The Dissertation Committee for Debashish Sahu Certifies that this is the approved version of the following Dissertation:

NMR STUDIES ON KINETICS OF PROCESSES INVOLVING DNA-BINDING PROTEINS

Committee:

Junji Iwahara, Ph. D., Supervisor

James C. Lee, Ph. D.

Wlodek M. Bujalowski, Ph. D.

Shankar Mitra, Ph. D.

Xiaodong Cheng, Ph. D.

John A. Putkey, Ph. D.

Dean, Graduate School

NMR STUDIES ON KINETICS OF PROCESSES INVOLVING DNA-BINDING PROTEINS

by

Debashish Sahu, B. Tech.

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas Medical Branch
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philisophy

Approved by the Supervisory Committee
Junji Iwahara, Ph.D.
James C. Lee, Ph.D.
Wlodek M. Bujalowski, Ph.D.
Sankar Mitra, Ph.D.
Xiaodong Cheng, Ph.D.
John A. Putkey, Ph.D.

The University of Texas Medical Branch

August 2011

Galveston, Texas

Keywords: Protein-DNA interactions, Kinetics, Translocation, Redox Reactions, Nuclear Magnetic Resonance (NMR)

To my parents, Dr. Hrushi Kesh Sahu and Jyotirmayee Sahu, for their constant support and encouragement in helping me persue my dreams.

ACKNOWLEDGEMENTS

This work would not be possible without the support of numerous people that have guided me through various obstacles during my graduate school. I'm in debt of their support and encouragement.

The journey of science in graduate school was very productive due to the unparalleled support and guidance by my mentor Dr. Junji Iwahara, from whom I have learned valuable lessons in approaching scientific problems and successfully tackling them. He has been a great role model and his constructive criticisms and advice will continually influence my scientific career choices. I thank my supervisory committee for their constant support and useful critiques, which has been crucial in achieving my research goals.

I also appreciate the support of current and past members of the Iwahara laboratory, where constant interactions on scientific issues have helped me troubleshoot problems while gaining a different perspective on the scientific matters. I would like to specifically acknowledge Dr. Yuki Takayama for the time and effort he spent to help me learn and hone the essential molecular biology skills needed to become a successful scientist.

The various biophysical approaches used in this dissertation have been made possible by the very skilled support by the facility managers of the core facilities at UTMB. I would also like to thank Dr. Luis Marcelo Holthauzen for supporting the solution biophysics core facility. I want to express my gratitude to Dr. Tianzhi Wang, current NMR facility manager and past managers who have gone above and beyond their duties to keep the NMR machines running. I would also like to acknowledge the Sealy Center for Structural Biology and John S. Dunn Gulf Coast Consortium for Magnetic Resonance for providing support and access to the valuable NMR instruments.

The department of biochemistry and molecular biology has played a crucial role in my scientific development. This department has constantly supported the current and past members of Biological Student Organization (BCSO) in their professional and social

development. A few people who have played vital roles in my progress are Drs. Wayne Bolen, Sarita Sastry, Kyung Choi, Lillian Chan and Andres Oberhauser.

I would like to express my thanks to all the administrative secretaries at Biochemistry and Molecular biology especially, Debora Botting, Rose Byrdlon-Griggs, Lori Blackwell, Angelina Johnson and Shirley Broz whose help is irreplaceable. I also thank Ms. Lisa Pipper for her computer-related support.

I would like to thank all my teachers who have played an instrumental role is shaping me as a scientist. Also, a very special thanks to all my friends for the scientific, social and moral support during my graduate school.

Finally I would like to acknowledge my source of inspiration, my parents who have supported my decisions and shaped me to become a better person. The unconditional love and support by my family, including my sister Ms. Ankita Sahu has been truly instrumental in my success.

NMR STUDIES ON KINETICS OF PROCESSES INVOLVING DNA-BINDING PROTEINS

Publication No._____

Debashish Sahu, Ph.D.

The University of Texas Medical Branch, 2011

Supervisor: Junji Iwahara

Macromolecular interactions do not just encompass the coming together of two or more macromolecules, but rather the speed of the interactions play a very important role in proper cellular function. In response to external stimuli, DNA binding proteins very efficiently carry out their cognate tasks by making use of different kinetic mechanisms. Here we quantitatively analyze the kinetics of various processes involving DNA binding proteins.

Chapter I provides an introduction to the different macromolecular process under investigation and the role of kinetics in determining their function. The primary means of investigation of these macromolecular interactions in terms of structure and kinetics is done using nuclear magnetic resonance (NMR) spectroscopic methods. Chapter II provides the theoretical description of how NMR based methods enable us to analyze kinetics involving macromolecules occurring at different time scales. Firstly we develop a TROSY based z-exchange methodology to analyze the kinetics of translocation of a

transcription factor, HoxD9 homeodomain, between two cognate DNA molecules in Chapter III. The facilitated target search process in a very challenging problem, which defines the function of different DNA binding proteins and the kinetic shortcuts taken by these macromolecules in speeding up the target location. In Chapter IV the NMR based methods used to dissect the roles of different translocation processes is extensively analyzed using simulations to verify their validity range. Here we also look into the effects of various microscopic events on the macroscopic rates measured using biophysical approaches. The redox regulatory proteins continually keep the cellular environment reductive and the redox state of DNA binding proteins such as HMGB1 could play an important role in its function. The reduction/oxidation kinetics of HMGB1 along with redox regulatory proteins are extensively analyzed in Chapter V using NMR based methods. This thesis provides an understanding into the physiologically relevant macromolecular interactions in terms of kinetics by the development of novel biophysical methods, simulations and experiments.

Table of Contents

List of Figures

CHAPTER I

INTRODUCTION

Living organisms carry out different biological processes for their survival and sustenance by depending on the interactions of macromolecules like proteins, DNAs, RNAs, lipids, etc. The inter-cellular communications occur in the form of external stimuli, instantly triggering a cascade of signal molecules that find their appropriate binding partners propagating the signal downstream to obtain the desired cellular response. It is important to note that macromolecules not only need to interact with each other but also need to carry out their respective functions very efficiently and effectively. Thus raising a very fascinating question of how do these macromolecules come together quickly in the right spatial position, orientation and conformation when there are lots of non-specific interactions present inside the cells. Often there are many obstacles on their path towards finding their final functional state (1, 2), such as other macromolecules, non-specific binding partners, and post-translational modifications. These hurdles are somehow overcome very efficiently, raising another important question, what are the kinetic shortcuts taken by biomolecules towards performing their functions?

