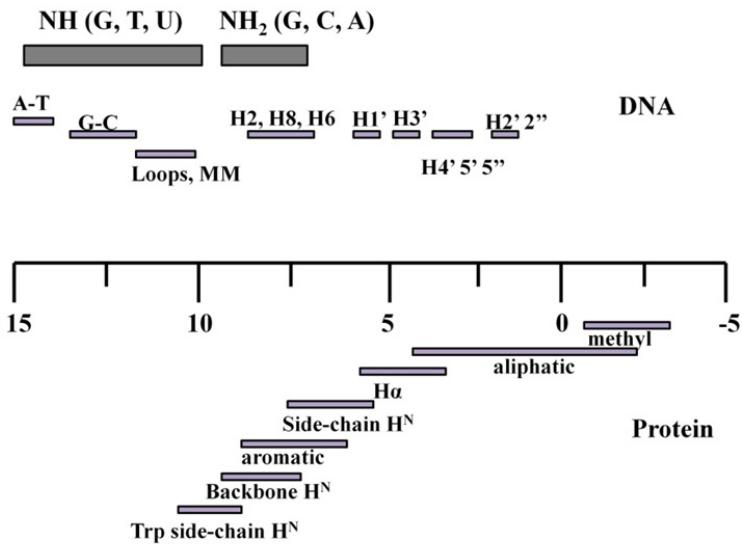


Figure 3.4. DNA and protein proton chemical shift distributions in spectrum region.

The chemical shift of protein is shown in Table 3.2. Note that the nitrogen attached side chain proton in tryptophan appears at larger ppm. And the HN of amino acid all present at the chemical shift of approximately 8ppm (Figure 3.4). Proline is the exception of amino acid which does not possess HN. In a folded protein spectrum, the chemical shifts of all amino acids will be distributed to a larger range (Figure 3.5, right panel). However, in a random coil peptide structure, all the resonances tend to allocate to one region (left panel), which brings difficulties for the assignments.

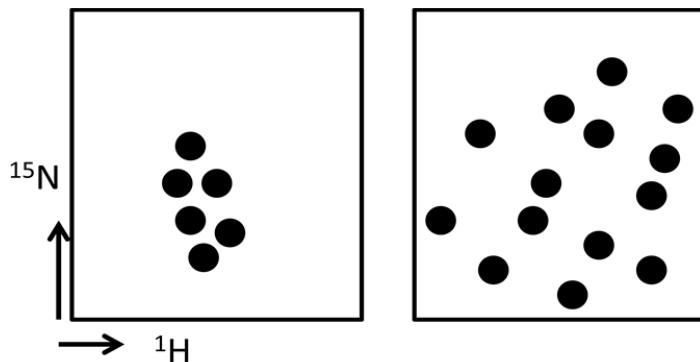


Figure 3.5 Illustration of unfolded protein spectrum (left) and folded protein in HSQC spectrum. The horizontal and vertical dimensions represent  $^1\text{H}$  and  $^{15}\text{N}$ .

Table 3.2 Typical proton chemical shifts (average) of amino acids within a protein, two digital statistics shifts are from BMRB.

Type	HN	HA	HB	other
GLY	8.3	3.9	-	
ALA	8.2	4.3	1.4	
VAL	8.3	4.2	2.0	0.83(methyl)
SER	8.3	4.5	3.7,	3.6b
THR	8.2	4.5	1.1(methyl)	
CYS	8.4	4.7		
ASP	8.3	4.6	2.7	
ASN	8.3	4.7	2.8,	NH <sub>2</sub> :7.3
GLU	8.3	4.3	2.0,	2.3 (H <sub>γ</sub> ),
GLN	8.2	4.3	2.0,	1.3 (H <sub>γ</sub> 1),      0.7 (methyl)      NH <sub>2</sub> :7.2
ILE	8.3	4.2	1.8	1.5 (H <sub>γ</sub> )      0.7 (methyl)
LEU	8.2	4.3	1.6,	1.5 (H <sub>γ</sub> ),      0.7 (methyl)
LYS	8.2	4.3	1.8,	1.4 (H <sub>γ</sub> )      1.6 (H <sub>δ</sub> ),      2.9 (H <sub>ε</sub> ),      NH <sub>3</sub> :7.4
ARG	8.2	4.3	1.8,	1.6 (H <sub>γ</sub> )      3.1(H <sub>δ</sub> ),      NH:7.1      NH <sub>2</sub> :6.8
MET	8.3	4.4	2.0,	2.4 (H <sub>γ</sub> )      2.2 (methyl)
PRO		4.4	2.0,	1.9 (H <sub>γ</sub> )      3.6 (H <sub>δ</sub> )
PHE	8.3	4.6	3.0,	ring:      6.5 to 7.6
TYR	8.3	4.6	2.9,	ring:      6.6 to 7.4
TRP	8.3	4.6	2.9,	ring:      6.5 to 7.9, NH: 10.0
HIS	8.2	4.6	3.0,	ring:      HD2, HE1: 7.0-8.0, HD1: 8.7 HE2: 9.7

Chemical shifts may vary by as much as approximately + 1.5, and sometimes more, depending on local structure. The significant down filed shifted NH of tryptophan at ~10ppm is marked.

### 3.1.2 Chemical exchange

Nuclear magnetic resonance parameters (e.g. chemical shift) are sensitive to environment. In a molecule, this can occur either due to a chemical reaction, conformational alterations as well as ligand interaction. The processes where the nuclei undergo different chemical environments are called ‘chemical exchange’. If the exchange between the environment is sufficiently rapid (fast exchange, ps-ns), the observed NMR parameters will be an average of the parameter values in each of the different environments or sites. And if it is in slow exchange ( $\mu$ s-ms), two distinct signals will be observed; Intermediate exchange gives broadened line width. In the case of the azapropazone derivative [92] shown in Figure 3.6, the two methyl groups that are attached to N undergo chemical exchange while they are rotating about the C-N bond. During the rotation, the methyl nuclei experience different chemical environments. At very low temperature (223K bottom of Figure 3.6) when exchange is absent, the  $^1\text{H}$  spectrum of the N-methyl group gives two sharp peaks which come from each of the methyl groups. With the temperature increased, the rate of exchange becomes faster; these two separate peaks first become broadened, and then merged into a single broad peak at 263K. When the temperature continues to rise, the broad peak becomes narrowed at the average position (of the two peaks) at 273K. From this, we can see that NMR is a very versatile tool and therefore can be used to measure chemical exchanges or molecular dynamics, and relaxations in solution.

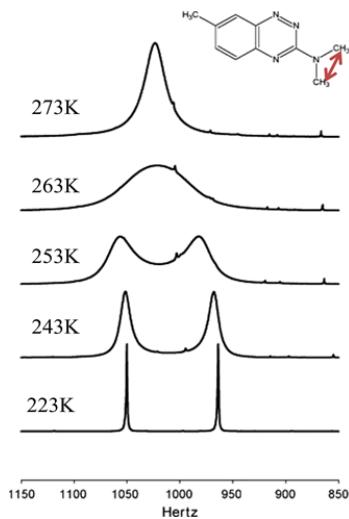


Figure 3.6: The N-methyl region from a 300MHz  $^1\text{H}$  NMR spectrum of an azapropazone derivative recorded at 223 (lowest), 243, 253, 263, and 273K (top). The figure is adapted from [92]

### 3.1.3 Nuclear Overhauser Effects SpectroscopY (NOESY)

The NOE effect normally occurs when a proton is close in space to another (or any other nuclei which have non zero spins. In common NMR, we consider nuclei with spin 1/2), their magnetic dipoles interact. NOE is distinct from scalar coupling (J coupling), which is not a through space

effect, but is mediated by polarization of bonding electrons in the molecule. The definition of NOE can be determined as: when a proton is saturated or inverted, spatially-close protons may experience an intensity enhancement. It depends only on the spatial proximity between protons, in other words, the strength of the NOE gives information on how close two protons are. And NOE may be observed between protons that are up to 5Å apart.

2D-NOESY pulse sequence (Figure 3.7, upper panel) contains three  $90^0$  pulses. The first  $90^0$  pulse creates the transverse spin magnetization which flips the magnetization from z-axis to x,y plane, followed by an evolution time  $t_1$ . This delay is varied systematically as usual to provide chemical shift information in the F1 domain. Then a second  $90^0$  pulse (applied on the y-axis) transmit some of the magnetization back to the z axis and during the following mixing period ( $\tau_m$ ), the non-equilibria z component will exchange magnetization through relaxation (dipole-dipole mechanism). The  $\tau_m$  is kept constant during the 2D experiment. Shorter  $\tau_m$  results in cross peaks directly between two nuclei (A→B) that are in short distance, and longer  $\tau_m$  could give rise to indirect connections (A→B→C). Such an effect is considered as ‘spin diffusion’ and can confuse the interpretation. In practice, several NOESY spectra are collected with different mixing times to be able to collect enough information. The third  $90^0$  pulse flips the magnetization back to the x, y plane, acquisition starts immediately after the third pulse and the transverse magnetization is observed as a function of  $t_2$ . 2D Fourier transform is then performed to generate the NOESY spectrum with respect to  $\omega_1$  and  $\omega_2$ .

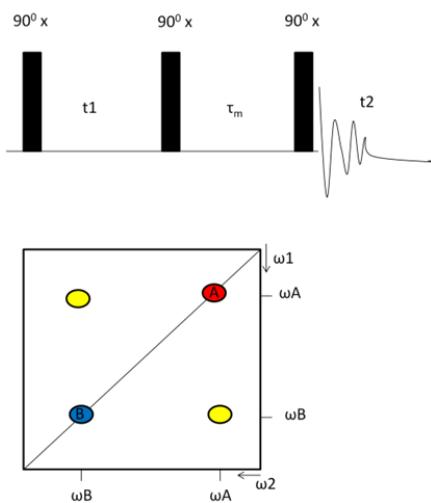


Figure 3.7. NOESY pulse sequence (upper panel) and spectrum (lower panel). A and B are two different nuclei that are close in distance. Yellow peaks are cross peaks between A and B. Vertical and horizontal dimensions are for  $\omega_1$  and  $\omega_2$ .

Back to the spectrum (Figure 3.7, lower panel), suppose that in some unspecified two-dimensional NOESY spectrum, a peak appears at  $\omega_1 = \omega_A$ ,  $\omega_2 = \omega_B$  (yellow peak). The interpretation of this peak is that a magnetization was present during  $t_1$  which evolved with a frequency of  $\omega_A$ . During the mixing time the same magnetization was transferred in some way to another nucleus (B) which evolved at  $\omega_B$  during  $t_2$ . Likewise, if there is a peak on the diagonal, at  $\omega_1=\omega_A$ ,  $\omega_2 = \omega_A$  the interpretation is that there was a signal evolving at  $\omega_A$  during  $t_1$  which was unaffected by the mixing period and continued to evolve at  $\omega_A$  during  $t_2$ . The cross peak intensity is proportional to  $1/r^6$ .  $r$  is the distance between the two nuclei.

The maximum possible NOE depends on the molecular correlation time (the inverse of the rate of molecular tumbling), which is in large part determined by the molecular weight and solvent viscosity. Larger molecular weights and higher viscosities lead to larger correlation times. The NOE is positive for small molecules ( $MW < 600$  Da), goes through zero for medium-sized molecules ( $MW$  range 700-1200), and becomes negative for large molecules ( $MW > 1200$ ). For medium sized molecules, the NOE may be (theoretically) zero (Figure 3.8).

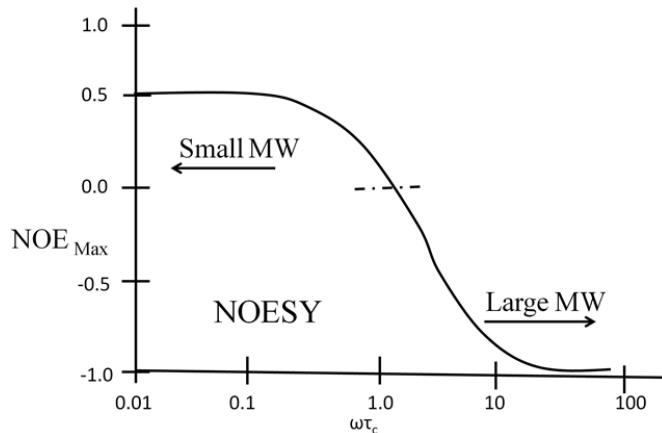


Figure 3.8. NOESY intensity vs. molecular correlation  $\tau_c$ . Small molecules give rise to positive NOEs, while big molecules show negative NOEs. The figure is adapted from Neuhaus and Williamson's text [93].

### 3.1.4 Experiments and assignments strategy

Besides the NOESY experiment, other types of experiments are also utilized to aid assignment before structure analysis. Such as TOCSY (Total Correlated Spectroscopy also known as HOHAHA-Homonuclear Hartmann Hahn) which is useful to divide the proton signals into groups or coupling networks. TOCSY unlike NOESY yields through bond correlations via spin-spin coupling. Correlations are seen throughout the coupling network. The intensity of cross peaks is not related in a simple fashion to the number of bonds connecting the protons; instead it

depends on the  $^3J$  constants and dihedral angles (Karplus equation). The requirement for the appearance of cross peaks in a proton TOCSY spectrum is the presence of bonds between neighboring protons. An asparagine residue with torsion angles of N-C $\alpha$ -C $\beta$ -C $\delta$  approximately  $180^\circ$  (Figure 3.9) gives seven TOCSY and eleven NOESY cross peaks. TOCSY cross peaks belong to HN-H $\alpha$ /H $\alpha'$ -H $\beta$ / $\beta'$ ; H $\alpha$ / $\alpha'$ -H $\beta$ / $\beta'$ ; between two H $\beta$ s and two H $\delta$ s. Note that there is no TOCSY cross peak between H $\beta$ / $\beta'$ -H $\delta$ / $\delta'$ , because there is no proton in between. However, the distance between H $\delta$ s and H $\beta$ s are short enough to give rise NOE cross peak.

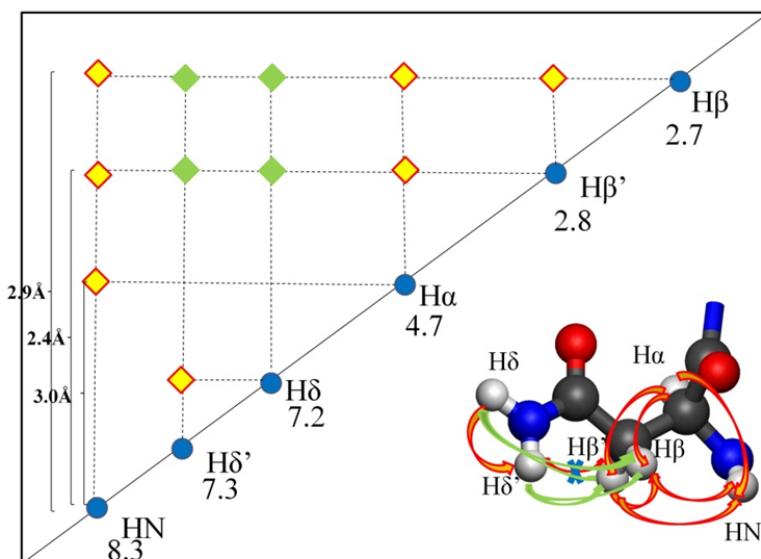


Figure 3.9. NOESY and TOCSY spectra for asparagine (ASN). Cross peaks color coding: red for only TOCSY; red outline and yellow filled peak: TOCSY and NOESY cross peaks; green: only NOESY; blue: diagonal peaks. The structure of ASN is illustrated on the right down corner. Red and green arrows represent TOCSY and NOESY connections respectively. Notice that there is no TOCSY cross peak between H $\delta$  and H $\beta$ . Spectra is only for the structure of ASN shown in this figure. The atom names are indicated on the structure of ASN and for the diagonal peaks. Part of the short distances between two protons in ASN is displayed in the spectrum.

Two dimensional COSY (CORrelation SpectroscopY) allows us to connect the protons that belong to the same systems. Similar to TOCSY, in COSY, cross peaks can be only seen if two atoms are bonded, differ from TOCSY, COSY gives cross peaks for nuclei that are only three bonds apart. COSY is basically a simple version of TOCSY, which gives a simplified connection network than that provided from TOCSY. For the same residue ASN shown, COSY cross peak only arises from HN to H $\alpha$  and no cross peaks to H $\beta$ s, this is because the number of bonds between HN and H $\beta$  exceeds three. In this thesis, unlabeled B-form and mismatched DNA, as well as lysozymes are assigned manually. The assignments for both biological molecule systems were carried out by combinations of NOESY, TOCSY and COSY. Examples for DNAs and lysozyme assignments are shown in Results.

## 3.2 Optical measurements

Apart from NMR spectroscopy, circular dichroism and fluorescence are considered as other important techniques which provide valuable results for this thesis study. Comprehensive reviews of these techniques can be found in references [94-96].

Electromagnetic radiation of wavelength  $\lambda$  can be seen as a wave propagating at the speed of light  $c$  with an electric and a magnetic field component oscillating in phase at the frequency  $v = c/\lambda$ .