In the case of Eukaryotic systems, there are huge cellular responses to external stimuli like hormones, neurotransmitters, growth and differentiation factors and metabolites. The cellular response to these stimuli often include activation of early genes and sometimes inducing cytosolic protein modifications like phosphorylation, acetylation, glycosylation, reduction/oxidation, degradation, etc. These protein modifications are crucial in generating proper response to the stimuli and it is very important for the cells to maintain appropriate populations of these functional states. For example, in eukaryotes, the ubiquitin molecule finds unwanted proteins in the cell

quickly, and tags them for degradation and recycling by proteasome (3). In order to study which macromolecules interact with each other, the classical approaches used are immunoprecipitation (4) or a pull-down assays (5). Even though these crude methods give us possible binding partners, the assays pick up only strong binding partners and sometimes the observed interactions might not be physiologically relevant. Though neglected often, weak interactions at times might play a vital role in the proper macromolecular function. For example, in the case of protein translocation on DNA, weak non-specific interactions with the DNA do contribute to facilitated target location (6) process. These macromolecular interactions are essential for the survival and proper functioning of an organism, furthermore the speed of these interactions also plays a crucial role in proper cellular response. Hence there is a need for the development of methods that can be used to measure the kinetics of these macromolecular interactions.

The kinetics of macromolecular interactions have been studied for a very long time using different biophysical approaches like fluorescence based experiments (7), single molecule experiments (8-10), gel-shift assays, and Nuclear Magnetic Resonance (NMR) experiments (6, 11-16). Sometimes, structural insights into the biological interactions are needed to understand the macromolecular interactions, which are mainly obtained by structural approaches like atomic force microscopy, cryo-electron microscopy, Small Angle X-ray Scattering (SAXS), NMR spectroscopy and X-ray crystallography. Interestingly NMR based approaches not only reveal microscopic information but also report on macroscopic events. Even though each of these analytical methods gives very useful information, they are not free from limitations either in instrumentation or experimental conditions; thus there is a need of developing new biophysical methods to analyze macromolecular interactions.

For instance, for studying protein-DNA interactions there have been a lot of biophysical approaches like fluorescence based experiments (7), single molecule experiments (8, 9), etc. In the fluorescence approach, often a fluorescent probe is attached to the DNA molecule and the effect of protein binding to the DNA is observed as a function of fluorescence intensity or anisotropy change. The typical working concentration of DNA used in these approaches is in nanomolar to micromolar range. On the other hand the DNA in the cell is confined to a very small compartment, and a simple mathematical calculation gives the nuclear DNA concentration to be in the range of 100-150 mM in terms of nucleotide base pairs (17). So the nuclear DNA concentration is in the range of 3-5 mM by converting the concentrations in terms of 30bp fragments, which are too high DNA concentrations to implement fluorescence based and single molecule approaches. One of the caveats of fluorescence based approaches to understand protein-DNA interactions is that we can only observe kinetics macroscopically and very little or sometimes no information is detected in the microscopic scale/atomic resolution. This limitation has been addressed by careful Förster Resonance Energy Transfer (FRET) based experiments to obtain various kinds of dynamic information on protein-DNA system (18, 19), but the perturbation created by the addition of a florescence donor/acceptor could interfere with the function of the protein. Recent advances in the field of single molecule based approaches (8-10) have given us a great insight into macromolecular interactions, but isolation of macromolecules could mask higher order interactions that are present at high concentrations of macromolecules under intracellular conditions.

Due to the recent advancements of NMR and X-ray crystallographic structural initiatives, there is a vast repository of structures of proteins alone and also with their binding partners in atomic resolution. These structures do give us structural insights into

these interactions but the information on the dynamic interaction process is often lost as they mostly reveal the initial and final state of the complex macromolecular interactions. Even though some aspects of these macromolecular interactions have been extensively studied using different biophysical methods, there are a few aforementioned limitations that create a gap in the knowledge of understanding complex interactions. To overcome some of these limitations we plan to use nuclear magnetic resonance (NMR) based methods whose merits are discussed below.

NMR based methods have been successfully used on various macromolecular systems (6, 12-16, 20) and useful information has been obtained on these macromolecular interactions in terms of structure, dynamics, kinetics and energetics. In order to observe a signal using NMR based methods, the macromolecules have to be isotope labeled to have odd number of protons and neutrons so that they possess a net magnetic moment. In the case of proteins, naturally abundant ^{12}C carbon atoms can be replaced with ^{13}C, and ^{14}N can be replaced with ^{15}N using isotope labeled substrates during the production of proteins so that they are detectable in NMR experiments. Making use of multidimensional NMR experiments, signals can be easily assigned to different atoms on the macromolecule; this feature can be exploited in observing the required interactions. In the case of ligand binding to proteins, NMR based methods have been successfully applied to reveal structural changes, binding affinities, specificity, and sometimes information on the location of ligand binding site on the protein. When applied to understand protein translocation on DNA (6, 12), we can use high concentrations of physiologically relevant unlabeled DNA concentrations (3-5 mM) with isotope labeled proteins to directly observe the protein translocation mechanisms operating inside the cells.

NMR based experiments on macromolecules give us a lot of structural information in atomic resolution, since we observe a signal corresponding each isotopically labeled atom of the macromolecule. These NMR signals arising from individual atoms are very sensitive to the physical environment around the nuclei, thus providing information on small changes surrounding the nuclei. NMR based methods have a minimal effect on the chemical and physical properties of macromolecules since this perturbation is limited only to the nuclei of the atoms, thus making it a very useful and powerful approach in studying biomolecules.

While performing their respective functions efficiently, biological macromolecules sometimes explore different environments in the form of chemical reaction or conformational change. And NMR based spectroscopic methods can be used on these macromolecules to observe the exchange between these states if the nucleus experiences different magnetic environments in each of these states, where this exchange is often referred to as chemical exchange. In the case of chemical equilibrium, the observed nucleus experiences different states when the system exchanges between two or more sites on a different molecule. Like in the case of conformational exchange, the nuclei exchanges between different conformations of the same macromolecule. An example of intermolecular exchange observed using NMR based methods would be the physical translocation of proteins between two different DNA molecules and that of an intramolecular conformation change would be domain motions of a multi-domain protein. NMR based methods can be used to observe exchange between different states as long as these states are magnetically different even if they are chemically indistinguishable. The nuclear spins systems are weekly coupled with chemical lattice, thus permitting the manipulation of nuclear magnetizations without affecting the chemical states. This is a very useful and unique feature as other biophysical approaches

used to study kinetics of exchange of macromolecules between two or more states require the monitoring of states that precedes chemical equilibrium.

We currently possess a wide variety of NMR based approaches (11) like zz-exchange experiments (21), CPMG relaxation dispersion experiments (22, 23), residual dipolar coupling (RDC) experiments (Reviewed in 24, 25), paramagnetic relaxation enhancement (PRE) experiments (26-29), etc which are used to look at different biological processes. Taking advantage of some of these NMR experiments, wide range of processes taking place in the biological systems like side chain motions, domain motions, folding/unfolding of proteins, conformational change, ligand binding, enzyme kinetics, etc have been studied.

Now that we are equipped with powerful biophysical tools to observe complex macromolecular processes, we can bridge the gap in the understanding of macromolecular interactions in terms of kinetics while getting an insight into their function and mechanism. The overall goal in this thesis is to understand the details of physiologically relevant macromolecular interactions in terms of kinetics, which is achieved by developing innovative biophysical methods, models and experiments.

TRANSLOCATION OF DNA BINDING PROTEINS

One of the first cellular responses to external stimuli is the activation of early genes that code for transcription factors. These DNA binding proteins are very efficient in translocating into the nuclei and finding their cognate DNA sites to express appropriate genes while maneuvering among an ocean of non-target DNA sites. There are lots of structural information on DNA binding proteins that reveal the formation of extensive contacts with the phosphate bases and nucleotides on the DNA (e.g., 30-32). These extensive contacts with the DNA are essential for the protein to distinguish between

cognate and non-cognate sites; on the other hand this could slow down the target search process if the protein spends too much time on non-target sites. This issue has boggled scientists and is referred to as the "speed-specificity" paradox wherein the protein has to bind to a specific site on the DNA rapidly while negotiation its path among an ocean of non-specific sites.

In this complex process of target search, if the protein translocation is assumed to be only diffusion limited, a simple theoretical calculation shows the time required to find a specific site on a DNA in a nucleus is in the order of hours. This is contrary to the time scales of target sequence location and activation by most DNA binding proteins *in vivo* (33), which happens to be in the order of a few seconds. Initial experiments on *lac* repressor revealed (34) that the *lac* protein can find its target site 1000 times faster than

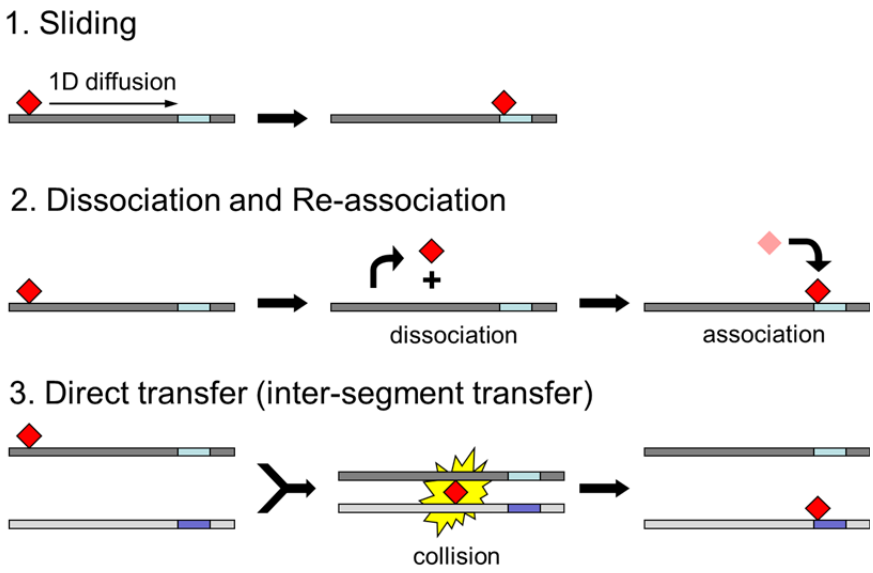

Figure 1.1: The three modes of translocation (i) Sliding (intra-segment transfer) (ii) dissociation followed by re-association and (iii) direct transfer (inter-segment transfer). The second mechanism consists of two steps dissociation and re-association, whereas the direct transfer consists of a single step second order reaction that occurs by collision between the complex and the free DNA.

just a three-dimensional diffusion and collision processes, showing the presence of some translocation mechanisms that facilitate the target search process. This was the first evidence implying that the DNA binding proteins take some kinetic shortcuts to locate their cognate sites. After the initial experiments on the *lac* repressor system, theoreticians have proposed various models through which proteins can perform efficient target search. Some of the widely accepted models for protein translocation mechanisms (34-41) to explain the enhanced rate of target search are, (i) one-dimensional sliding, (ii) hopping (dissociation and re-association) and (iii) direct transfer (also called inter-segment transfer), which are schematically shown in Figure 1.1.

The first mechanism is where the protein sides on DNA, initially binding to non-specific site on the DNA and then slides across the DNA without dissociating to find its target site. This sliding mechanism might be advantageous since the protein scans for target site in only one dimension, thus greatly reducing the search space. This is a first order kinetic process in which the protein slides until it finds its target site. Recently this phenomenon was directly observed by monitoring unbiased movement of transcription factors on DNA using single-molecule experiments (e.g. 9, 42-46) and the one dimensional diffusion rates constant have been determined to be ~300-1000 bp s^{-1}. The second mechanism is the process where the protein dissociates from the DNA and diffuses in the three-dimensional space, which is followed by the process of re-association. Inside the cells the concentrations of DNA is high (17) and under this condition, the initial first order process of dissociation is the rate limiting step which is followed by a very quick second order re-association process. This phenomenon has been extensively studied as bulk measurements using various spectroscopic methods since this translocation mechanism is predominant under low concentrations of DNA and can be easily observed. The third mechanism is a single-step second order process that involves