When a molecule in an initial state  $S_i$  is exposed to electromagnetic radiation, it may absorb a photon and make a transition to a higher-energy final state  $S_f$  with the given frequency,  $v$ , of the light that satisfies the frequency condition  $\Delta E = hv$ , where  $\Delta E$  represents the energy difference between the two states. These kinds of measurements, although giving less structural details than X-ray and NMR spectroscopy, provide valuable complementary information on drug-DNA interactions and are considered as quick methods which do not require large amounts of sample. Thus sample condition surveys are usually done by these methods, e.g. salt concentration, pH, and temperature *etc.*

The theory of NMR is similar to that of optical spectroscopy. However, some differences should be mentioned. For example in UV-visible absorption spectroscopy, which occurs as a result of electronic transitions, where all of the molecules will be in the ground electronic state at room temperature, the energy difference between the ground and excited states is large. However, in NMR, the difference in energy in the two spin states (spin up and down) is negligible. Therefore the population difference is also small (about 1 in 10,000 for  $^1\text{H}$  in an 11.74 T magnetic field). Note that the population difference determines the intensity of signals; NMR is thus inherently a less sensitive technique than many other spectrometers.

### 3.2.1 Circular Dichroism

For circularly polarized light, the electric and magnetic field vectors have a constant length, but rotate about their axis of propagation. They thus form ‘chiral’ (left and right handed) helices propagating through space and will interact differently, and give rise to a non-vanishing CD signal, only with molecules which are themselves chiral or perturbed by a chiral environment. The structured right-handed DNA helices produce a strong CD signal from the intrinsically achiral nucleobases because of their electronic interaction, and likewise in DNA-drug non-covalent complexes, CD signals can be induced also for achiral molecules.

### 3.2.2 Fluorescence

Fluorescence is a spectroscopic method to analyze the molecules that are excited by irradiation at a certain wavelength and emit radiation at a different wavelength. The Jablonski diagram in

Figure 3.10 illustrates processes that may follow the electronic excitation (absorption). The ground, first and second excited states are represented by  $S_0$ ,  $S_1$  and  $S_2$ , respectively. Each state is divided into several vibrational states. A photon of energy  $h\nu_A$  is absorbed by a molecule, which is excited from  $S_0$  to a higher electronic singlet state. The molecule is relaxed to the lowest vibrational level of  $S_1$  by releasing energy as heat through vibrational relaxation and internal conversion. The molecule can then return to the ground state  $S_0$  via either radiative or non-radiative processes. The latter include either emitting fluorescence ( $h\nu_F$ ), i.e. relaxation directly from the singlet state; or by undergoing intersystem crossing to the first triplet excited state  $T_1$ , which can then be converted to the ground state by emission of phosphorescence ( $h\nu_p$ ) [94].

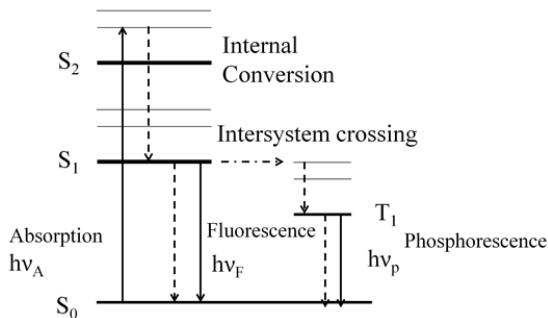


Figure 3.10. the Jablonski diagram showing excitation and relaxation pathways of the electrons.

Emission of light takes place on a time scale much slower than absorption which allows a wider range of interactions and perturbations to influence the spectrum. As mentioned above, there are several different pathways that are involved to depopulate the excited state. The quantum yield or quantum efficiency for fluorescence is the ratio of the number of molecules that fluoresces to the total number of excited molecules. The ratio between radioactive and non-radiative relaxation is defined as the efficiency of fluorescence or quantum yield:

$$\Phi = \frac{\text{number of photons emitted}}{\text{number of photons absorbed}} = \frac{k_F}{k_F + k_{IC} + k_{IS} + k_Q}$$

Where  $k_F$ ,  $k_{IC}$ ,  $k_{IS}$  are rate constants for fluorescence, internal conversion and intersystem crossing, respectively.  $k_Q$  represents the sum of rate constants through quenching, a process where the energy loses through interactions with other molecules.

### 3.3 Computational structure determinations

NMR provides important geometric information about how the 3D structure looks like for macromolecules. How can we use this information to obtain the 3D structure? Given a sufficiently complete list of assigned chemical shifts and one or several lists of cross-peak positions provided by two-, three-, or four-dimensional nuclear overhauser effect spectroscopy (NOESY), the 3D structure of the protein or DNA in solution can be automatically calculated

with CYANA [97] (Combined assignment and dynamics algorithm for NMR applications). The structure calculation procedure is shown below.

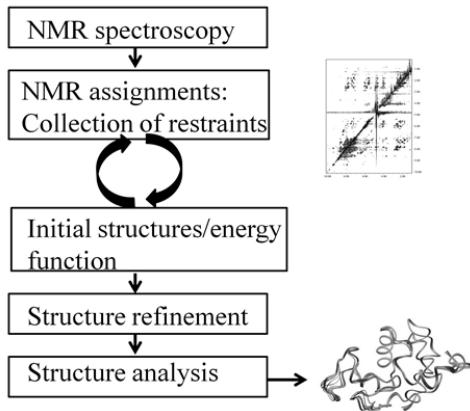


Figure 3.11. General scheme of automated combined NOESY assignments and structure calculation.

Structure determination using NMR data relies on a network of restraints. Distance restraints are derived from nuclear Overhauser effects (NOEs) that can be observed between nearby hydrogen atoms recorded in NOESY spectra. The signal intensity is directly related to the distance ‘r’ between the protons involved ( $I_{NOE}=1/r^6$ ) and each restraint connects two hydrogens that are spatially separated by less than about 5Å. In this way, 3D structures can be calculated (Figure 3.12).

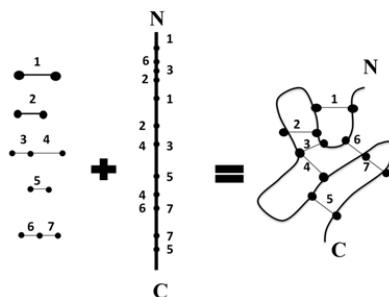


Figure 3.12. Principle of the structure determination by NMR: schematic drawing of the use of distance restraints for the determination of the 3D structure of biological macromolecules. On the left, distances obtained from NOESY spectra for seven pairs of protons are indicated with numbers from 1 to 7. In the middle figure, the straight line represents the polypeptide chain, and the black dots for the sequential positions of the individual protons which define the seven distances on the left figure. N and C mark the N and C terminus. In the right hand figure, adding the information on the distance between pairs of protons and their position in the polypeptide sequence allows construction of possible arrangements where all the distance constraints are fulfilled.

Both energy minimization (see structural refinement in Figure 3.11) and simulated annealing [98] have the same goal to change the atomic coordinates in order to obtain a structure which has the lowest value for either an experimental or non-experimental (energy) function. During energy

minimization, the atomic coordinates are adjusted in multiple steps. However, the energy surface as a function of atomic coordinates is complex; a simple minimization of the energy will inevitably reach a local minimum, rather than the global minimum, as shown in Fig. 3.13. Simulated annealing is used to overcome the problem of the structure being trapped in a local energy minimum. The structure is heated to high temperatures (where atoms have high kinetic energy) to facilitate atomic rearrangements and then cooled or annealed to more stable structures. At high temperatures the atoms will have high kinetic energy, and will be able to overcome the energy barrier between minima. The motion or trajectories of the atoms during simulating annealing are determined by molecular dynamics calculations. A set of initial coordinates and velocities (which depends on temperature) are given,  $x_0$  and  $v_0$ . The coordinates at time  $t+\Delta t$  are calculated from the coordinate at time  $t$ , using Newtonian mechanics where  $-\nabla E$  is equal to the force applied to the atoms,  $E$  represents energy:

$$x_{i+1} = x_i + v_i \Delta t - \nabla E \frac{\Delta t^2}{2m}$$

$$v_{i+1} = v_i - \frac{\nabla E}{m} \Delta t$$

After applying these two equations for thousands of times, each time calculating new coordinates for each atom, low-energy structure can be obtained.

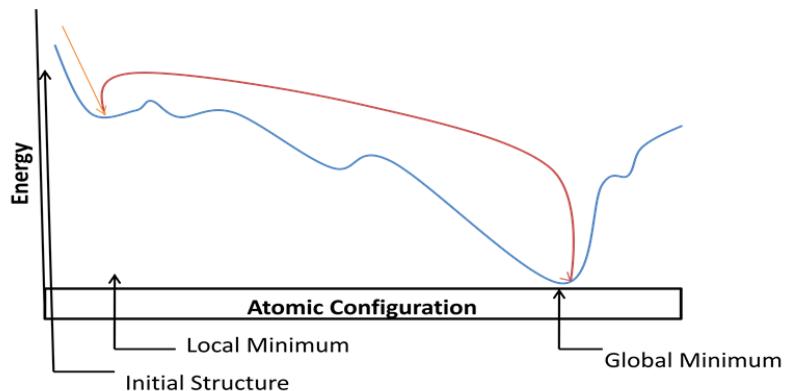


Figure 3.13 Energy changes during simulated annealing. The initial structure is minimized to the local minimum by minimization process (orange arrow), after annealing (red line), the new structure should correspond to energy minimum.

The minimization algorithm of CYANA is based on the idea of simulated annealing by molecular dynamics simulation in torsion angle space independently. The standard simulated annealing protocol in the program CYANA starts from a conformation with all torsion angles treated as random variables. Before CYANA calculations, the experimental input data is read in; this input file contains a library file, which stores the geometry information of nucleotides or

amino acids; a sequence file which lists the residues in the proper order; upper and lower limit files which define NOE restraints and possibly other distances such as the hydrogen bonds of B-DNA or secondary structures. Van der Waals restraints are produced by the programme. After reading these restraints, CYANA starts to find a best structure which has the lowest target function (the sum of violations of the restraints for NOEs, H-bonds and Van der Waals *etc*) from a random structure. By annealing several random starting structures, some of the resulting structures may have a low target functions and can be selected for description of the structure. The efficiency of the torsion angle dynamics algorithm implemented in the program CYANA is high because it requires a computational effort that increases only linearly [97] with the system size, and the NOE part of the target function defines global minimum more uniquely.



## 4 RESULTS

The aim of this thesis is to investigate the interaction of macromolecules, such as DNA and protein, with organic ligand (dimeric ruthenium compounds, Hoechst 33258, disaccharide) at atomic level by NMR spectroscopy, and with the help of some optical spectroscopic methods. This chapter summarizes the main result of *paper I-IV* and also presents some relevant unpublished data.

In *paper I*, we demonstrated a groove binding mode of binuclear intercalator  $\Lambda,\Lambda\text{-B}$  to palindromic DNA d[(CGCGAATTGCG)]<sub>2</sub>: one of the positively charged Ru<sup>2+</sup> moieties locates near the AATT region, leaving the other near the end of the duplex. Energy calculations revealed that electrostatic and van der Waals contacts are the main interactions. Two hydrogen bonds are formed between N atoms of upper and lower dppz moieties to G2 and G4. The intermolecular hydrogen bonds help to define a unique position of  $\Lambda,\Lambda\text{-B}$ .

In *Paper II*, the final intercalation state by X-ray crystallography was resolved by Miquel Coll's group in Barcelona. Results showed that half of the Ru compound is intercalated into a hexamer oligonucleotide duplex from the minor groove, extruding the AT base pair towards the widened minor groove. Meanwhile, the other half is recruiting a second neighboring duplex via end-stacking. Optical studies indicate that such interactions are present also in dilute aqueous solution.

In the context of threading intercalation, solution NMR studies of the same binuclear Ru compound as in *paper I* with a 12-mer mismatched DNA were carried out. Preliminary results indicate that one binuclear Ru compound has intercalated into two DNA duplexes simultaneously. Complete data analyses (NMR assignments) are still ongoing.

*Paper III* studies Hoechst 33258 in the context of DNA-mediated ligand-ligand interactions, where 1D-<sup>1</sup>H-NMR was used in conjunction with optical spectroscopy methods and microcalorimetry. The NMR results provided supports for the consecutive non-overlapping binding of two Hoechst molecules to the oligonucleotide duplex.

In *Paper IV*, we found that lysozyme undergoes domain-wide structural rearrangements in accordance with the deprotonation process of the catalytic residue Glu 35. Studies on binding of a disaccharide molecule, which is a fragment of the cell wall from bacteria *Klebsiella pneumoniae* to both hen egg white lysozyme (HEWL) and human lysozyme indicated that the electrostatic potential and hydrophobicity of individual protein play important roles.

#### 4.1 Groove binding

From earlier extensive studies, it has been shown that  $\Delta,\Delta$ -enantiomer of  $[\mu\text{-} \text{bidppz(phen)}_4\text{Ru}_2]^{4+}$  reorganizes from the initial surface bound state to intercalation in complex with calf thymus DNA (ct-DNA) [20].

The goal of *Paper I* was to capture and characterize the surface bound state before intercalation takes place. In the present NMR study, a 12-mer Dickerson-Drew symmetric DNA  $[\text{d(CGCGAATTCGCG)}_2]$  was selected. According to previous findings, both the length of AT and sequence prevent Ru compound from being intercalated, the initial surface bound state was therefore captured prior to intercalation [99].

Initially, the Ru-compound  $\Delta,\Delta\text{-}[\mu\text{-} \text{bidppz(phen)}_4\text{Ru}_2]^{4+}$  which was shown to thread intercalate into DNA in LD studies was chosen as a candidate. However, with the high concentration (mM) required in NMR, the selected compound appeared to be aggregated. Therefore, a second binuclear Ru-compound  $\Lambda,\Lambda\text{-B}$  ( $[\Lambda,\Lambda\text{-}\mu\text{-} \text{bidppz(bipy)}_4\text{Ru}_2]^{4+}$ ) which can be dissolved at the required concentration, was prepared for NMR studies.

Optical measurements were carried out to obtain preliminary results for binding of the metal compound to DNA. Circular dichroism spectra (Figure 4.1) show that the curve of DNA- $\Lambda,\Lambda\text{-B}$  complex (green) is strikingly similar to that of ct-DNA- $\Lambda,\Lambda\text{-B}$  (red). However, it significantly differs from that of  $[\text{poly(dAdT)}_2]\text{-}\Lambda,\Lambda\text{-B}$  (blue), the latter was determined for an intercalation mode. Results suggested that  $\Lambda,\Lambda\text{-B}$  binds to 12-mer DNA in a similar way as the initial surface binding mode to ct-DNA. This conclusion was also supported by luminescence spectra: six times lower magnitude of luminescence intensity for both ct-DNA and  $[\text{d(CGCGAATTCGCG)}_2]$  compared to  $[\text{poly(dAdT)}_2]$  (Figure 4.1, inset).

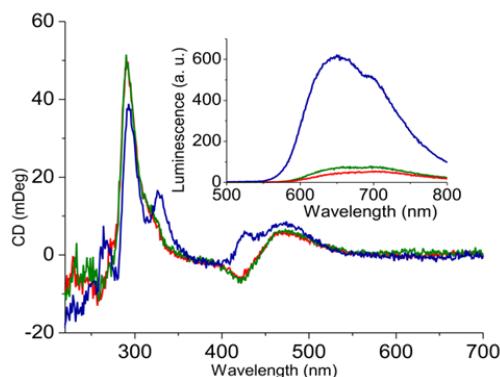


Figure 4.1. Circular dichroism spectra of DNA mixtures with  $50 \mu\text{M}$   $\Lambda,\Lambda\text{-B}$  in  $20\text{mM}$   $\text{pH}=6.5$  phosphate buffer,  $100\text{mM}$   $\text{NaCl}$  water solution:  $[\text{d(CGCGAATTCGCG)}_2]$  (green), ct-DNA (red, mostly covered by green curve), and  $[\text{poly(dAdT)}_2]$  (blue). The  $[\text{nucleotides}]/[\Lambda,\Lambda\text{-B}]$  ratio is 24:1 for all mixtures. Inset: Luminescence spectra of corresponding mixtures with the same colour coding, excitation wavelength:  $410\text{nm}$ . All spectra were recorded about half an hour after mixing at  $25^\circ$ .

With the indication for a surface bound geometry from optical measurements, NMR studies were performed to determine the 3-D structure.  $^1\text{H}$  1D, 2D NOESY and TOCSY NMR spectra were recorded for  $\Lambda,\Lambda$ -B alone, and for a DNA- $\Lambda,\Lambda$ -B mixture. For the compound alone, the molecule itself can be considered as a dimer of two monomers, which are identical to each other. This means that only one setting of signals should be observed in NMR. However, the two axial ( $\text{A}, \text{A}'$ ) and the two equatorial ( $\text{B}, \text{B}'$ ) pyridine rings show slightly different chemical shifts due to the breaking of the monomer 2-fold symmetry by the pivot bond, this results in two sets of peaks for the pyridine rings shown in Figure 4.2. The two dppz motifs appear to be identical. Thus one set of peaks are shown. The assignments started by connecting spin systems in the TOCSY spectrum first, for the A and B rings, four four-spin systems (four protons attached) were connected, and for the C ring, one three-spin system was shown. And then individual assignment was completed by utilizing the NOESY spectrum. Figure 4.2 shows the assignments for individual proton. Note that some of the resonances of free compound possess unusually high chemical shift value ( $\text{H}4\text{C}, \text{H}4\text{C}'$ ).

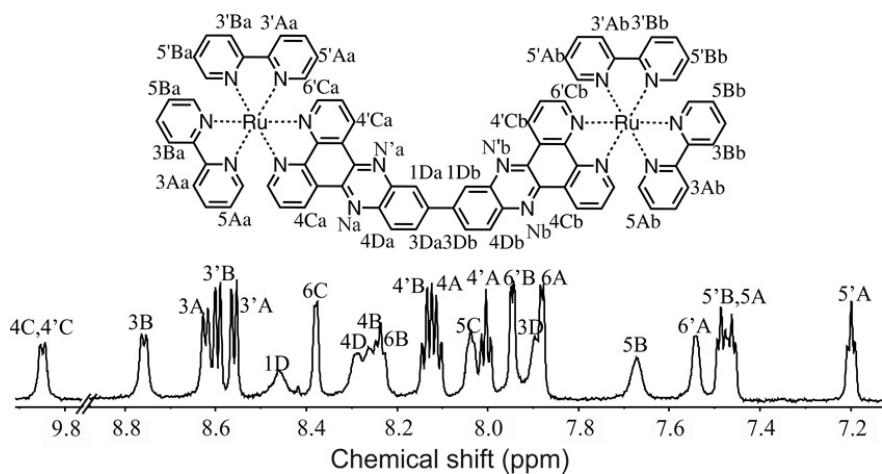


Figure 4.2.  $^1\text{H}$ -NMR assignments of 1 mM  $\Lambda,\Lambda$ -B in 20 mM phosphate pH=6.5, 100 mM NaCl water solution. Resonance assignments refer to positional symbols defined in the chemical structure of  $\Lambda,\Lambda$ -B with positional symbols at carbons at the top. Initial numbers (with or without primes) indicate position within the rings, which are identified by capital letters. A and B are the axial and equatorial (with respect to the dppz plane) pyridine rings, respectively, of the bipyridine ligands; C denotes the pyridine and D denotes the benzene rings of the dppz ligand. Primed and non-primed numbers discriminate the two bipyridines attached to the ruthenium atoms. Lower case letters define the two equivalent monomer units. Spectrum was plotted in Mnova [100].

The 12-mer DNA, which contains two symmetric single strands, was assigned by analyzing NOESY and TOCSY spectra. The full assignments were obtained by following the ‘sequential walk’ [101], an example of which is shown in Figure 4.3C. Different regions of the spectrum are marked with rectangular frames (Figure 4.3A). Strong thymidine methyl-adenine H8 cross peaks

and strong deoxyribose-H2', H2''- base H8 (purine) and H6 (pyrimidine) protons (Figure 4.3B) indicate that the DNA is in B-conformation [102]. This is because of the corresponding pattern of close inter-proton interactions, and the H2'-*endo* conformation of the sugar ring with *anti* base conformation in the B form. The starting point for the assignments is the sequential connection network: H8/6<sub>(i)</sub>-H2'/2''<sub>(i)</sub>-H8/6<sub>(i+1)</sub>-H2'/2''<sub>(i+1)</sub>-H8/6<sub>(i+2)</sub>-H2'/2''<sub>(i+2)</sub> (Figure 4.3C, red lines). Other sequential connection network include: H6/8<sub>(i)</sub>-H1'<sub>(i)</sub>-H6/8<sub>(i+1)</sub>-H1'<sub>(i+1)</sub>-H6/8<sub>(i+2)</sub>-H1'<sub>(i+2)</sub>; the case of consecutive thymine residues, the methyl groups of each thymine residue provide extra sequential assignments pathways: Me<sub>(i)</sub>-H6<sub>(i)</sub>-Me<sub>(i+1)</sub>-H6<sub>(i+1)</sub> (Figure 4.3C, blue lines). For intra-residue connections, TOCSY/COSY spectra are commonly in use, e.g. the sugar ring protons: H1'-H2'/2''-H3'-H4'-H5'/5'' are connected in both types of spectra. However, for the nucleotide base, only signals from cytosine and thymine can be seen: H41-H42, H5-H6 of C, Me-H6 of T. The TOCSY cross peaks of these nuclei can be found in NOESY spectra as well and they are usually located in isolated regions. Due to the close distances, their NOE are strong. Both Cs and Ts can be distinguished from the other nucleotides. Imino protons can also be connected sequentially. The presence of imino cross peaks can provide the evidence that the two single-stranded DNA are base paired (Figure 4.4). This is because these types of protons are attached to N, and only when they form hydrogen bonds (to their counter bases or to ligand), their signals can be observed. When they are base paired, their chemical shifts will appear at larger ppm, (larger than 12ppm). The sequential connections from nucleotide (i) to (i+1) as well as the integrity of imino cross peaks for each base pair prove that the duplex conformations are formed under NMR experimental conditions. Following completion of the assignments for free compound and free DNA, next step is to mix these two.

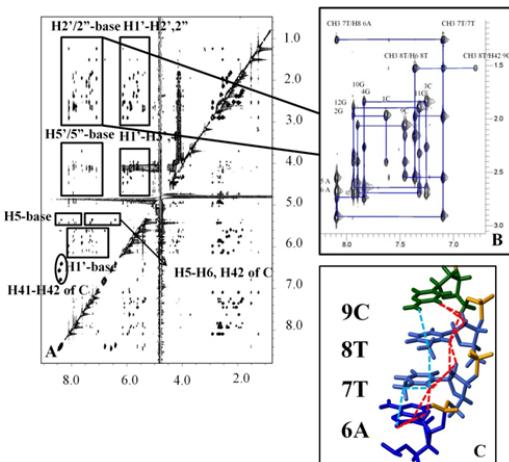


Figure 4.3. NOESY spectrum of 12-mer DNA alone (left), different regions are indicated in panel A, the base-sugar proton region assignment is displayed in panel B. The assignment connecting H2'/2'' to base protons as well as connections from methyl to H2'/H2'' are shown in C (red dashed lines for methyl-H2'/H2'' network; cyan ones are for base-H2'/H2''). Phosphorus bonds are coloured orange.

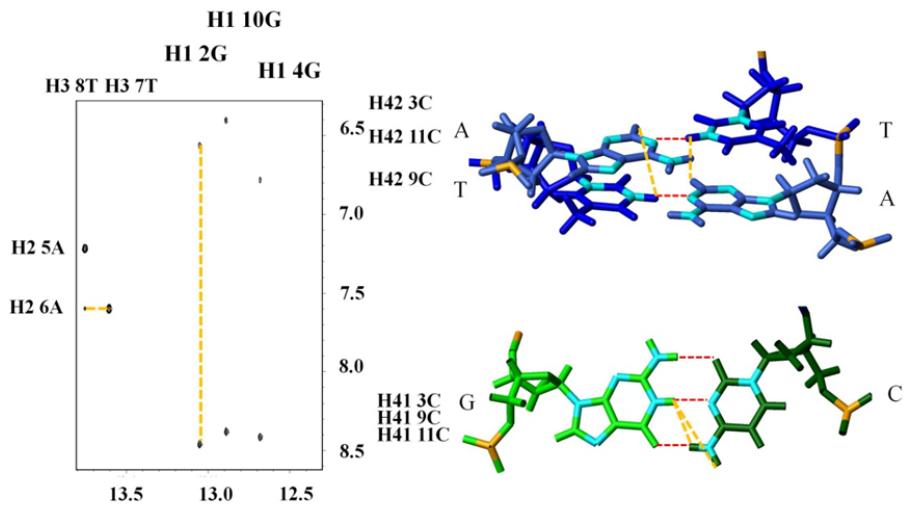


Figure 4.4. Imino proton region in NOESY spectrum (left) of  $[(\text{CGCGAATTGCG})_2]$ , assignments for individual peak are displayed on the border of the spectrum: assignments for vertical dimension at the top and for horizontal dimension at the left and right border. The short distances ( $\leq 4\text{\AA}$ ) for A-T and C-G base pairs are shown with orange dashed lines in both left and right panels. H-bonds are indicated with red lines on the DNA structure. N and P atoms are coloured with cyan and orange, respectively.

Upon titration of  $\Lambda,\Lambda\text{-B}$  to free DNA, two sets of signals become visible: resonances from free DNA, as well as from the complex. The intensity of the complex peaks appear to be increasing with increasing amount of  $\Lambda,\Lambda\text{-B}$ , while peaks from free DNA have decreased but are still present (Figure 4.5). Examples for these two species are marked with black and red rectangular frames. The presence of new peaks at 10-9ppm (signals from  $\Lambda,\Lambda\text{-B}$ ) and at imino region indicate the binding is in slow exchange, which will be discussed later. To characterize the interaction mode for the free DNA, series of 2D spectra were recorded at different ligand to DNA ratios: 0:1, 0.5:1, and 1:1 (Figure 4.6). Resonances from the free DNA appear to shift from low ligand-DNA ratio (magenta spectrum) to high ratio (blue). Thus, the interaction mode for the symmetric DNA with  $\Lambda,\Lambda\text{-B}$  is in fast-exchange. Same as shown in 1D, in 2D NOESY spectrum, two species of DNA duplexes are visible: the symmetric duplex (assignments are shown in Figure 4.3B), which undergoes fast exchange interaction, and an asymmetric DNA which binds to double strands (strands  $\alpha$  and  $\beta$ ). Assignments for the latter are shown in Figure 4.7. Orange and green lines connect each of the single strands.

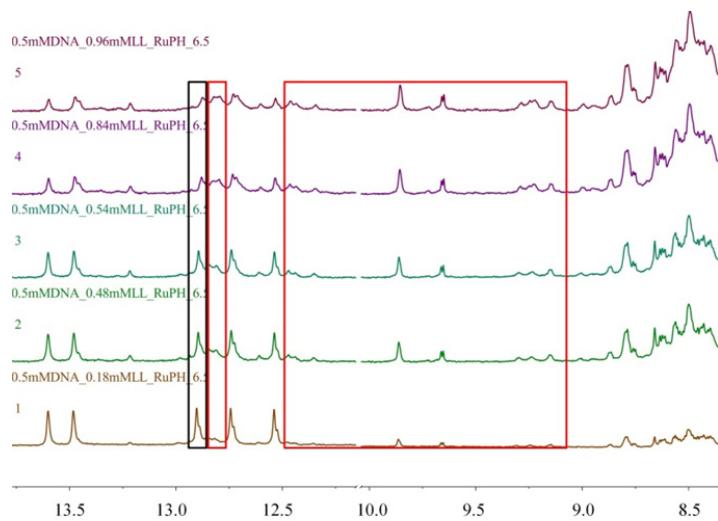


Figure 4.5. 1D-NMR titration of  $\Lambda,\Lambda$ -B into 0.5 mM DNA at  $25^{\circ}$ , in 20mM pH 6.5 phosphate buffer solution, NaCl concentration: 100mM. The black and red rectangular frames mark the free and bound complex respectively. When the ratio of [DNA]/[ $\Lambda,\Lambda$ -B] exceeds 1:1, the line width of DNA imino protons become broadened. The  $\Lambda,\Lambda$ -B signals between 10-8 ppm remain unchanged by adding more  $\Lambda,\Lambda$ -B.

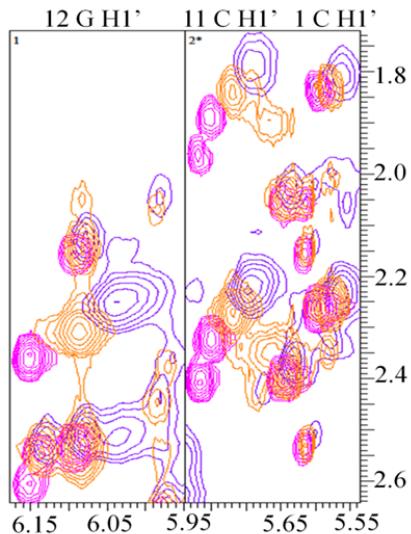


Figure 4.6. Fast exchange evidence for  $\Lambda,\Lambda$ -B and DNA: Overlay of three  ${}^1\text{H}$  NOESY spectra ( $\tau_m = 100$  ms): 1mM DNA-only (magenta), 0.5mM  $\Lambda,\Lambda$ -B mixed with 1mM DNA (orange); mixture of 1mM  $\Lambda,\Lambda$ -B and 1mM DNA (blue). Spectra are overlaid in CCPNmr [103].

Due to the overlapping signals of bound DNA with free DNA, the terminal bases were not assigned. Complete or near complete connectivity of the asymmetric duplex indicate that the DNA remains in B conformation (Figure 4.8, left panel). This contradicts intercalation mode: at least one of the base pair stack has to open up to allow the bulky Ru center to thread in. Plot of

chemical shifts versus residue numbers of these two asymmetric strands show larger perturbations for one of the single strands (Figure 4.8, right panel) than the other, which indicates that the location of  $\Lambda,\Lambda$ -B is in closer contact to one strand than to the other.

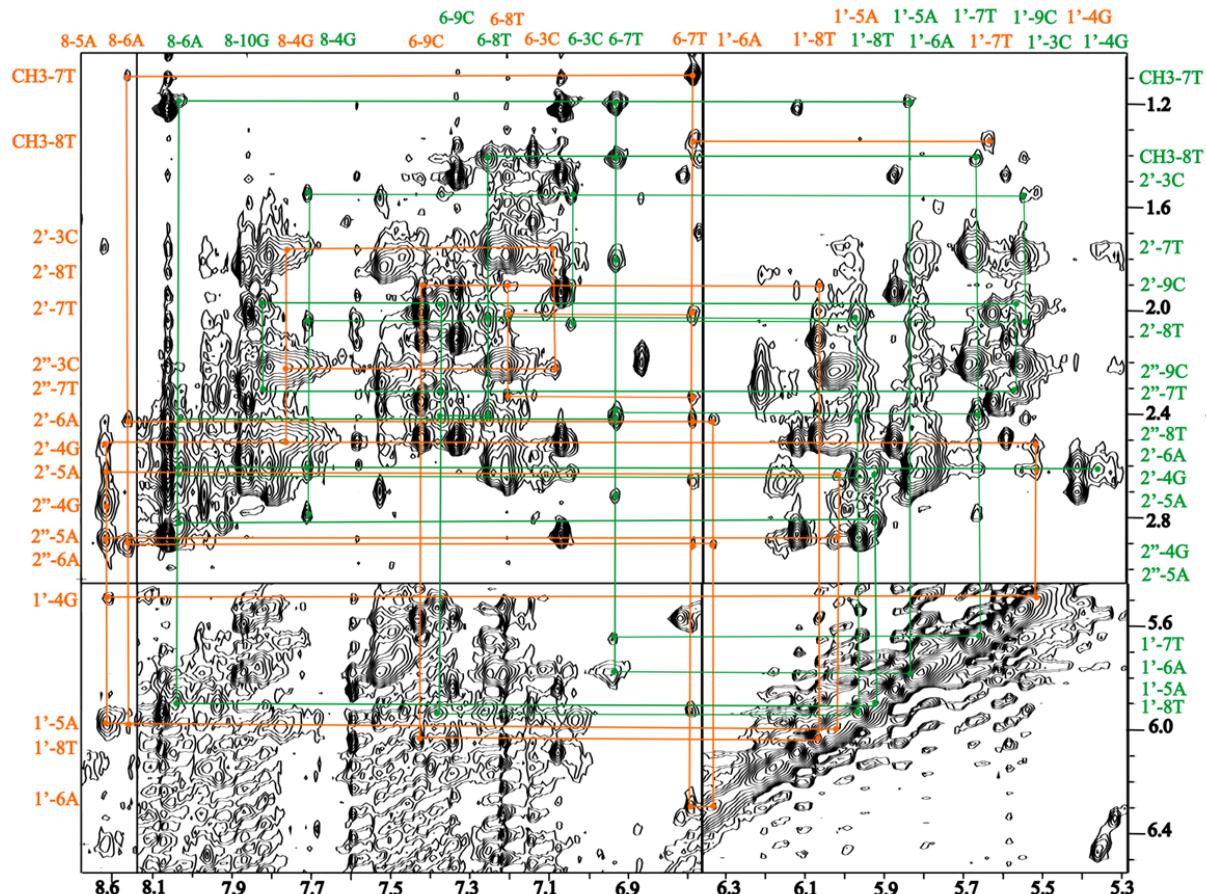


Figure 4.7. Selected regions of the NOESY spectrum of a 1:1 mixture of 1 mM  $\Lambda,\Lambda$ -B with  $[d(CGCGAATTCGCG)]_2$ , showing connections between base and sugar protons. Sequential walks are presented for the  $\alpha$ -strand (green) and the  $\beta$ -strand (orange). Individual cross peaks are identified on the left and top borders with numbers indicating the positions in the sugar or base rings, followed by the nucleotide number and type

According to TOCSY, resonances on the left border in Figure 4.9 (at 9.14 and 9.35 ppm) are part of three-spin systems thus they cannot come from DNA. For normal B-DNA, the chemical shifts of base protons rarely exceed 9ppm. For the assignment of peaks at 9.14 and 9.35ppm, several possible candidates were considered. The most probable ones were designated to H4C and H4'C, this is because their chemical shift is close to 9.14 and 9.35ppm when free in buffer solution. Larger shift changes ( $>1$ ppm) are unlikely in surface bound system. Upon binding to DNA, the symmetry of dppz moiety is broken. Four probable proton candidates for each resonance yield sixteen possibilities. The four with the same protons assigned can immediately be excluded. The

remaining twelve combinations of six consist of identical pairs due to the symmetry of  $\Lambda,\Lambda$ -B. The assigned DNA resonances ( $H1'$  of 2G and  $H1'$  of 5A) were connected to the rest of DNA to verify the validity of their assignments. CYANA was used to find the best complex structure with the lowest target function from all combinations. MD (molecular dynamic) simulation was performed to refine the obtained CYANA structure with the present of water molecule in the system. And finally, a minor groove binding structure was determined:  $\Lambda,\Lambda$ -B binds in the minor groove of the DNA, near to one end of the duplex, leaving one of the Ru centers near the electronegative AATT region in the minor groove (Figure 4.10). Energy calculations showed that besides the electrostatic attraction between  $\Lambda,\Lambda$ -B which carries four positive charges and the strongly negatively charged DNA, the observed minor-groove structure is also stabilized by burying the highly hydrophobic surface of  $\Lambda,\Lambda$ -B to DNA. Moreover, two hydrogen bonds are formed between  $\Lambda,\Lambda$ -B and the phenazine nitrogen atom of G. Interestingly, the binding stoichiometry was determined to be 1:1 instead of 2:1 ( $\Lambda,\Lambda$ -B: DNA), the latter was proposed because of the existence of DNA two ends. A symmetric second binding of  $\Lambda,\Lambda$ -B to the other end of DNA was excluded, although no steric clash between two  $\Lambda,\Lambda$ -B molecules would hinder such a 2:1 complex. The interpretation for this could be because binding  $\Lambda,\Lambda$ -B to one end in the minor groove results in a changed groove at the other end. Lacking of full assignments of the asymmetric DNA for the terminal base pairs makes the actual groove width measurements impossible.