the collision of free DNA with the protein-DNA complex, where the protein translocates from one DNA to another without going through a free state. This intersegment transfer process was proposed by theoreticians (39, 47-49) where the protein can take kinetic shortcuts to translocate long distances in terms of DNA sequence space. Since this mechanism does not involve the protein to go through a free state, the protein is thought to form a transient bridge between two DNA molecules while not requiring the breakage of all hydrogen bonds with DNA. Until recently this phenomenon was not well characterized due to the limitations of different biophysical approaches, but NMR based experiments have been able to provide the evidence of intersegment transfer mechanism (6, 12). These three major translocation mechanisms are different in terms of kinetics and hence can be distinguished from one another. Since the only process that is dependent on free DNA concentration is the intersegment transfer process, its presence can be monitored by observing a change in the translocation kinetics by varying free DNA concentration. The NMR method used to measure the intersegment transfer kinetics of protein translocating between two DNA molecules is *z*-exchange spectroscopy, which is used for systems under the slow exchanging regime. While providing useful kinetic information this methodology has its limitations as well. Since the signals obtained from this method are only as strong as a Heteronuclear Single Quantum Correlation (HSQC) experiment, this limits the application of this method on larger molecular systems and is restricted to a narrow working temperature range. In order to overcome some of these limitations, there is a need to develop new NMR based methods that would provide the same information with increased sensitivity.

In order to systematically understand the translocation mechanism, we choose the human transcription factor HoxD9 homeodomain as our model system. HoxD9 homeodomain (Figure 1.2) is a helix-turn-helix single domain DNA binding transcription

factor that is highly conserved in humans and fortunately there are lots of structural information available (6, 12). Previous paramagnetic relaxation enhancement (PRE) based NMR measurements on the HoxD9 homeodomain have revealed that the structure of HoxD9 does not change much when bound to specific and non-specific DNA sequences (6, 12, 13), where it makes extensive contacts with the DNA. The translocation of HoxD9 between two different cognate DNA molecules was previously examined using *z*-exchange spectroscopy (12), where the increase in free DNA concentration resulted in an increase in the translocation rate constant showing the existence of intersegment/direct transfer mechanism. On the other hand, the determination of activation free energy was not feasible due to broadening of signals during *z*-exchange spectroscopy at low temperatures. In order to determine the activation free energy required for HoxD9 to translocate from one cognate DNA to another, we need to develop a NMR based tool to carry out *z*-exchange in a broader range of temperatures.

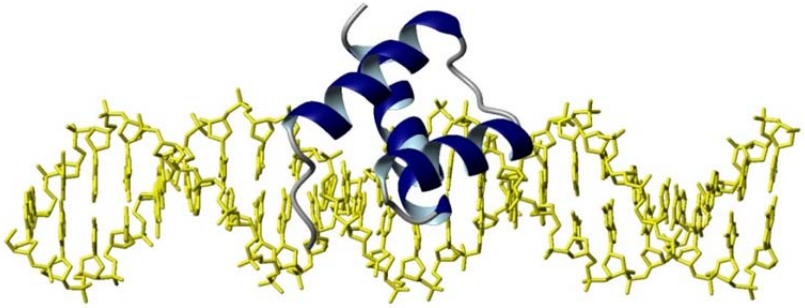

Figure 1.2: Human HoxD9 homeodomain bound to 24bp DNA containing the cognate site TAATGG.

HoxD9 makes nine ion pairs with the phosphate bases and nucleotides via salt bridges and hydrogen bonds which are associated with free energies in the range of 0.4 to 2 kcal/mol each, adding up to ~20 kcal/mol. For the dissociation of HoxD9 followed by re-association to be the predominant mechanism, it is imperative that the activation energy for the translocation of HoxD9 between two DNA molecules containing the

specific DNA binding sites should be ~20kcal/mol. Our strategy to address this issue was to implement the Transverse Relaxation Optimized Spectroscopy (TROSY) principle that increases the sensitivity of the conventional *z*-exchange experiment, and then to monitor the translocation kinetics of HoxD9 homeodomain between cognate DNA molecules at different temperatures. These can then be used in an Eyring plot yielding the activation energy for translocation, thereby confirming the presence or absence of kinetic shortcuts like intersegment transfer.

The process of development of new NMR based method to analyze translocation kinetics in slow exchanging systems will be discussed in detail in Chapter 3 of this thesis. Upon the successful implementation of TROSY principle into z-exchange spectroscopy, the activation free energy associated with the translocation of protein from one target DNA to another can be determined.

VALIDATION OF NMR BASED APPROACH

Previous NMR based approaches developed (6, 12-15) have dissected the different modes of protein translocation on DNA. The approach used to observe the translocation of protein between two DNA molecules was the "mixture approach" (6, 12, 13), where 1:1 ratios of two different unlabeled DNA molecules are mixed with isotope labeled protein. The consequences of the mixture of different DNA molecules results in the translocation of protein from one DNA to another which can be monitored using NMR based methods if the protein experiences slightly different magnetic environments when bound to each DNA. When the translocation is in the slow exchange regime, two sets of signals arising from the protein bound to individual DNA molecules are observed. These slow exchanging systems can be analyzed using different *z*-exchange NMR methods (12, 50), yielding the intermolecular translocation rate constant. Similarly when

the translocation is in the fast exchanging regime, only one signal is observed as a population weighted average of the two states. The apparent relaxation for a two-state fast exchanging system can be analyzed using Lorentzian line-shape analysis and with the use of mixture approach and Reuben and Fiat approximation, the intersegment translocation rates can be determined (6). The Reuben and Fiat approximation (51) used in the NMR based mixture approach (6) to determine the translocation rate is valid only for a two-state exchanging system, the apparent translocation kinetics obtained could be affected by multiple microscopic states accessed by the protein on DNA. On the microscopic scale the protein could be bound to different positions on the DNA molecule playing a vital role in finding its cognate sites. In order to verify the validity range of these NMR methods, there was a need to develop an N-site theoretical model to describe this highly complex translocation process on DNA composed of multiple microscopic states. Additionally it was also necessary to consider the relationship between the apparent rate constants measured using the NMR based methods and the macroscopic rate constants (Figure 1.3) affected by various microscopic states.

The development of a model to describe the protein performing different translocation events between N-sites is challenging because the energy landscapes of a protein-DNA system are supposed to be rugged (52) and the kinetic rate constants should satisfy the detailed balance. Our strategy is to develop a model that is based on McConnell equations and Kramers' theory for an N-state exchanging system and simulate the NMR approaches used to interpret the translocation mechanisms. These simulations would provide us with valuable knowledge on the validity range of our NMR based approach for the kinetic investigation of protein translocation on DNA.