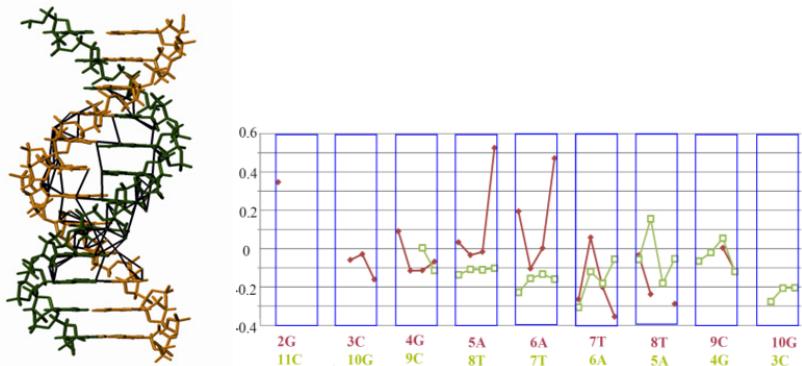


Figure 4.8. Chemical shift differences between free  $[d(CGCGAATTCTCGCG)]_2$  and strand  $\alpha$  (green), strand  $\beta$  (red) as a function of the DNA sequence. Peaks inside each blue box are (from left to right):  $H1'$ ,  $H2'$ ,  $H2''$  and  $H6/H8$ .

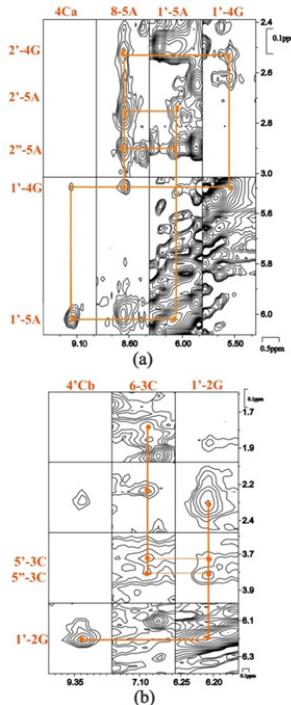


Figure 4.9. Spectral regions from the same NOESY as in Fig. 4.7. Panel (a) shows the NOEs between the  $\Lambda,\Lambda$ -B proton 4Ca (see Fig. 4.7) at 9.14 ppm to H1' of 4G and H1' of 5A; panel (b) shows the NOE between the  $\Lambda,\Lambda$ -B proton 4'Cb at 9.35 ppm to H1' of 2G. All NOEs are to the  $\beta$ -strand. To confirm the DNA assignments, these intermolecular NOEs are connected to the intramolecular network of DNA NOEs. The cross peaks are identified on the left and top borders.

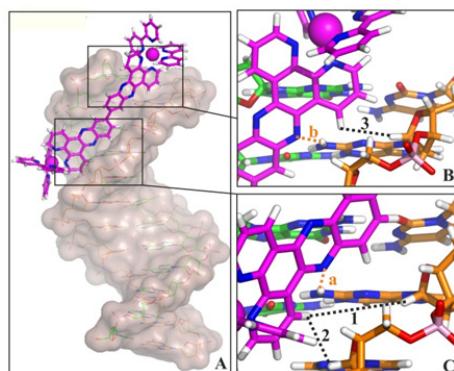


Figure 4.10. Average DNA- $\Lambda,\Lambda$ -B structure of the 5 ns productive, unrestrained trajectory from MD simulation. Panel A: Overall side view. Panels B and C: details of the short distances of three intermolecular NOEs shown in Table 1. Colour coding: green for strand  $\alpha$ , orange for strand  $\beta$ . Dotted lines indicate NOEs (black, numbered 1-3) and potential hydrogen bonds (orange, labelled a and b: Na and N'b of  $\Lambda,\Lambda$ -B with the amino groups of the bases of guanines 4 and 2, respectively).

## 4.2 Binding to mismatch DNA (unpublished data)

In the previous study, we successfully characterized the surface binding properties of Λ,Λ-B with a 12-mer B-DNA duplex. In order to fully understand the mechanism, the intercalation state needs to be studied in detail. Based on the knowledge that was gained from the groove binding study: Ru center is near AATT region, and making hydrogen bonds with the second guanine from the terminus, we thus designed a 12-mer mismatched DNA: 5'-GGCCTTTACGCG-3', 3'-CCGGAATGCGC-5', where the AATT still remains in the middle, and four flanking base pairs in the ends. Mismatched DNA is less stable than normal B-DNA. We have chosen an asymmetric duplex to avoid self-complementing base pairing, thus avoiding hairpin formation.

In order to obtain the intercalated state of Λ,Λ-B-mismatched DNA complex, it is necessary to make sure that the mismatched DNA contains destabilized base pairs to allow a Ru center to thread in, and most importantly the overall duplex is stable at room temperature. For this sample, a series of 1D spectra were recorded upon temperature variations. At the lowest temperature (10°C, bottom spectrum in Figure 4.11) nine imino protons are expected out of eleven (except for the T-T mismatch pair), the terminal two could be missing due to base pair opening. In Figure 4.11, the GC imino protons (13.3-12.4ppm) except of the terminal ones are present in a temperature range of 10-30°C. However, the stability of the AATT region (14.4-13.7 ppm) depends strongly on temperature: at 10°C, the T imino protons become visible. When the temperature was raised above 20°C, all the thymine imino peaks disappear.

At 10°C, DNA is kept in duplex conformation at both GC and AT regions. Between 10-30°C, GC regions are still base paired, but no evidence for the AT base pairing above 10°C. Due to slow motions of molecule tumbling at low temperature, the line width of the spectrum becomes broadened. This could bring difficulties for full 2D spectra assignments. Thus 2D NOESY spectra (Figure 4.12) were recorded at both 25°C and at 30°C to have better line separations. From 1D and 2D spectra, the missing thymine imino protons indicate the absent of hydrogen bonds between adenines and thymines. Moreover, the sequential walk of the DNA base protons-H2/2" stops at the AATT region which means that the consecutive connections between adjacent bases are lost, on the contrast, the GC region was assigned sequentially.

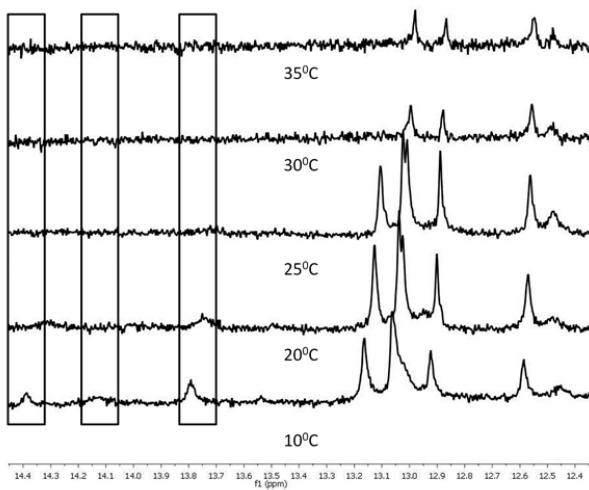


Figure 4.11. 1D temperature studies of DNA alone with 20mM tris-HCl buffer at pH 6.5 with the present of 100mM NaCl on 800MHz Varian instrument. Rectangular frames mark the T imino peaks ranging 14.4-13.7ppm, the G imino protons appear at 13.2-12.5ppm. Line width is broadened when the temperature is below 25°C.

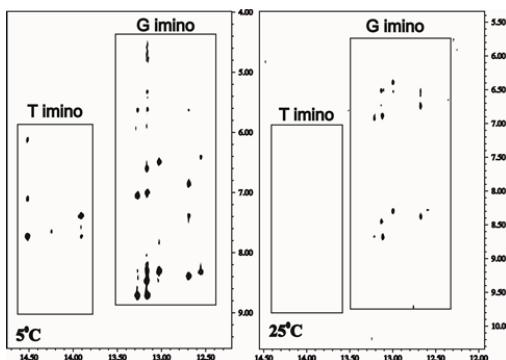


Figure 4.12. 2D NOESY spectra of DNA alone at 5°C (left) and 25°C (right) at imino proton region. The thymine and guanine imino protons are marked with rectangular frames in both spectra.

After characterizing the thermodynamics for DNA alone (0.5mM), next step is to mix the two molecules at room temperature. With the increased amount of  $\Lambda,\Lambda$ -B, both new settings of DNA and  $\Lambda,\Lambda$ -B are shown (Figure 4.13), while the free DNA signals disappeared at the molar ratio of labeled with 0.2:0.12. This indicates that the binding of  $\Lambda,\Lambda$ -B to DNA is in slow exchange. The ratio suggests that the binding stoichiometry is 2:1 (DNA duplex to  $\Lambda,\Lambda$ -B). Recent 2D gCOSY (gradient COSY spectra) analysis shows that in the 1:1 complex (corresponding to 0.2: 0.2 in the last spectrum of Figure 4.13), the free  $\Lambda,\Lambda$ -B is still present. This confirms that one  $\Lambda,\Lambda$ -B binds to multiple duplexes. New samples containing the 0.2:0.12 complexes will be prepared and analyzed in the near future. For the 1D titration, interestingly, one of the missing T imino peaks at 13.7ppm has re-appeared upon binding to  $\Lambda,\Lambda$ -B. More discussions on this can be found in assignment of the mixture (see below).

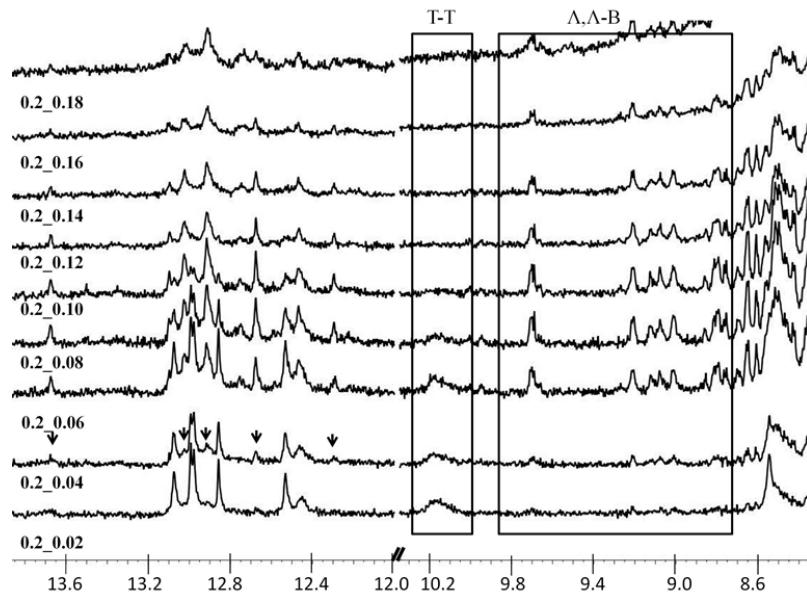


Figure 4.13. 1D titration spectra of DNA-  $\Lambda,\Lambda$ -B at room temperature with 20mM Tris-HCl, pH 6.5. The concentrations for both DNA and  $\Lambda,\Lambda$ -B are shown on the left. Rectangular frames mark the T-T mismatch peaks at 10.18ppm, and also the signals from  $\Lambda,\Lambda$ -B. Arrows indicate new DNA peaks that are growing in the imino region.  $\Lambda,\Lambda$ -B concentration is increased from the bottom to the top spectrum, final ratio is 1:1 (0.5mM).

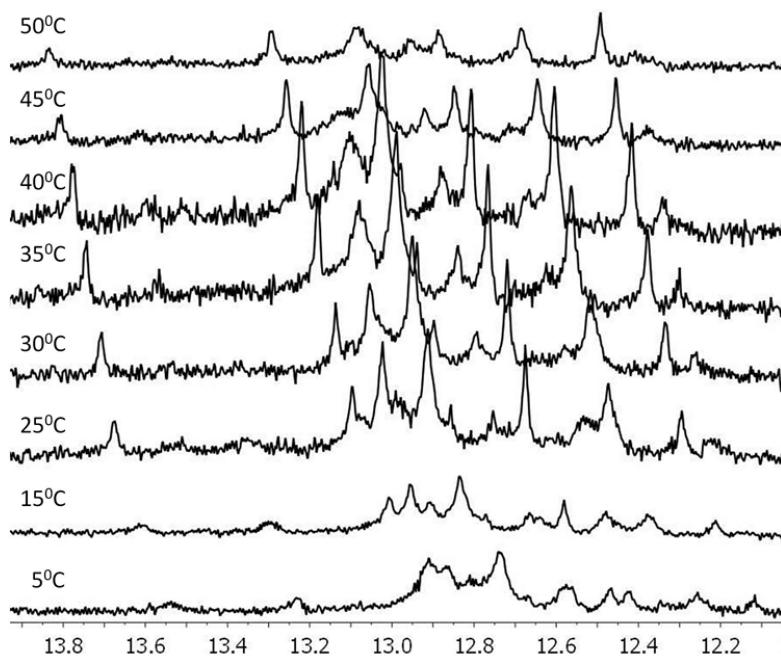


Figure 4.14. 1D temperature titration spectra of 1:1 DNA-  $\Lambda,\Lambda$ -B.

To be able to study the thermodynamic stability of the complex, 1D temperature titration spectra were collected for the 1:1 complexes (Figure 4.14). In the near future temperature studies would be carried out for the 2:1 mixture (mentioned above). Results show that all the present imino peaks, especially the newly appeared one at 13.7ppm are temperature independent, which indicates that the complex is thermodynamically stable, differing from the groove bound complex, where the complex signals intensities decreased with the increased temperature (data now shown). Detailed structure information about the complex has to be obtained from NOESY spectra analysis. In order to accomplish this, DNA and  $\Lambda,\Lambda$ -B has to be assigned individually. Analyzing the TOCSY and gCOSY spectra of the complex, twelve four-spin and five three-spin systems were obtained. Among these, four four-spin and one three-spin systems were designated to the unbound  $\Lambda,\Lambda$ -B, which remains in symmetric. And the rest of the eight four-spin and four three-spin systems were assigned to the bound  $\Lambda,\Lambda$ -B, the symmetry of the bound molecule is broken. Assignments of the bound duplex were completed for the GC region (assignment pathway is shown in the left panel in Figure 4.15), leaving the AT region unassigned. Intermolecular NOEs between  $\Lambda,\Lambda$ -B and DNA are marked with rectangular frames. The horizontal and vertical axes represent  $\Lambda,\Lambda$ -B and DNA dimension respectively. Cross peaks at 2-4ppm (vertical dimension) and at 8-9ppm (horizontal dimension) are derived from the intermolecular connections between sugar protons of the duplex and  $\Lambda,\Lambda$ -B. The ones that are located at 6-7ppm (vertical dimension) and 9-9.5ppm (horizontal dimension) can be ascribed to connections between base protons and  $\Lambda,\Lambda$ -B. No intermolecular NOEs are obtained for the free  $\Lambda,\Lambda$ -B to the duplex.

The increased stabilities in the mismatched- $\Lambda,\Lambda$ -B complex, compared to the minor groove binding system and the 2:1 binding stoichiometry, make the former surface binding mode unlikely. On the other hand, the 1:2 stoichiometry could be a result of intercalation:  $\Lambda,\Lambda$ -B inserts one of the dppz plane into one of the DNA duplex, the other dppz motif associates with a second neighboring one. Interestingly, only one setting of GC imino protons is observed. The reason for this can be speculated:  $\Lambda,\Lambda$ -B binds to the second duplex in a similar manner as to the other one, leaving the GC region symmetric in both duplexes. The final structural characterizations of course will be gained a lot from access to complete assignments, which will be worked on in the future.