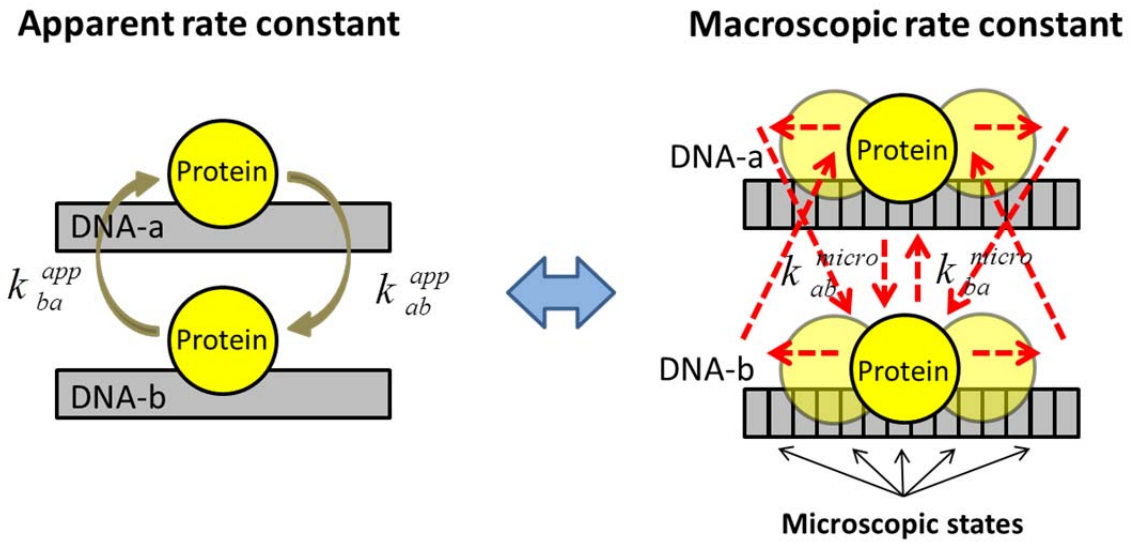

Figure 1.3: The accuracy and validity range of NMR based methods needs to be examined for experimentally observed apparent macroscopic rate constant (left) and the macroscopic rate constant affected by different microscopic states.

Upon successfully developing the model, simulation of various plausible experimental conditions should give us ideas on building new experimental tools to decipher the significance of one-dimensional sliding. The simulations can then be extended to describe systems involving large DNA binding proteins whose sliding rates might be slower and predict the outcomes of the NMR experiments on these systems. The details of these simulations will be discussed in chapter 4 of this thesis.

We have tackled the challenges in developing new NMR based methods to understand the protein translocation process and building a generalized N-site model to describe the protein translocation on DNA so as to verify the accuracy and validity range of NMR experiments. Next we intend to systematically dissect the different states explored by DNA binding proteins inside the cell as they could play an important role in its function. Since the cells spend enormous amount of energy in keeping the intracellular

environment reductive, there is a need to understand the role of various redox states of DNA binding proteins.

REDOX STATE OF DNA BINDING PROTEINS

In most proteins, spatially close cystines can form disulphide bonds which help them maintain structure under oxidative environments. The formations of disulphide bonds could help stabilize the structure of proteins under oxidative conditions, but once formed the available conformational space is restricted. Living cells possess a whole arsenal of redox protein machinery such as Glutathione, Thioredoxin, Nucleoredoxin, etc to keep the intracellular environment reductive (53, 54). The redox state of the proteins inside the cells could play a role in its interactions with other biomolecules and determine its function (54, 55). Thus it is important to study the kinetics of reduction/oxidation of these proteins by cellular redox machineries. When the cells are under oxidative stress, there are high amounts of oxidative species (like superoxide anion, hydrogen peroxide, etc) and low levels of reduced proteins present in the cell that could affect the functions

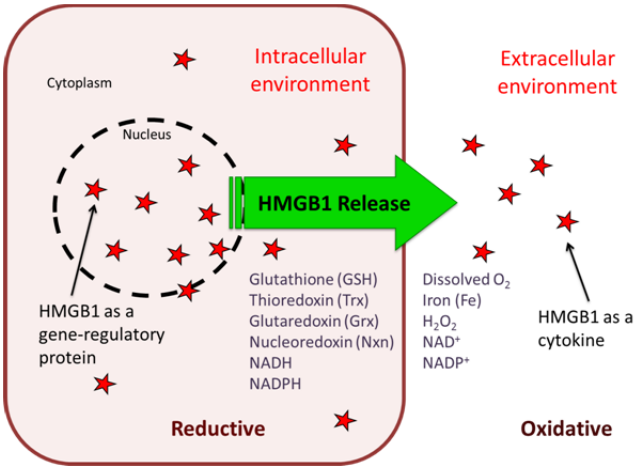

Figure 1.4: Different roles of HMGB1: Inside the cell primary function of HMG1 is DNA binding and regulation of genes. Outside the cell HMGB1 acts a cytokine interacting with RAGE and TLR receptors.

of many important protein machineries. Hence there is a need to understand the redox states of biologically important proteins that could lead to the understanding of the root cause of a complex disease.

A very abundant protein, High Mobility Group Box 1 (HMGB1) is a unique DNA binding protein (15, 56-60) where it bends the DNA upon binding, facilitating other transcription factor in finding their target sites and also acts as a cytokine invoking the necrosis pathway extracellularly (Figure 1.3). HMGB1 is made up of two very similar DNA binding domains (A-box and B-box) and a highly acidic tail (59, 60). The full length HMGB1 protein contains three cystines, where C22 and C44 on A-box form a disulphide bond (Figure 1.4) and C105 on B-box remains reduced. In the extracellular environment HMGB1 is believed to act as damage associated molecular pattern (DAMP) molecule which interacts with receptors, like receptor for advanced glycation end-products (RAGE) and toll-like receptors (TLR), playing an important role in oxidative stress and apoptosis (61). High levels of HMGB1 expression have also been associated with many forms of cancer and inflammatory diseases like asthma (61). The redox state of this protein is poorly characterized even though HMGB1 has been attributed to many diseases, thus creating a gap in knowledge on the function of the protein.