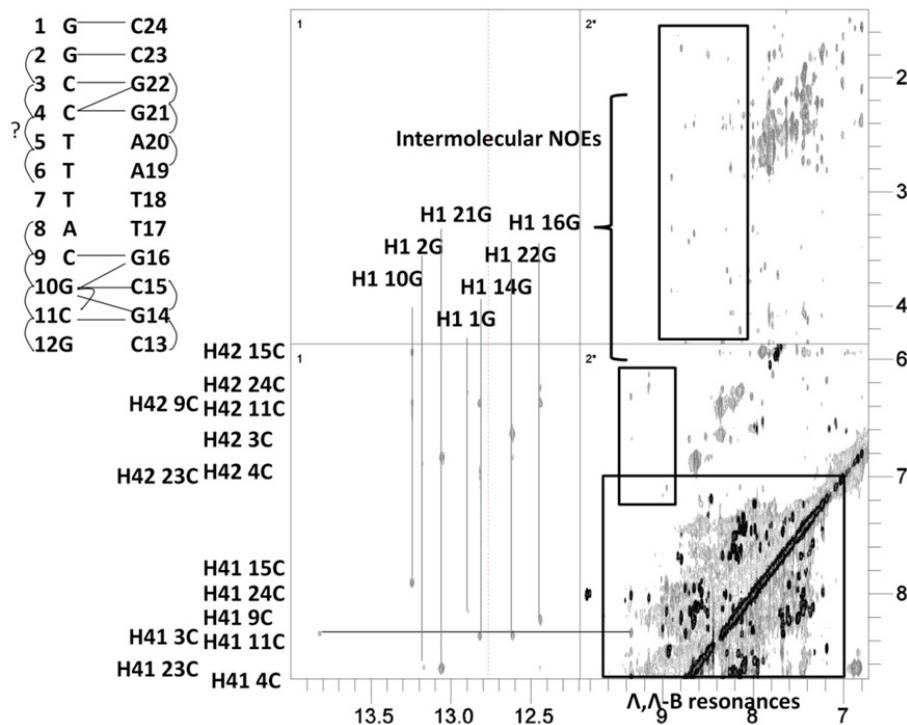


Figure 4.15. NMR spectra of the imino (14-12ppm) and  $\Lambda,\Lambda$ -B region (9.5-6.8ppm). The dark black and grey spectra are NOESY and gCOSY of the 1:1 mixture at 25°C with the present of 20mM Tris-HCl buffer, pH 6.5, respectively. The intermolecular NOEs and  $\Lambda,\Lambda$ -B resonances are indicated with rectangular frames (right panel). Short distances that give rise to NOEs are marked with straight lines for imino region on the duplex sequence and with curved lines for H<sub>2'/</sub>H<sub>2''</sub>-Base or H<sub>5'/</sub>H<sub>5''</sub>-Base sequential connections (in the left upper panel). Assignments of cross peaks of imino resonance are displayed with texts: horizontal dimension on the left of the spectrum and vertical dimension on the top.

#### 4.3 Intercalation into hexamer DNA

Earlier studies suggested that binuclear Ru-compounds thread intercalate into alternating AT DNA, the minimum length of the central AT sequence was determined to be six base pairs [55]. The X-ray study performed in the laboratory of Miquel Coll in Barcelona shows that  $[\mu\text{-}(bidoz)(phen)<sub>4</sub>Ru<sub>2</sub>]<sup>4+</sup>$  ( $\Delta,\Delta$ -P) inserts into a hexamer DNA [(CGTACG)]<sub>2</sub>, which only contains one AT base pair in the middle. The presence of AT is believed to be essential for intercalation, although mismatch is not required. Different from the mismatched DNA- $\Lambda,\Lambda$ -B complex, the stoichiometry between  $\Delta,\Delta$ -P and duplex was determined to be 1:1. Solubility of the 1:1 complex differs from the system described in *paper I*, the  $\Delta,\Delta$ -P-DNA complex has the tendency to form aggregates, because of this, solution NMR studies could not be carried out. Fortunately, fluorescence and absorption experiments provided valuable results which were in great accordance to the structure determined from X-ray analysis.

To understand the aggregation processes, light scattering spectra were collected at 680nm (light scattering region, Figure 4.16). Results show the formation of aggregation depends on the ratio between DNA:  $\Delta,\Delta$ -P. At higher ratio (4:1, blue curve), no precipitates were found; smaller ratio results in larger amount of precipitates. Formation of precipitates are consistent with a step-wise build up of the crystal lattice, proceeding by formation of complexes consisting of two neighboring duplexes held together by two binuclear ruthenium compounds.

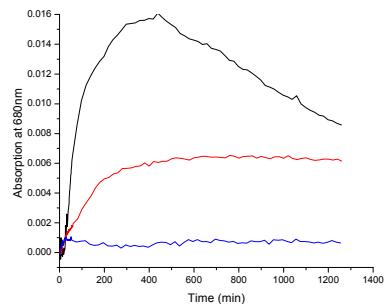


Figure 4.16. Light scattering absorption changes by subtracting the first spectrum from each one recorded at different time (X-axis) for three [DNA base]: $\Delta,\Delta$ -P molar ratios: black, 1:1; red, 2:1; blue, 4:1 at 680nm (light scattering region) 25°C. Sodium chloride concentration was kept in 150mM.

Emission spectra of equivalent amount of  $\Delta,\Delta$ -P in complex with  $[\text{poly(dAdT)}]_2$  (molar ratio per nucleotide) and with the hexamer were recorded and compared (Figure 4.17). Emission intensity with the present of hexamer (red, blue, green curves for different ratio) are 100 times lower than that with  $[\text{poly(dAdT)}]_2$ . The low emission quantum yield can be ascribed to the exposure of dppz moiety to water. The former is expected to be facilitated by the base pair opening. Interestingly, in the mixture of  $\Delta,\Delta$ -B and 12-mer mismatch oligonucleotide (data not shown), low emission intensity was also observed, which indicates the present of exposed dppz moiety.

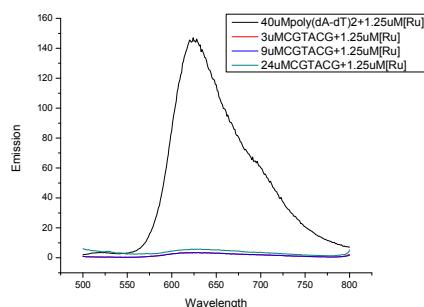


Figure 4.17. Emission spectra of the various mixtures in the present of the same amount of  $\Delta,\Delta$ -P (1.25 $\mu\text{M}$ ) in 20mM phosphate buffer, at room temperature. Black spectrum was recorded in the present of  $[\text{poly(dAdT)}]_2$ , molar ratio of  $[\text{poly(dAdT)}]_2 : \Delta,\Delta$ -P was 2:1. The red, blue and green spectra represent the mixture of [duplex:  $\Delta,\Delta$ -P] in the ratio of 1:5, 1:1.6, and 2:1 respectively.

#### 4.4 Binding studies of Hoechst 33258 to AT DNA

Besides the binuclear ruthenium compounds which are shown to bind in the minor groove of DNA, another small molecule, Hoechst 33258, a well known minor groove binder with a binding preference to AT/rich sequence, is also studied. Despite numerous studies that have already been carried out, the binding property for higher ligand concentration remains unclear. Optical measurements and NMR are combined to understand the binding geometry of a second Hoechst 33258 into three different 12-mer oligonucleotides with different length of ATs in the middle of the sequences [104]:  $[d(CGCGAATTCGCG)]_2$  ( $A_2T_2$ ),  $[d(CGCAAATTTCG)]_2$  ( $A_3T_3$ ) and  $[d(CGAAAATTTCG)]_2$  ( $A_4T_4$ ). Similar binding geometries were found in all oligonucleotides when ligand-to-duplex ratio is equal to 1:1.

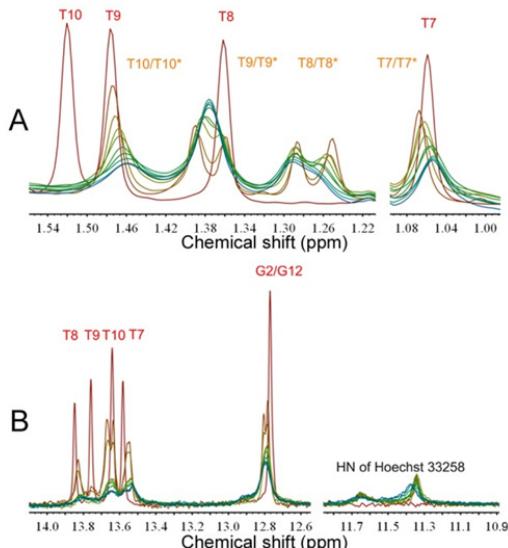


Figure 4.18.  $^1\text{H}$ -NMR spectra and assignments (800 MHz). The thymine methyl region (1.02–1.55 ppm) is shown in (A), while the DNA imino region (12.6–14.1 ppm) and the region with HN of the Hoechst molecule (11.1–11.8 ppm) are shown in (B). The increasing Hoechst 33258 to DNA ratio is illustrated by color: 0:1 (dark red), 1:1 (dark orange), 1.1:1, 1.2:1, 1.3:1, 1.5:1 and 2:1 (blue).

Further structural studies at higher ligand-DNA ratios (2:1) were also carried out by NMR, ITC, and optical spectroscopic measurements. In this thesis, only NMR results are presented. Due to the binding dynamics at ratios beyond 1:1, the broadened spectra lines make the assignments for mixtures that contain ligand-DNA ratio above 1:1 difficult to obtain. Therefore, the binding studies were monitored via the chemical shift changes of certain signals, preferably the methyl groups, since they appear in an isolated spectrum region. Although these methyl groups are located in the major groove, many NMR studies have shown their chemical shift changes upon minor groove binding [105–107]. Addition of equivalent amounts of Hoechst 33258 to  $A_4T_4$  results in strong shifts and splitting of T8 and T9 (dark orange in Figure 4.18), while the more

peripheral ones T7 and T10 are less perturbed. This suggests that the asymmetric ligand binds in the middle of the AT region, which breaks the symmetry of the DNA. Thus eight methyl peaks are shown instead of four. Continuing titration of Hoechst 33258 to DNA leads to the DNA symmetry resumed: the split T8 and T9 peaks in the dark orange spectrum become one peak in the blue spectrum. The regained symmetry is supported by the two HN resonances of Hoechst 33258 (Figure 4.18B) which do not show splitting and only moved marginally upon titration. The small shifts of HN peak also exclude the assumption that was gained from previous study: two Hoechst 33258 bind into the minor groove in a sandwich manner (model A in Figure 4.19). In such a model, significant chemical shifts are expected because of the ring-ring stacking effect. Model C shown below in Figure 4.19 is considered to be the correct one.

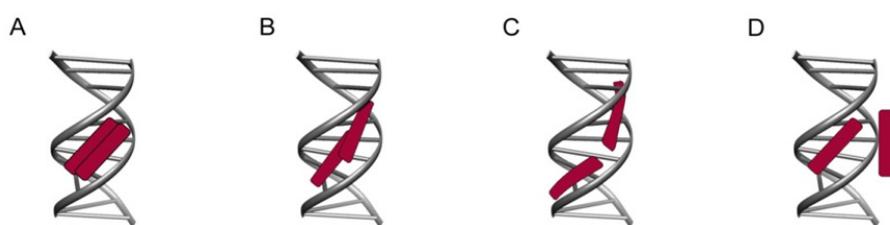


Figure 4.19. Schematic illustrations of possible binding sites for two Hoechst 33258 molecules in the minor groove for the A<sub>4</sub>T<sub>4</sub> dodecamer. The second Hoechst 33258 could either bind as a sandwich, on top of the first molecule (Model A), or parallel in the minor groove and slightly overlapping the first molecule (Model B), or in the minor groove but with a slight distance from the first Hoechst 33258 (Model C). Another putative mode of binding site is that it binds externally, to the negatively charged backbone (Model D).

#### 4.5 Correlation of human lysozyme structural changes with the deprotonation process of Glu 35.

Apart from structural characterizations of small molecules binding to DNA, structural studies on another type of biomacromolecule, protein, provide better understanding on correlation of molecule structural responses to biophysical environment changes. In this paper, NMR is utilized to investigate the influence of pH on the structure of human lysozyme. Besides this, studies of the binding properties of the lipopolysaccharide fragment propyl 3-O-( $\alpha$ -D-glucopyranosyl)- $\beta$ -D-galactofuranoside (Figure 4.20) with both hen egg white (HEWL) and human lysozyme (HL) are present.

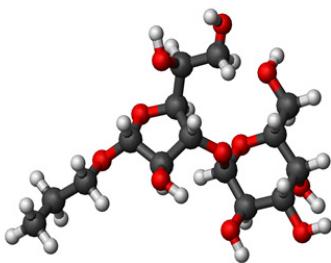


Figure 4.20: Structure of propyl 3-O-( $\alpha$ -D-glucopyranosyl)- $\beta$ -D-galactofuranoside, compound I.

$^1\text{H}$  resonance assignments were obtained for HL and HEWL by transferring chemical shifts from the BMRB entries 5130 for HL [108] and 1093 for HEWL [109] to the NOESY and TOCSY spectra recorded for different lysozymes at different pH. All HN and  $\text{H}\beta$  resonances are assigned (100%), but due to overlap with the water line, ten  $\text{H}\alpha$  assignments are missing. For these ones, their surrounding protons are all successfully assigned.

$^1\text{H}$ -1D NMR titration of HL alone is shown in Figure 4.21. Most of the protein resonances are at the overlapped region below 9 ppm. 1D-NMR spectra were recorded for 21 different pH values in the range of 3.17-8.5 with a separation of ~0.2 units. Four out of five tryptophan  $\text{He1}$  resonances above 9.3 ppm can be monitored in 1D (Figure 4.21). The  $\text{He1}$  resonance of the fifth tryptophan residue (Trp 28) appears in the overlapped region. From 2D titration (below) and 1D, both Trp 28 and 64 do not titrate. The other assigned signals shift with the largest steps in the pH range of 6.7 to 7.7. The most obvious one is  $\text{He1}$  of Trp 109. Note that the line width at pH around 6.8 of Trp 109 is twice as large as the line width at low and high pH (Figure 4.22). This line width broadening between pH 6.1-7.7 can be explained by the dynamics of structural changes, more explanation can be found below. Other resonances such as HN of Cys 77 and Ala 111 also display sizable chemical shift changes. For Ala 111, part of the signals is overlapped. In order to resolve overlapped resonance, ten 2D NOESY titration spectra were recorded in the pH range of 3.8-8.1. Complete assignments allow us to monitor individual shifting resonances. NOESY spectra of selected residues are shown as an example in Figure 4.23. From 2D titration, all the shifted resonances are listed in Table 4.1. Those listed in the upper panel of Table 4.1 are gathered at the binding active site, near Glu 35 (Figure 4.24). Similar pKa values are also obtained for these residues (shown in Figure 4.25, black curves).

From Figure 4.23, HN of Ala 111 and  $\text{He1}$  of Trp 109 have been shown to shift the most among all listed residues in Table 4.1 (Figure 4.23, A1 A2). The significant chemical shift changes of both residues can be ascribed to their close distance to the catalytic residue Glu 35. Hydrogen bond formation is indicated between Ala 111 and Glu 35 (3.15 Å from HN of Ala 111 to Glu 35 oxygen atom in PDB entry 1lzs). Interestingly, line width broadening is also found for Ala 111, but occurs at higher pH 7.4 (black spectrum in Figure 4.23). Different rotational states of Glu 35

side chain were found by comparing the structures with PDB entries of 1jsf and 1lzs. As an assumption, the equilibrium of this motion can lead to an unstable hydrogen bond between HN of Ala 111 and oxygen of Glu 35 at pH 7.4, leading line width broadened for Ala 111.

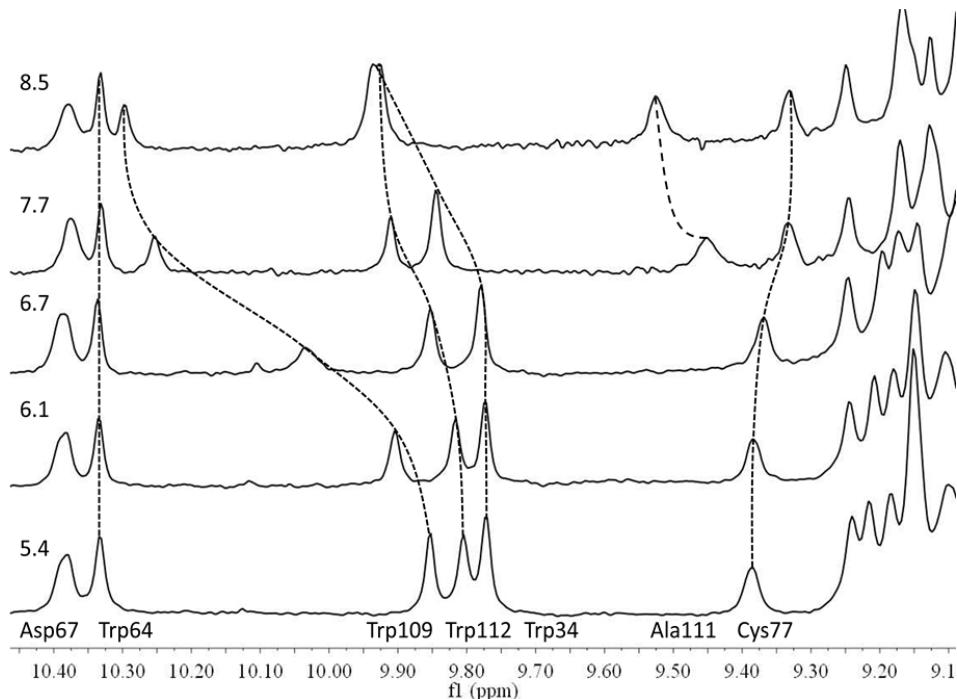


Figure 4.21. 1D NMR titration of HL at spectrum region above 9ppm. Spectra were recorded with 800MHz Varian spectrometer at 35°C. Four out of five Trp residues are marked with assignments. The fifth Trp 28 cannot be monitored since it locates at the overlap region at 9.13ppm. The H $\epsilon$ 1 peak of Trp109, 112, 34 titrate with pH, as well as the HNs of Ala 111, and Cys 77.