The two distinct functions of HMGB1 could be ascribed to the different redox environments inside and outside the cells. Since these two distinct functionally important states of HMGB1 have not been analyzed, there was a need to address the question and determine the population of individual redox states. This information on the population of individual redox state of the protein can be calculated by determining the redox potential of the protein, and more importantly we get valuable information on the rate of reduction of HMGB1 by redox regulatory systems.

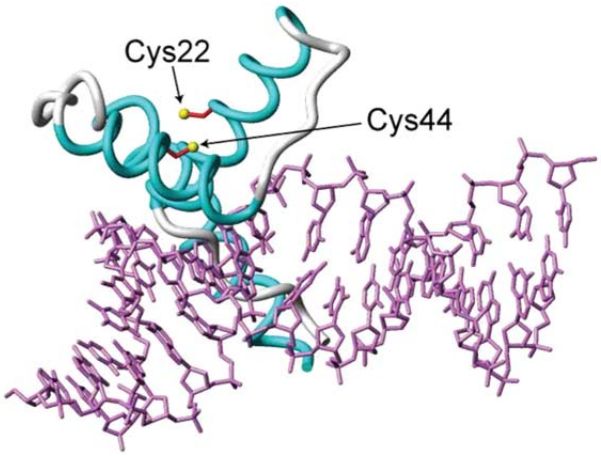

Figure 1.5: HMGB1 A-box bound to DNA. Highlighted are Cys22 and Cys44 that form the disulphide bond

Our strategy to obtain the redox potential of HMGB1 is to monitor redox reactions involving HMGB1 and glutathione, since glutathione system is well understood and is the most abundant redox machinery in the cells (54). In parallel we also intend to understand the reduction of HMGB1 by another redox regulatory system, thioredoxin system comprising of thioredoxin, thioredoxin reductase and NADPH (53).

If the redox potentials of the proteins involved in the redox reactions are comparable, the reduction/oxidation reactions would occur in minute-hour timescale. In order to monitor reactions happening in the minute-hour time scales, NMR based real-time kinetics experiments can be used to gain valuable knowledge on the kinetics of redox reactions. From these real-time experiments we would acquire an insight into the redox state of HMGB1 in cells under oxidative stress once the kinetics of reduction/oxidation is understood, giving us an idea of the efficiency of the redox regulatory machineries to keep proteins reduced inside the cells. The experimental approaches and kinetics of the redox reactions governing the fate of HMGB1's function is described in detail in chapter 5.

The overall goal of this work is to understand the physical basis of the various macromolecular interactions in terms of kinetics, mainly for biologically relevant DNA binding proteins. Since NMR is one of the very powerful tools to achieve this goal, we present the development and application of various NMR based methods to decipher the details of these macromolecular interactions.

CHAPTER II

THEORETICAL BACKGROUND: NMR AS A TOOL FOR KINETIC STUDIES

INTRODUCTION

For understanding the details of reaction mechanisms it is important to comprehensively dissect the kinetic details of the reaction, which have been extensively studied by different analytical techniques. Among the various analytical methods used to study kinetics of chemical reactions, Nuclear Magnetic Resonance (NMR) is one of the very powerful tools that not only gives information on kinetics but also provide structural and dynamic insights with atomic resolution. The signals observed using NMR based methods are very sensitive to the chemical environment around the observed atom which can also be used effectively to gain information on the temperature and pH of the chemical reaction.

NMR based methods can be used to study macromolecular interactions when the system under observation exchanges between two or more distinct magnetic environments, which is referred to as chemical exchange. The contribution to the chemical exchange could be due to intramolecular exchanges like local unfolding/refolding, side-chain motions, etc or intermolecular exchanges like enzyme catalysis, ligand binding, etc. The NMR based methods can be used for observing the changes in the NMR signals due to kinetic processes under dynamic equilibrium, where the exchange is happens to take place at a noticeable rate even though there is no net reaction. This is possible to observe because NMR detects the molecular motions rather than the individual number of molecules in different states.

Often from understanding systems under equilibrium, NMR based methods can be utilized to obtain kinetic information from transient chemical reactions. The system

can be brought to a non-equilibrium point and monitored while it reaches its equilibrium state as a function of time.

A generalized theory on chemical reaction rates, kinetics and its detection using NMR based methods are described below. This description covers the theory on the applicability of NMR based methods on systems under non-equilibrium and equilibrium conditions.

GENERALIZED DESCRIPTION OF REACTION AND KINETICS

Most macromolecular interactions involve two or more molecules to interact with each other, in order to generate appropriate cellular response. Most biological reactions observed in nature are reversible and the different molecules interconvert to exist under dynamic equilibrium. Sometimes the concentrations of one of the macromolecule could influence the fate of another molecule that is connected via a network of chemical reactions. So it is important to quantitatively analyze these dynamic reaction networks revealed by the kinetic rate constants governing their inter-conversion rates to get some insights into their cognate cellular function.

In order to describe reaction networks in classical kinetics, let us consider a system consisting of J molecular species A_j involved in L forward and backwards reactions. The forward and backward reactions are considered separate because the rates could be different under non-equilibrium conditions. The stoichiometry of the reaction can then be formulated by 2L series of linear equations explicitly given by

$$\begin{aligned} \upsilon_{11}\mathrm{A}_1 &+ \upsilon_{21}\mathrm{A}_2 &&+\cdots+ \upsilon_{\mathrm{J}1}\mathrm{A}_\mathrm{J} &&= 0 \\ \upsilon_{12}\mathrm{A}_1 &+ \upsilon_{22}\mathrm{A}_2 &&+\cdots+ \upsilon_{\mathrm{J}2}\mathrm{A}_\mathrm{J} &&= 0 \\ \vdots &\quad \vdots &&\quad \vdots &&\quad \vdots \\ \upsilon_{1,2\mathrm{L}}\mathrm{A}_1 &+ \upsilon_{2,2\mathrm{L}}\mathrm{A}_2 &&+\cdots+ \upsilon_{\mathrm{J},2\mathrm{L}}\mathrm{A}_\mathrm{J} &&= 0 \end{aligned} \qquad (2.1)$$

where forward reactions are depicted by l=1, 2, … , L, backward reactions are given by l=L+1, L+2, ... , 2L and the stoichiometric coefficient involving the j^{th} molecule in the l^{th} reaction is denoted by ν_{jl}. These generalized stoichiometric equations can be rewritten in the matrix form as,

$$\vec{\mathrm{A}}\mathbb{N} = 0 \tag{2.2}$$

where $\vec{\mathrm{A}}$ is a row vector with A_j as its elements and $\mathbb{N}$ is the stoichiometric matrix consisting of the coefficients of A_j in equation 2.1 as its elements.

$$(A_1 \quad A_2 \quad \cdots \quad A_{j-1} \quad A_j)\cdot\begin{pmatrix} \nu_{11} & \nu_{21} & \cdots & \nu_{J-1,1} & \nu_{J1} \\ \nu_{12} & \ddots & & \ddots & \nu_{J2} \\ \vdots & & \ddots & & \vdots \\ \nu_{1,2L-1} & \ddots & & \ddots & \nu_{J-1,2L-1} \\ \nu_{1,2L} & \nu_{2,2L} & \cdots & \nu_{J-1,2L} & \nu_{J,2L} \end{pmatrix} = 0 \tag{2.3}$$

For a reaction l, the reaction number $\xi_l(t)$ gives the extent to which the reaction has progressed, and is the same as the weight per unit volume of the reactants that have reacted in time t to give the desired products. So a vector $\xi(t)$ can be created comprising of 2L reaction number corresponding to 2L reactions. The time dependence of the concentration $[A_j]$ is described by

$$[A](t) = [A](0) + \mathbb{N}\xi(t) \tag{2.4}$$

The rate of change of the concentration is then given by

$$\frac{d}{dt}[A](t) = \mathbb{N}\frac{d}{dt}\xi(t) \tag{2.5}$$

Since the forward and backward reactions are distinguished in this case, we can separate the stoichiometric matrix into sum of forward and backward matrices. The stoichiometric coefficients for the reactants are negative since they are consumed throughout the reaction and those for products are positive as they are continuously produced. The positive coefficients ν_{jl}^{+} can be combined into a matrix $\mathbb{N}^{+}$ and the

negative coefficients corresponding to the magnitude $|\nu_{jl}^{-}|$ can be combined into a matrix $\mathbb{N}^{-}$,

$$\mathbb{N} = \mathbb{N}^{+} + \mathbb{N}^{-} \qquad (2.6)$$

The rate of a reaction $\frac{d}{dt}\xi_l(t)$ depends on the concentrations $[A_j]$ of J species, which can be expressed in the mass-action formalism. This reaction rate is determined mainly by rate constant k_l of the reaction, concentration of reactants $[A_j]$ and their stoichiometric coefficients ν_{jl}^{-}

$$\frac{d}{dt}\xi_l = k_l \prod_{j=1}^{J} \left[A_j\right]^{-\upsilon_{jl}^{-}} \qquad (2.7)$$

Upon the substitution of the above equation 2.7 into equation 2.5 we obtain a series of simultaneous nonlinear differential equations. An analytical solution is very tough to obtain unless a very simplistic system is under investigation, but normally numerical solutions have been calculated to yield the results for these differential equations.

For a first-order reaction, the reaction rate ($\frac{d}{dt}\xi_{jr}$) for production of A_r from A_j is directly proportional to $[A_j]$ and given by,

$$\frac{d}{dt}\xi_{jr} = k_{jr}[A_j] \qquad (2.8)$$

The rate of depletion of A_j is given by

$$\frac{d}{dt}[A_j](t) = -\left(\sum_{r\neq j} k_{jr}\right)[A_j](t) + \sum_{r\neq j} k_{rj}[A_r](t) \qquad (2.9)$$

The above equation 2.9 can be further simplified by defining the kinetic terms into a kinetic matrix. The kinetic matrix is composed of diagonal elements that are the

sum of all reaction rates from *j* to *r* and the off diagonal elements are made up kinetic term describing the *r* to *j* process.

$$\mathbb{K}_{jr} = k_{rj},\ r \neq j \text{ and } \mathbb{K}_{jj} = -\sum_{r \neq j} k_{jr} \tag{2.10}$$

For a first-order reaction process, Eq 2.9 leads to a much generalized form

$$\frac{d}{dt}[A] = \mathbb{K}[A] \tag{2.11}$$

whose solution is readily given by

$$[A](t) = e^{\mathbb{K}t}[A](0) \tag{2.12}$$

EXCHANGE IN SYSTEMS WITHOUT SPIN-SPIN COUPLING

For the spin systems without spin-spin coupling, the chemical exchange effects are described by modified Bloch equations, also known as McConnell equations. The McConnell equations for first-order exchange reactions and single spin systems involved in higher order reactions are described below.

The magnetization (M_j) of a spin system consisting of *j* chemical species follows the Bloch equations (62, 63) in the absence of chemical reactions.

$$\frac{d}{dx}\mathbb{M}_j(t) = \gamma(1-\sigma_j)\mathbb{M}_j(t) \times \mathbb{B}(t) - \mathbb{R}_j\{\mathbb{M}_j(t) - \mathbb{M}_j(0)\} \tag{2.13}$$

where γ is the gyromagnetic ratio of the observed species, σ_j is the chemical shielding constant and the $\mathbb{R}_j$ is the relaxation matrix,

$$\mathbb{R}_j = \begin{pmatrix} R_{2j} & 0 & 0 \\ 0 & R_{2j} & 0 \\ 0 & 0 & R_{1j} \end{pmatrix}. \tag{2.14}$$