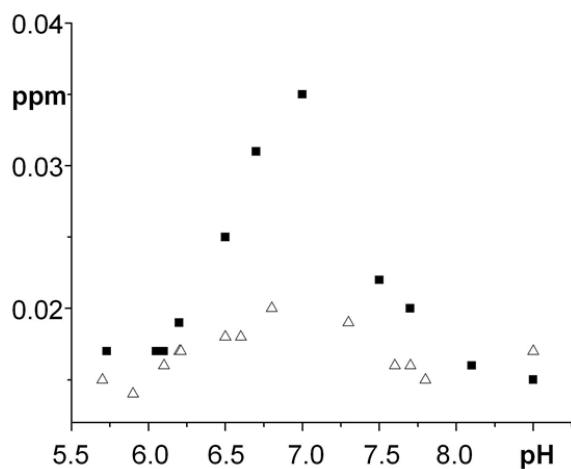


Figure 4.22. Line widths for the resonances for H $\epsilon$ 1 of Trp 109 versus pH for pure HL (filled squares) and for a 1:1 mixture of HL with compound I (triangles). Line widths are only reported when the resonance is not overlapped.

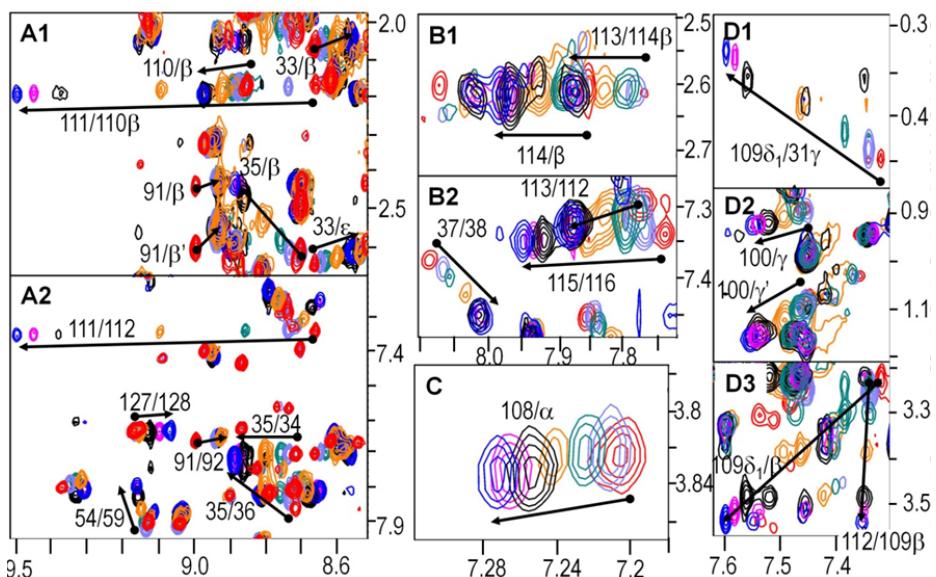


Figure 4.23. Selected 2D NOESY regions of HL alone. Spectra at different pH values are coloured as follows: 3.8 red, 5.0 light blue, 5.5 green, 6.8 orange, 7.4 black, 7.7 purple, 8.1 dark blue. Chemical shifts that vary with pH are indicated by arrows on or next to corresponding peaks with different colours. Peaks are labeled by residue number plus atom name for the assignment on the horizontal axis followed by that for the vertical axis. For space reasons the following items are omitted: residues number of the second atom if it is the same as for the first; atom names in the case of HN. Peak contours are calibrated such that the intensities of HN-HN cross-peaks of helix C (unchanged region) are constant over all pH values.

Besides residues located on the helices near Glu 35, other residues e.g. Ile 59, and Gln 58, Val 100 and Ala 108 also display shift changes. Ring current effect is considered as the reason for their shift changes: the HNs of residues 100 and 108 are found to be in the plane of Trp 109 side chain; the same holds for Gln 58, and in a different crystal structure, also for the one from Ile 59. Plotting the chemical shifts (vertical) versus pH (horizontal) for shifted residues with exception of Aps 91 from Table 4.1 results Figure 4.25. It can be seen from this plot that residues from the upper part of Table 4.1 exhibit identical pKa at 6.8 (black curves), suggesting that they are affected by the same type of chemical or structural changes. In conclusion, there is a close interdependence of the (de)protonation of Glu 35, which is in turn coupled to enzyme activity, and the structural arrangement of the surrounding helices B and D as well as alterations of surrounding hydrogen bonds.

Other residues listed in the lower panel of Table 4.1 are not marked onto the 3D structure. Their pKa values were determined to be different from the ones in the upper part. Reasons for their different pKa values ascribe from different processes. For example, the titratable amino acid His 78 is located near Leu 79 and Asn 75. Earlier NMR results reported a pKa of 7.6 for His 78 [110].

The results for residues 79 and 75 from Figure 4.25 (red and yellow curves) suggest a reduced pKa value of 7.1. Residues 34, 39, 40, and 127 possess unusual high pKa values. Protein denaturation is considered to be the responsible reason. The pKa of Asp 91 was observed to be below 4, a precise pKa could not be obtained, since the titration stopped at pH 3.1. Low pKa value of Asp 91 may be associated to protonation of acidic residues and concomitant protein expansion as proposed earlier [111].

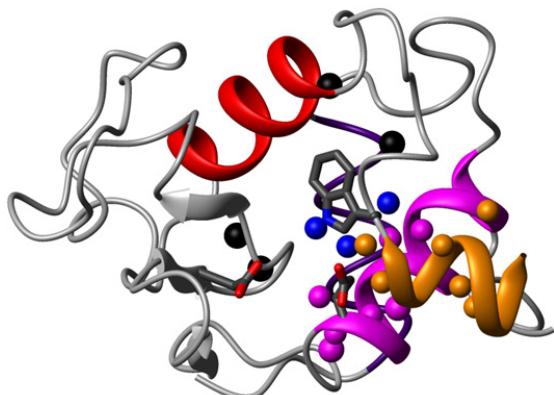


Figure 4.24. Mapping of titrating resonances, represented as spheres, from the upper part of Table 1 onto the 3D structure of HL. Spheres are colour-coded as follows: atoms on helix B are magenta, atoms on helix D are orange, and side chain atoms of Trp 109 are blue. In addition, black spheres also mark the positions of the HNs from the following residues: 58 and 59 (near D53), 100 (at the end of the red helix C), and HN 108 (before Trp 109). The structure is rotated by 30° around a vertical axis with respect to Fig. 1A; the helix colouring is the same, helix A is drawn as a thin curve for clarity.

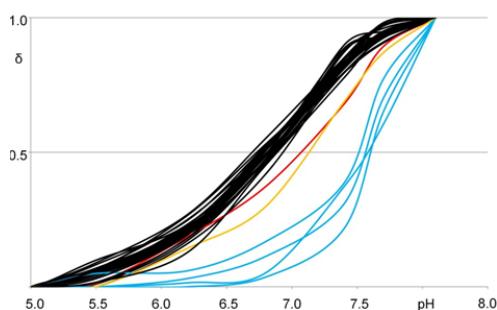


Figure 4.25. Titration curves for the resonances of Table 1 (without the one for HN 91 with a pKa<4), determined from the 2D spectra recorded at various pH values (for selected regions see Figure 3). Resonances listed in the top part of Table 1: black curves, resonances of Asn 75 and Leu 79: yellow and red curves; resonances with pKa >7.5: cyan curves. The vertical axis shows normalised chemical shifts,  $\delta N$ , where for all curves the chemical shifts at pH 5.0 and pH 8.1 are set to 0 and 1, respectively. Note that this provides incorrect intersection points at  $\delta N=0.5$  for high pKa values (>7.5, cyan curves) due to missing measurements at pH>8.1.

**Table 4.1: Titrating resonances in HL upon pH variation in the range 3.8-8.1**

<u>Residue</u>	<u>Atom</u> <sup>1</sup>	<u><math>\Delta\delta_{tot}</math></u> <sup>2</sup>	<u>Comment</u>
----------------	--------------------------	-----------------------------------------------------	----------------

*Titrations with pKa 6.8 near Glu 35 and/or on helices B or D*

31	L	HN	0.09	helix B
31	L	H $\beta$	0.13	helix B
31	L	H $\gamma$	0.12	helix B
33	K	HN	0.10	helix B
35	E	HN	0.17	catalytic residue
35	E	H $\beta$	0.21	catalytic residue
35	E	H $\beta'$	0.14	catalytic residue
35	E	H $\gamma$	0.08	catalytic residue
36	S	HN	0.13	helix B
38	Y	HN	0.08	helix B
58	Q	HN	0.08	on $\beta$ -domain, contact distance to Trp 109
59	I	HN	0.12	on $\beta$ -domain, contact distance to Trp 109
100	V	HN	0.10	contact distance to Trp 109
108	A	HN	0.07	loop preceding helix D
109	W	H $\delta$ 1	0.28	Trp side chain in binding cleft
109	W	H $\epsilon$ 1	0.48	Trp side chain in binding cleft
109	W	H $\beta$	0.32	Trp side chain in binding cleft
110	V	HN	0.12	helix D, shift change also at pH 3.8
111	A	HN	0.84	helix D
112	W	H $\epsilon$ 1	0.13	helix D
113	R	HN	0.21	helix D
114	N	HN	0.13	helix D
115	R	HN	0.11	helix D, interacts with helix B [112]

*Other titrations*

79	L	HN	0.08	pKa 7.1
75	N	H $\delta$	0.078	pKa 7.1
34	W	H $\epsilon$ 1		pKa >7.5
39	N	H $\alpha$		pKa >7.5
40	T	HN		pKa >7.5
127	G	HN		pKa >7.5
91	D	HN		pKa <4.0

<sup>1</sup> Methylene protons are not stereospecifically assigned.

<sup>2</sup> Total change of chemical shift with the pH range 3.8-8.1.

Besides the structural characterizations of pure HL, the binding properties of both HEWL and HL with compound I were demonstrated by chemical shift changes. For HL, binding events were detected by comparing the free lysozyme and mixture spectra at both low pH 3.8 and neutral pH 5.5. At low pH, there was no chemical shift change detected, which means that there is no binding event. In contrast, at this low pH affected residues in HEWL spread in wide range involving a larger number of residues, covering the length of the entire enzymatic cleft (blue spheres in Figure 4.26). At neutral pH 5.5, the same binding sites are revealed for both HL and HEWL (green spheres in Figure 4.26), and the binding occurs at the active site, covering sites D-F. The observation of shift changes for Arg 112 of HEWL, and Arg 113 of HL which are remote from any binding site and quite buried behind the preceding loop residues at pH 5.5 is in agreement with a coupling between ligand binding and conformational effects.

From the results, we can conclude that the binding properties of compound I to HEWL and HL are different: for HEWL at low pH, binding occurs at all subsites, and no binding can be observed for HL. Comparing all 15 different C-type lysozymes from different species (Table 4.2), whose structures are available from the PDB, reveals that differences in electrostatic potential and hydrophobicity of individual lysozyme are responsible for the distinct binding (Figure 4.27 and 4.28). All the 15 different lysozymes are listed with their species names in Table 4.2.

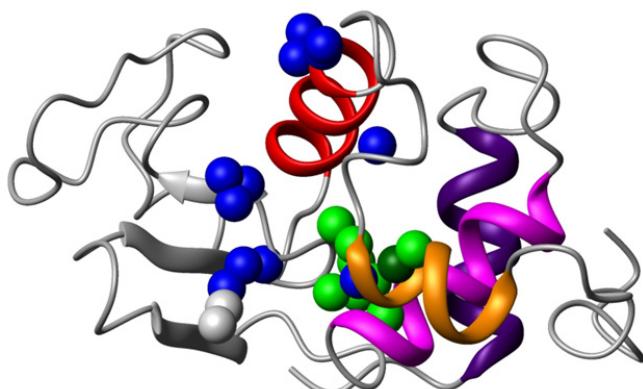
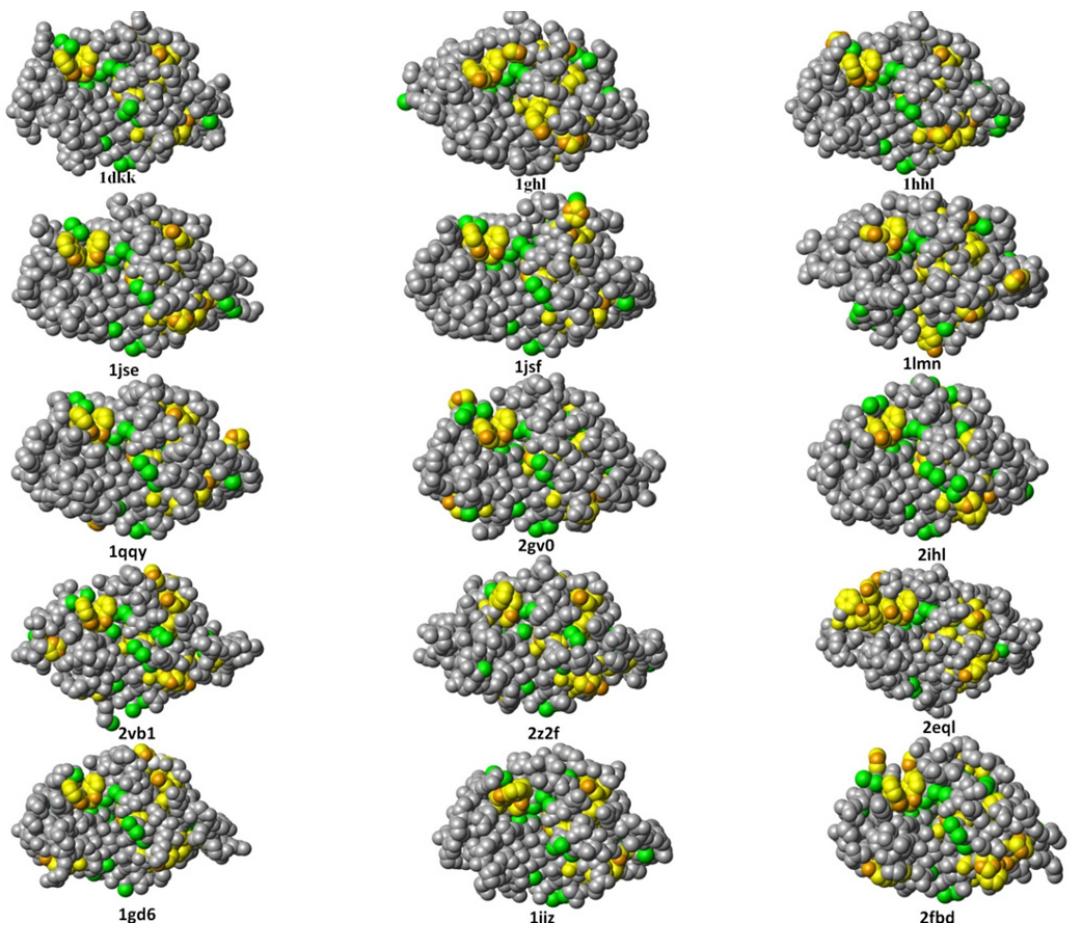


Figure 4.26: Mapping of resonances with chemical shift changes upon addition of compound I onto the 3D structure of lysozyme. Spheres are colored as follows: blue for shift changes observed only for HEWL at pH 3.8, dark green for shift changes observed only for HEWL at pH 5.5, light green for shift changes of both HL and HEWL at pH 5.5, and white for shift changes in all of the three situations.

Lysozyme is a highly common antimicrobial agent against bacteria; for example, in human airways and alveoli it is the most abundant antimicrobial peptide. A prerequisite for the design of novel compounds with improved antibacterial properties, for pharmaceutical as well as food preservation applications. Solid understanding of all aspects, including surface properties and structural rearrangements as illustrated in this study, are relevant for the biological function of lysozyme and its response to exposure to other molecules.

Table 4.2 List of 15 C-type lysozymes with PDB entries and species names.

PDB entry	Species
1dkk	Bobwhite quail
1ghl	Pheasant and guinea-fowl <sup>1</sup>
1hh1	Pheasant and guinea-fowl <sup>2</sup>
1jse	Turky
1jsf	Human
1lmn	Rainbow trout
1qqy	Canine milk
2gv0	Soft-shell turtle
2ihl	Japanese quail
2vb1	HEWL
2z2f	Bovine stomach
2eql	Equine milk
1gd6	Bombyx mori
1iiz	Tasar silkworm
2fdb	Musca domestica



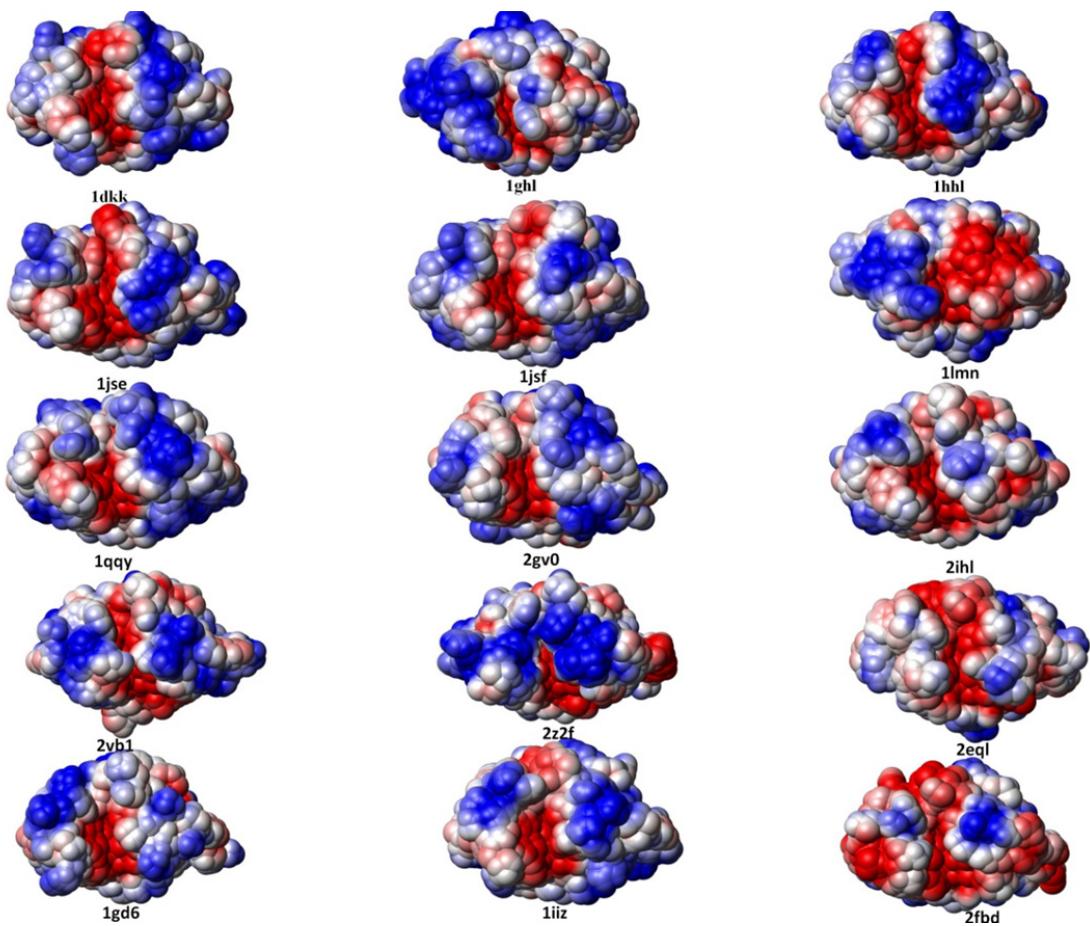


Figure 4.27 Surface representation of 15-C type lysozymes with hydrophobic patches (top) and electrostatic (below). For the former, methyl groups are shown in green and aromatic rings in yellow; orange are side chain oxygen (Tyr) and nitrogen (Trp, His) in aromatic rings. Electrostatic potentials were calculated with MOLMOL using charged Asp, Glu, Arg, Lys and uncharged His residues, with surfaces representing the centre positions of a solvent molecule rolling on the protein atoms for electrostatic potentials. Hydrophobic patches are shown using CPK representations of heavy atoms.

## 5 CONCLUDING REMARKS

NMR structural studies of ligand-DNA interactions have been presented and discussed in this thesis. Important insights on ruthenium compounds binding to DNA were obtained. It has been shown from earlier studies that dimeric Ru-compounds intercalate into DNA, following a fast surface association. The intercalation of the bulky Ru-center requires at least one of the base pairs to open up. Much more severe DNA structural distortions are expected compared to intercalation of monomeric Ru-compounds. To be able to understand the intercalation processes, characterization of the initial interactions is necessary. We have successfully captured the surface bound state with the Dickerson-Drew DNA, prior to intercalation. In this structure, the binuclear Ru compound  $[\mu\text{-}(\text{bidppz})(\text{bpy})_4\text{Ru}_2]^{4+}$  ( $\Delta,\Delta\text{-B}$ ) binds in the minor groove of the DNA, leaving one of the Ru centers near the AATT region, while the other monomer is stabilized by hydrogen bonds to the guanine at the end of the DNA, which helps to define a unique position of the Ru-compound. Most of the interaction enthalpy originates from electrostatic and van der Waals contacts. In another study, a T-T mismatched DNA sequence was designed to include guanines at the ends and adenines in the middle of the sequence. Preliminary results indicate that the same Ru-compound  $\Delta,\Delta\text{-B}$  intercalated to the mismatch DNA at the AATT region. Interestingly 1D NMR-titration suggested that one  $\Delta,\Delta\text{-B}$  associates two DNA duplexes. Future work will focus on full NOESY assignments, which are required to be able to perform detailed structural calculations.

Insertion of a second Ru-compound,  $[\mu\text{-}(\text{bidppz})(\text{phen})_4\text{Ru}_2]^{2+}$  ( $\Delta,\Delta\text{-P}$ ) into a hexamer DNA was studied by X-ray crystallography. The 3D structure showed that  $\Delta,\Delta\text{-P}$  inserts itself into DNA from the minor groove, extruding one TA base pair into the minor groove, leaving dppz exposed to solvent. Meanwhile, the second half of dppz moiety protruded from the opposite major groove, recruiting another DNA duplex. A cross-linking like structure was proposed. Low emission quantum yield confirmed that the intercalated dppz motif is indeed exposed to solvent.

In another study, two Hoechst 33258 molecules were indicated to bind in the minor groove of DNA consecutively with small separation apart, one after another, and not in a sandwich structure as previously proposed. This shows that multi ligand molecules can accommodate within a close distance of DNA.

Besides the work in the context of DNA-ligand interactions presented in this thesis, relations between the active-site state and the overall structure of human lysozyme were also assessed. Domain-wide structural rearrangements were found to accompany the deprotonation of Glu 35, including movements of the neighboring helices, reorientation of Trp 109, and variations of hydrogen bonds pattern. Similar structural changes were also observed upon binding of a LPS fragment from *Klebsiella pneumoniae* to various lysozymes. Protein surface properties such as

electrostatic potential and hydrophobic patches significantly modulate the interaction of a disaccharide near the remote subsites A and B, while binding affinity near the active site appears less dependent. A better understanding of all aspects of the response of lysozyme to the contact with other particles is a prerequisite for the design of novel antibacterial compounds with altered and improved properties.

## **6 ACKNOWLEDGEMENTS**

I would like to thank my main supervisor Prof. Bengt Nordén for introducing me to the biophysical chemistry research field, and providing supports, and guidance. Thank you for always being very enthusiasm, humble and thank you for leading me to see the gems in the piles of gravel.

I would like to thank my co-supervisor Prof. Per Lincoln, who brings me in the ruthenium world-a world that is filled with rocks and sands. Without your patience and encouragements, I could hardly step out. Thank you for always believing in me!

My co-supervisor, Prof. Martin Billeter, deserves my special thanks. Thank you to give me golden knowledge about NMR. It is you who has always inspired me and let me see the brightness before the sun rises.

I would like to express my gratitude to Prof. Hans-Christian Siebert, thank you to inspire us with the new projects in lysozyme filed, and thank you for your contributions and suggestions for the paper preparation.

I would like to thank my colleges and my friends Pär Nordell, who taught me how to operate on CD and other instruments in physical chemistry department, Johanna Anderson, Nils Carlsson who are very warm hearted, and helped me in many aspects throughout my PhD. And the co-authors, Anna Reymer who did the MD simulation, and Prof. Tom Brown who provided us valuable materials for experiments and suggestions for publication.

I am very grateful for the helps from the administrators in KB department, Gunilla Saethe, Anna Molander, Carina Pettersson. I must also express my appreciation to the Swedish NMR center, especially to Profs. Göran Karlsson, Vladislav Orekhov and Max Meyer, Cecilia Persson, Diana Bernin and Anders Pedersen, thank you for helping me!

My life away from the lab has been colorful, thank you my dear friends, Lulu, Xianchan, Weizhen and many others, I also would like to thank the Chinese school to provide me the opportunity to become a good Chinese teacher!

Finally I would like to thank my family members for their unconditional love, support and encouragement. Baba, mama, xie xie ni men! This thesis is dedicated for you and to my fiancé, my best friend, Joakim Åling.



## 7 REFERENCES

1. McCulloch, S. D., and Kunkel, T. A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.* 18 (1), 148-161.
2. Mizrahi, V., Henrie, R. N., Marlier, J. F., Johnson, K. A., and Benkovic, S. J. (1985). Rate-limiting steps in the DNA polymerase I reaction pathway. *Biochemistry*, 24 (15), 4010–4018.
3. Bruck, I., and O'Donnell, M. (2001). The ring-type polymerase sliding clamp family. *Genome Biol.* 2(1), Reviews3001.
4. The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.
5. Neidle, S., and Thurston, D. E. (2005). Chemical approaches to the discovery and development of cancer therapies. *Nat. Rev. Cancer*, 5(4), 285-296.
6. Wong, E., and Giandomenico, C. M. (1999). Current status of platinum-based antitumor drugs. *Chem. Rev.* 99(9), 2451-2466.
7. Hurley, L. H. (2002). DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer*, 2(3), 188-200.
8. Smith, C. K., Brannigan, J. A., and Moore, M. H. (1996). Factors effect DNA sequence selectivity of nogalamycin intercalation: the crystal structure of d(TGTACA)2-nogalamycin(2). *J. Mol. Biol.*, 263, 237-258.
9. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., and Dickerson, R. E. (1985). The molecular-origin of DNA drug specificity in Netrosin and Distamycin. *Proc. Natl. Acad. Sci.* 82, 1376-1380.
10. Abu-Daya, A., Brown, P. M., and Fox, K. R. (1995). DNA sequence preferences of several AT-selective minor groove binding ligands. *Nucleic Acids Res.* 23, 3385-3392.
11. Abu-Daya, A., and Fox, K. R. (1997). Interaction of minor groove binding ligands with long AT tracts. *Nucleic. Acids. Res.* 25, 4962-4969.
12. Kubista, M., Åkerman, B., and Nordén, B. (1987). Characterization of Interaction between DNA and 4',6-Diamidino-2-phenylindole by opticalspectroscopy. *Biochemistry*, 26, 4545-4553.
13. Wilson, W. D., Tanious, F. A., Barton, H. J., Jones, R. L., Fox, K., Wydra, R. L., and Strekowski, L. (1990). DNA sequence dependent binding modes of 4', 6-diamidino-2-phenylindole (DAPI). *Biochemistry*, 29, 8452-8461.
14. Eriksson, S., Kim, S. K., Subista, M., and Nordén, B. (1993). Binding of 4',6-diamidino-2-phenylindole (DAPI) to DNA AT regions of DNA-Evidence for an allosteric conformatyional change. *Biochemistry*, 32, 2987-2998.
15. Accelrys Draw 4.1. Release Notes. Accelrys Software Inc. 2012.
16. Niyazi, H., Hall, J. P., Sullivan, K. O., Winter, G., Sorensen, T., Kelly, J. M., and Cardin, C. J. (2012). Crystal structure of Lambda-[Ru(phen)(2)dppz](2+) with oligonucleotides containing TA/TA and AT/AT steps show two intercalation mode. *Nat. Chem.* 4, 621 –628.

17. Song, H.; Kaiser, J. T.; and Barton, J. K. (2012). Crystal structure of Delta-[Ru(bpy)(2)dppz(2+)] bound to mismatched DNA reveals side-by-side metalloinsertion and intercalation. *Nat. Chem.* 4, 615 – 620.
18. Nordell, P., Westerlund, F., Wilhelmsson, M., Nordén, B., and Lincoln, P. (2007). Kinetic recognition of AT-rich DNA by ruthenium complexes. *Angewandte. Chemie.* 46, 2203-2206.
19. Önfelt, B., Lincoln, P., and Nordén, B. (1999). A molecular stple for DNA: Threading bis-intercalating [Ru(phen)(2)dppz](2+) dimer . *J. Am. Chem. Soc.* 121, 10846-10847.
20. Wilhelmsson, L. M., Westerlund, F., Lincoln, P., and Nordén, B. (2002). DNA-binding of semirigid binuclear ruthenium complex delta,delta-[mu-(11,11'-bidppz)(phen)4Ru2]4+: extremely slow intercalation kinetics. *J. Am. Chem. Soc.* 124, 12092 –12093.
21. Proctor, V. A., and Cunningham, F. E. (1988). The chemistry of lysozyme and its use as a food preservative and a pharmaceutical. *Crit. Rev. Food Sci. Nutr.* 26, 359–395.
22. Koradi, R., Billeter, M. and Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 51-55.
23. Brown, T., Hunter, W. N., Kneale, G., and Kennard, O. (1986). Molecular structure of the G.A base pair in DNA and its implications for the mechanism of transversion mutations. *Proc. Natl. Acad. Sci.* 83(8), 2402-2406.
24. Skelly, J. V., Edwards, K. J., Jenkins, T. C., and Neidle, S. (1993). Crystal structure of an oligonucleotide duplex containing G.G base pairs: Influence of mispairing on DNA backbone conformation. *Proc. Natl. Acad. Sci.* 90(3), 804-808.
25. Hunter, W. N., Brown, T., and Kennard, O. (1987). Structural features and hydration of a dodecamer duplex containing two C.A mispairs. *Nucleic Acids Res.* 15(16), 6589-6605.
26. Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D., and Kennard, O. (1987). The structure of guanosine-thymidine mismatches in B-DNA at 2.5-a resolution. *J. Biol. Chem.* 262(21), 9962-9970.
27. Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia. J. (1999). Nearest-neighbor thermodynamics and NMR of DNA sequences with internal AA, CC, GG, and TT mismatches. *Biochemistry*, 38(12), 3468-3477.
28. SantaLucia, J., and Hicks, D. (2004). The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* 33(1), 415-440.
29. Arnold, F. H., Wolk, S., Cruz, P., and Tinoco, I. (1987). Structure, dynamics, and thermodynamics of mismatched DNA oligonucleotide duplexes d(CCCAGGG)2 and d(CCCTGGG)2. *Biochemistry*, 26(13), 4068-4075.
30. Patel, D. J., Kozlowski, S. A., and Ikuta, K. (1984). Dynamics of DNA duplexes containing internal GT, GA, AC, and TC pairs-hydrogen-exchange at and adjacent to mismatch sites. *Fed. Proc.* 43(11), 2663-2670.
31. Li, G. M. (2008). Mechanisms and functions of DNA mismatch repair. *Cell Res.* 18, 85-98.

32. Kumar, C. V., Barton, J. K., and Turro, N. J. (1985). Photophysics of Ruthenium complexes bound to double helical DNA. *J. Am. Chem. Soc.* 107, 5518-5523.
33. Barton, J. K., Goldberg, J. M., Kumar, C. V., and Turro, N. J. (1986). Binding modes and base specificity of Tris(phenanthroline)ruthenium (II) enantiomers with nucleic-acids-tuning the stereoselectivity. *J. Am. Chem. Soc.* 108, 2081-2088.
34. Rehmann, J. P., and Barton, J. K. (1990). H-1-NMR studies of Tris(phenanthroline) metal complexes bound to oligonucleotides- characterization of binding modes. *Biochemistry*, 29, 1707-1709.
35. Hiort, C., Nordén B., and Rodger, A. (1990). Enantioselective DNA-binding of [Ru(II)(1,10-phenanthroline)3]2+ studied with linear and circular-dichroism. *J. Am. Chem. Soc.* 112, 1971-1982.
36. Satyanarayana, S., Dabrowiak, J. C., and Chaires, J. B. (1992). Neither delta- nor lambda-tris(phenanthroline)ruthenium (II) binds to DNA by classical intercalation. *Biochemistry*, 31, 9319-9324.
37. Satyanarayana, S., Dabrowiak, J. C., and Chaires, J. B. (1993). Tris(phenanthroline) ruthenium (II) enantiomer interactions with DNA: mode and specificity of binding. *Biochemistry*, 32, 2573-2584.
38. Eriksson, M., Leijon, M., Hiort, C., Nordén, B., and Gräslund, A. (1994). Binding of delta-[Ru(phen)3]2+ and lambda-[Ru(phen)3]2+ to [d(CGCGATCGCG)]2 studied by NMR. *Biochemistry*, 33, 5031-5040.
39. Barton, J. K., Goldberg, J. M., Kumar, C. V., and Turro, N. J. (1986). Binding modes and base specificity of tris(phenanthroline)ruthenium(II) enantiomers with nucleic-acids-tuning the spectroselectivity. *J. Am. Chem. Soc.* 108, 2081-2088.
40. Chambron, J. C., Sauvage, J. P., Amouyal, E., and Koffi, P. (1985). Ru(Bipy)2(Dipyridophenazine)2+ - a complex with a long-range directed charge transfer excited-state. *New. J. Chem.* 9, 527-529.
41. Friedman, A. E., Chambron, J. C., Sauvage, J. P., Turro, N. J., and Barton, J. K. (1990). Molecular light switch for DNA – Ru(bpy)2(dppz)2. *J. Am. Chem. Soc.* 112, 4960-4962.
42. Jenkins, Y., Friedman, A. E., Turro, N. J., and Barton, J. K. (1992). Characterization of dipyridophenazine complexes of ruthenium (II)- the light switch effect as a function of nucleic-acid sequence and conformation. *Biochemistry*, 31, 10809-10816.
43. Olson, E. J. C., Hu, D., Ormann, A. H., Jonkman, A. M., Arkin, M. R., Stemp, E. D. A., Barton, J. K., and Barbara. P. F. (1997). First observation of the key intermediate in the “light-switch” mechanism of [Ru(phen)2dppz]2+. *J. Am. Chem. Soc.* 119(47):11458-11467.
44. Önfelt, B., Lincoln, P., Nordén, B., Spencer Baskin, J., and Zewail, A. H. (2000). Femtosecond linear dichroism of DNA-intercalating chromophores: Solvation and charge separation dynamics of [Ru(phen)2dppz]2+ systems. *Proc. Natl. Acad. Sci.* 97(11), 5708-5713.

45. Dupureur, C. M., and Barton, J. K. (1994). Use of selective deuteration and  $^1\text{H}$  NMR in demonstrating major groove binding of Delta-[Ru(phen)2dppz] $^{2+}$  to d(GTCGAC)2. *J. Am. Chem. Soc.* 116, 10286-10287.
46. Lincoln, P., Broo, A., and Nordén, B. (1996). Diastereomeric DNA-binding geometries of intercalated ruthenium (II) trischelates probed by linear dichroism: [Ru(phen)2DPPZ] $^{2+}$  and [Ru(phen)2BDPPZ] $^{2+}$ . *J. Am. Chem. Soc.* 118, 2644-2653.
47. Tuite, E., Loncoln, P., and Nordén, B. (1997). Photophysical evidence that Delta- and Lamdba- [Ru(phen)2(dppz)] $^{2+}$  intercalate DNA from the minor groove. *J. Am. Chem. Soc.* 119, 239-240.
48. Holmlin, R. R., Stemp, E. D. A., and Barton, J. K. (1998). Ru(phen)(2)dppz( $^{2+}$ ) luminescence: Dependence on DNA sequences and groove-binding agents. *Inorg. Chem.* 37, 29-34.
49. Hall, J. P., Beer, H., Buchner, K., Cardin, D., and Cardin, C. (2013). Preferred orientation in an angled intercalation site of a chloro-substituted  $\Lambda$ -[Ru(TAP)2(dppz)] $^{2+}$  complex bound to d(TCGGCGCCGA)2. *Phil. Trans. R. Soc. A* 371: 20120525.
50. Müller, W., and Crothers, D. M. (1968). Studies of the binding of actinomycin and related compounds to DNA. *J. Mol. Biol.* 35, 251-290.
51. Önfelt B., Lincoln, P., and Nordén, B. (1999). A molecular staple for DNA: threading bisintercalating [Ru(phen)2dppz] $^{2+}$  dimer. *J. Am. Chem. Soc.* 121, 10846-10847.
52. Önfelt B., Lincoln, P., and Nordén, B. (2001). Enantioselective DNA threading dynamics by phenazine-linked. *J. Am. Chem. Soc.* 123, 3630-3637.
53. Önfelt B. (2002). Chiral ruthenium dipyridophenazine complexes: DNA binding and excited state relaxation; Chalmers university of Technology: Gothenburg.
54. Wilhelmsson, L. M., Esbjörner, E., Westerlund, F., Nordén, B., and Lincoln, P. (2003). Meso stereoisomer as a probe of enantioselective threading intercalation of semirigid ruthenium complex [ $\mu$ -(11,11'-bidppz)(phen)4Ru2] $^{4+}$ . *J. Phys. Chem. B.* 107, 11784-11793.
55. Nordell, P., Westerlund, F., Reymer, A., El-sagheer, A. H., Brown, T., Nordén, B., and Lincoln, P. (2008). DNA polymorphism as an origin of adenine-thymine tract length-dependent threading intercalation rate. *J. Am. Chem. Soc.* 130, 14651-14658.
56. Kogan, M., Nordén, B., Lincoln, P., and Nordell, P. (2011). Transition state of rare event base pair opening probed by threading into looped DNA. *Chem. Bio. Chem.* 12, 2001-2006.
57. Ojha, H., Murari, B. M., Anand, S., Hassan, M. I., Ahmad, F., and Chaudhury, N. K. (2009) Interaction of DNA Minor Groove Binder Hoechst 33258 with Bovine Serum Albumin. *Chem. Pharm. Bull.* 57, 481-486.
58. Joubert, A., Sun, X.-W., Johansson, E., Bailly, C., Mann, J., and Neidle, S. (2003). Sequence-Selective Targeting of Long Stretches of the DNA Minor Groove by a Novel Dimeric Bis-benzimidazole. *Biochemistry*, 42, 5984-5992.

59. Baraldi, P. G., Bovero, A., Fruttarolo, F., Preti, D., Tabrizi, M. A., Pavani, M. G., and Romagnoli, R. (2004). DNA Minor Groove Binders as Potential Antitumor and Antimicrobial Agents. *Med. Res. Rev.* 24, 475–528.
60. Chiang, S. Y., Welch, J., Rauscher, F. J., and Beerman, T. (1994). Effects of Minor Groove Binding Drugs on the Interaction of TATA Box Binding Protein and TFIIA with DNA. *Biochemistry* 33, 7033–7040.
61. Downs, T. R., and Wilfinger, W. W. (1983). Fluorometric Quantification of DNA in Cells and Tissue. *Anal. Biochem.* 131, 538–547.
62. Goracci, L., Germani, R., Savelli, G., and Bassani, D. M. (2005). Hoechst 33258 as a pH-Sensitive Probe to Study the Interaction of Amine Oxide Surfactants with DNA. *Chem. Bio. Chem.* 6, 197–203.
63. Barooah, N., Mohanty, J., Pal, H., Sarkar, S. K., Mukherjee, T., and Bhasikuttan, A. C. (2011). pH and Temperature Dependent Relaxation Dynamics of Hoechst-33258: a Time Resolved Fluorescence Study. *Photochem. Photobiol. Sci.* 10, 35–41.
64. Neidle, S. (2001). DNA Minor-Groove Recognition by Small Molecules. *Nat. Prod. Rep.* 18, 291–309.
65. Moon, J. H., Kim, S. K., Sehlstedt, U., Rodger, A., and Nordén, B. (1996). DNA Structural Features Responsible for Sequence-Dependent Binding Geometries of Hoechst 33258. *Biopolymers*. 38, 593–606.
66. Pjura, P. E., Grzeskowiak, K., and Dickerson, R. E. (1987). Binding of Hoechst 33258 to the Minor Groove of B-DNA. *J. Mol. Biol.* 197, 257–271.
67. Fede, A., Billeter, M., Leupin, W., and Wüthrich, K. (1993). Determination of the NMR Solution Structure of the Hoechst 33258-d-(GTGGAATTCCAC)2 Complex and Comparison with the X-ray. *Crystal Structure. Structure* (Cambridge, MA, U. S.) 1, 177–186.
68. Blake, C. C. F., Koenig, D. F., Mair, G.A., North, A. C. T., Phillips, D. C., and Sarma, V.R. (1965). Structure of hen egg-white lysozyme. *Nature (London)*, 206, 757-761.
69. Wang, J., Dauter, M., Alkire, R., Joachimiak, A., and Dauter, Z. (2007). Triclinic lysozyme at 0.65Å resolution. *Acta Crystallogr. Sect. D*, 63, 1254.
70. Cheetham, J. C., Artymiuk, P. J., and Phillips, D. C. (1992) Refinement of an enzyme complex with inhibitor bound at partial occupancy. *J. Mol. Biol.* 224 613–628.
71. Vocadlo, D. J., Davies, G. J., Laine, R., and Withers, S. G. (2001). Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate; *Nature (London)*. 412, 835–838.
72. Kuramitsu, S., Ikeda, K., Hamaguchi, K., Fujio, H., Amano, T-. Miwa, S., and Nishina T. (1974). Ionization constants of Glu35 and Asp52 in hen, turkey, and human lysozymes. *J. Biochem.*, 76, 671-683.
73. Reitamo, S., Klochars, M., Adinolfi, M., and Osserman, E. F. (1978). Human lysozyme (origin and distribution in health and disease). *Ric. Clin. Lab.* 8, 211-231.

74. Hansen, N. E., Karle, H., Andersen, V., and Olgaard, K. (1972). Lysozyme turnover in man. *J. Clin. Invest.* 51, 1146-1155.
75. Blake, C. C. F., Grace, D. E., Johnson, L. N., Perkins, S. J., Phillips, D. C., Cassels, R., Dobson, C. M., Poulsen, F. M., and Williams, R. J. (1977). Physical and chemical properties of lysozyme. *Ciba Found. Symp.* 137-185.
76. Song, H., Inaka, K., Maenaka, K., and Matsushima, M. (1994) Structural changes of Active site cleft and different saccharide binding modes in human lysozyme co-crystallized with hexa-N-acetyl-chitohexaose at pH 4.0. *J. Mol. Biol.* 244, 522-540.
77. Artymiuk, P. J., and Blake, C. C. F. (1981). Refinement of human lysozyme at 1.5 Å resolution. Analysis of non-bonded and hydrogen-bond interactions. *J. Mol. Biol.* 152, 737-762.
78. Mine, S., Ueda, T., Hahimoto, Y., and Imoto, T., (2000) Analysis of the internal motion of free and ligand-bound human lysozyme by use of <sup>15</sup>N NMR relaxation measurement: A comparison with those of hen lysozyme. *Protein Sci.* 9, 1669-1684.
79. Kumeta, H., Miura, A., Kobashigawa, Y., Miura, K., Oka, C., Nemoto, N., Nitta, K., and Tsuda, S. (2003). Low-temperature-induced structural changes in human lysozyme elucidated by three-dimensional NMR spectroscopy. *Biochemistry* 2003, 42, 1209-1216.
80. Young, A. C. M., Tilton, R. F. and Dewan, J. C. (1994). Thermal expansion of hen egg-white lysozyme. *J. Mol. Biol.* 235, 302-317.
81. Refaee, M., Tezuka, T., Akasaka, K., and Williamson, M.P. (2003). Pressure-dependent changes in the solution structure of hen egg-white lysozyme. *J. Mol. Biol.* 327, 857-885.
82. Deckers, D., Vanlint, D., Callewaert, L., Aertsen, A., and Michiels, C. W. (2008). Role of Ivy lysozyme inhibitor in growth or survival of *Escherichia coli* and *Pseudomonas aeruginosa* in hen egg albumen and in human saliva and breast milk. *Appl. Environ. Microbiol.* 74, 4434-4439.
83. Maga, E. A., Walker, R. L., Anderson, G. B., and Murray, J. D. (2006). Consumption of milk from transgenic goats expressing human lysozyme in the mammary gland results in the modulation of intestinal microflora. *Transgenic Res.* 15, 515-519.
84. Brundige, D. R., Maga, E. A., Klasing, K. C., and Murray, J. D. (2008). Lysozyme transgenic goats' milk influences gastrointestinal morphology in young pigs; *J. Nutr.* 138, 921-926.
85. Cole, A. M., Liao, H-I., Stuchlik, O., Tilan, J., Pohl, J., and Ganz, T. (2002). Cationic polypeptides are required for antibacterial activity of human airway fluid. *J. Immunol.* 169, 6985-6991.
86. Tenovuo J. (2002). Clinical applications of antimicrobial host proteins lactoperoxidase, lysozyme and lactoferrin in xerostomia: efficacy and safety. *Oral diseases.* 8, 23-29.
87. Herreweghe, J. M. V., and Michiels, C. (2012). Invertebrate lysozymes: Diversity and distribution, molecular mechanism and in vivo function. *J. Biosci.* 37(2) 327-348.

88. Sava, G., Pacor, S., and Dasci G. (1995). Lysozyme stimulates lymphocyte response to ConA and IL-2 and potentiates 5-fluorouracil action on advanced carcinomas. *Anticancer research*, 15, 5B, 1883-1888.
89. Lee-Huang S., Huang P.L., Sun, Y., Kung, H. F., Blithe, D. L., and Chen, H.C. (1999). Lysozyme and RNases as anti-HIV components in beta-core preparations of human chorionic gonadotropin. *Proc. Natl. Acad. Sci.* 96(6):2678-2681.
90. Rule, G. S., and Hitchens, T. K. (2006). Fundamentals of protein NMR spectroscopy. 1<sup>st</sup> ed. Springer, Netherlands.
91. Barber, J. C., and Parkinson, J. A. (1993). High-Resolution NMR of DNA and Drug-DNA Interactions. In: *Methods in Molecular Biology, Spectroscopic Methods and Analyses: NMR, Mass Spectrometry, and Metalloprotein Techniques*.
92. Bain, A. D. Chemical exchange in NMR. *Prog. Nucl. Mag. Res. SP.* 43 (2003) 63-103.
93. Neuhaus, D., and Williamson, M. (1989). The nuclear overhauser effect in structural and conformational analysis. VCH, UK.
94. Lakowicz, J. R. (2006). Principles of fluorescence spectroscopy. 3<sup>rd</sup> ed. New York: Springer.
95. Nordén, B., Rodger, A., and Dafforn, T. (2010). Linear dichroism and circular dichroism. A textbook on polarized-light spectroscopy. Cambridge: RSC publishing.
96. Hollas, J. M. (1995). Morden spectroscopy 3<sup>rd</sup> ed. Chichester: J. Wiley.
97. Güntert, P. (2003). Automated NMR protein structure calculation. *Prog. Nucl. Mag. Res. SP.* 43, 105-125.
98. Kirkpatrick, S., Gelatt, C.D. and Vecchi, M. P.(1983). Optimization by simulated annealing. *Science*. 220, 671-680.
99. Wu, L., Reymer, A., Persson, C., Kazimierczuk, K., Brown, T., Lincoln, P., Nordén, B., and Billeter, M. (2013). Initial DNA Interactions of the Binuclear Threading Intercalator,  $\Lambda,\Lambda\text{-}[\mu\text{-}bipy_4Ru_2]^{4+}$ : An NMR Study with [d(CGCGAATTCTCGCG]2. *Chem. Eur. J.* 19, 5401-5410.
100. Claridge, T. (2009). Mnova: NMR data processing, analysis, and prediction software. *J. Chem. Inf. Model.* 49, 1136-1137.
101. Evans, J. N. S. (1995). Biomolecular NMR spectroscopy. 1<sup>st</sup> ed. Oxford University Press, UK.
102. Cohen, J. S. (1987). 2D-NOESY of DNA: as easy as A, B, Z. *Trends. Biochem. Sci.* 121, 33-135.
103. Vranken, W. F., Boucher, W., Stevens, T., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E., Markley, J. L., Ionides, J., and Laue, E. D. (2005), *Proteins Struct. Funct. Bioinf.* 59, 687-696.
104. Fornander, L. H., Wu, L., Billeter, M., Lincoln, P., and Nordén, B. (2013). Minor-groove binding drugs: where is the second Hoechst 33258 molecule? *J. Phys. Chem. B.* 117, 5820-5830.

105. Higgins, L. D., and Searle, M. S. (1999). Site-Specificity of Bis-benzimidazole Hoechst 33258 in A-Tract Recognition of the DNA Dodecamer Duplex d(GCAAAATTTGC)<sub>2</sub>. *Chem. Commun.* 1861–1862.
106. Fede, A., Labhardt, A., Bannwarth, W., and Leupin, W. (1991). Dynamics and Binding Mode of Hoechst 33258 to d(GTGGAAATTCCAC)2 in the 1:1 Solution Complex as Determined by Two-Dimensional <sup>1</sup>H NMR. *Biochemistry* 30, 11377–11388.
107. Embrey, K. J., Searle, M. S., and Craik, D. J. (1993). Interaction of Hoechst 33258 with the Minor Groove of the AT-Rich DNA Duplex d(GGTAATTACC)2 Studied in Solution by NMR Spectroscopy. *Eur. J. Biochem.* 211, 437–447.
108. Redfield, C., and Dobson, C. M. (1990). <sup>1</sup>H NMR studies of Human Lysozyme: Spectral assignment and comparison with hen lysozyme. *Biochemistry* 29, 7201-7214.
109. Redfiled, C., and Dobson, C. M. (1988). Sequential <sup>1</sup>H NMR assignments and secondary structure of Hen Egg White Lysozyme in solution. *Biochemistry* 27, 122-136.
110. Meadows, D.H., Markley, J. L., Cohen, H. S., and Jardetzky, O. (1967). Nuclear magnetic resonance studies of the structure and binding sites of enzymes, I. histidine residues. *Proc. Natl. Acad. Sci.* 58, 1307-1313.
111. Bonincontro, A., Francesco, A. De., and Onori, G. (1998). Influence of pH on lysozyme conformation revealed by dielectric spectroscopy. *Colloids Surf. B* 12 , 1-5.
112. Harata, K., Muraki, M. and Jigami, Y. (1993). Role of Arg 115 in the catalytic action of human lysozyme X-ray structure of His115 and Glu115 mutants. *J. Mol. Biol.* 233, 524-535.