When there is a chemical reaction between J species, the magnetization is transferred between the different species giving rise to additional kinetic term in the Bloch equations (62, 63). Thus,

$$\frac{d}{dx}\mathbf{M}_j(t) = \gamma(1-\sigma_j)\mathbf{M}_j(t)\times\mathbf{B}(t) - \mathbf{R}_j\{\mathbf{M}_j(t) - \mathbf{M}_{j0}(t)\} + \sum_r K_{jr}\mathbf{M}_r(t) \qquad (2.15)$$

where $\mathbf{K}$ is the kinetic matrix with elements consisting of K_{jr}, which are connected to the chemical rate constants k_{rj} using Eq. 2.10. The z-magnetization at magnetic equilibrium M_{jo}(t) is proportional to the concentration $[A_j]$(t)

$$\mathbf{M}_{jo}(t) = \mathbf{M}_0\frac{[A_j](t)}{\sum_i [A_i]} \qquad (2.16)$$

During the course of the free precession periods, chemical reaction processes generally occur in the absence of r.f. fields and is observed in one- and two- dimensional Fourier spectroscopy. The evolution of transverse and longitudinal magnetization during these intervals can be described (50) in the rotating frame with frequency ω,

$$\frac{d}{dt}M_j^+ = -\left(R_{2,j} - i\Omega_j\right)M_j^+ + \sum_r K_{jr}M_r^+ \qquad (2.17)$$

$$\frac{d}{dt}M_{jz} = -R_{1,j}\left(M_{jz} - M_{j0}(t)\right) + \sum_r K_{jr}M_{rz} \qquad (2.18)$$

where the transverse magnetization is represented by $M_j^+ = M_{jx} + iM_{jy}$, and the chemical shift frequency is $\Omega_j = -\gamma(1-\sigma_j)B_0$. The magnetization components, M_j^+, M_{j0} and M_{jz} can be combined into a matrix form represented by the magnetization vectors $\mathbf{M}^+$, $\mathbf{M}_z$ and $\mathbf{M}_0$,

$$\frac{d}{dt}\mathbf{M}^+(t) = \mathbf{L}_{tr}\mathbf{M}^+(t) \qquad (2.19)$$

$$\frac{d}{dt}\mathbf{M}_z(t) = \mathbf{L}_{lo}\{\mathbf{M}_z(t) - \mathbf{M}_0(t)\} + \mathbf{K}\cdot\mathbf{M}_0(t) \qquad (2.20)$$

The last term in the previous equation disappears in chemical equilibrium whose details are described below. The term $\mathbf{L}_{tr}$ and $\mathbf{L}_{lo}$ describe the chemical shift, relaxation and kinetics of the processes being observed.

$$\mathbf{L}_{tr} = -\mathbf{R}_{tr} + \mathbf{K} + i\boldsymbol{\Omega} \tag{2.21}$$

$$\mathbf{L}_{lo} = -\mathbf{R}_{lo} + \mathbf{K} \tag{2.22}$$

where the diagonal elements of the chemical shift matrix ($\boldsymbol{\Omega}$) correspond to the chemical shifts of J individual states Ω_j. Similarly the transverse relaxation matrix ($\mathbf{R}_{tr}$) is only made up of diagonal elements which are represented by $\mathbf{R}_{tr,ij} = \delta_{ij} R_{2,j}$. The longitudinal relaxation matrix ($\mathbf{R}_{lo}$) consists of diagonal elements representing longitudinal relaxation of individual states and the off-diagonal elements represent the cross-relaxation between nuclei of different species. In a system consisting of J components and undergoing chemical reactions in the first-order, the evolution of transverse magnetization in the matrix form is given by

$$\frac{d}{dt}\begin{pmatrix} M_1^+ \\ \vdots \\ \vdots \\ M_j^+ \end{pmatrix} = -\left\{ \begin{bmatrix} R_{2,1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & R_{2,j} \end{bmatrix} - i\begin{bmatrix} \Omega_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Omega_j \end{bmatrix} + \begin{bmatrix} \sum_{l=1} k_{1l} & -k_{21} & \cdots & -k_{j1} \\ -k_{12} & \ddots & & \vdots \\ \vdots & & \ddots & -k_{j(j-1)} \\ -k_{1j} & \cdots & -k_{(j-1)j} & \sum_{l=j} k_{jl} \end{bmatrix} \right\} \begin{pmatrix} M_1^+ \\ \vdots \\ \vdots \\ M_j^+ \end{pmatrix} \tag{2.23}$$

Such parallels to the longitudinal magnetization can also be made using the equations 2.20 and 2.22. For a system that is not in equilibrium the time dependence of equilibrium magnetization M_{jo} connected to A_j species is directly proportional to the concentration $[A_j]$. On the other hand for the systems in dynamic equilibrium the

equilibrium magnetization M_0 is zero since there is no net concentration change, thus reducing the equation 2.16 to

$$\frac{d}{dt}\Delta M_{lo} = L_{lo} \cdot \Delta M_{lo} \tag{2.24}$$

where $\Delta M_{lo} = M_{lo} - M_0$ that represents the variation in the Boltzmann distribution of nuclear polarizations in a system under equilibrium.

HIGHER-ORDER EXCHANGE WITHOUT SPIN-SPIN COUPLING

The various chemical reactions under investigation using NMR methods might involve more than one reactants whose spins are uncoupled and could act as tracers under different molecular environments. In the case of only one nuclear spin, let us assume that the pathway involves J molecular events and the reaction rates are given by k_{rj} for conversion of A_r to A_j. A schematic representation of such a scheme is described below

$$A_1 + B \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} A_2$$

$$A_2 + C \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} A_3 + D$$

$$\vdots$$

$$A_{j-1} \underset{k_{j,j-1}}{\overset{k_{j-1,j}}{\rightleftharpoons}} A_j + Z$$

In the perspective of the tracer nucleus, the reaction seems to be similar to first-order reaction with an exception that the reaction rates are expressed as derivatives of

reaction numbers that are dependent on concentrations of various macromolecules involved in the reaction. This can be numerically calculated by solving a series of simultaneous equation as described in equation 2.1.

Kinetic rate constants k_{rj}(t) can be derived by dividing the reaction rates $\frac{d}{dt}\xi_{rj}(t)$ by the concentration of reactants $[A_r](t)$

$$k_{rj}(t) = \frac{\frac{d}{dt}\xi_{rj}(t)}{[A_r](t)} \tag{2.25}$$

We have similar first-order formalism for the kinetic matrix as in equation 2.10, which is now dependent on time. This leads to partial differential equation describing the fate of transverse and longitudinal magnetization (11, 64),

$$\frac{d}{dt}M^{+}(t) = \boldsymbol{L}_{tr}(t)M^{+}(t) \tag{2.26}$$

$$\frac{d}{dt}M_z(t) = \boldsymbol{L}_{lo}(t)\{M_z(t) - M_0(t)\} + K(t)\cdot M_0(t) \tag{2.27}$$

These equations are similar to the equations 2.19 and 2.20, but now the $\boldsymbol{L}_{tr}$ and $\boldsymbol{L}_{lo}$ are dependent on time. But when a system is under equilibrium, the concentrations and the reaction rates do not change with time, and a time-independent kinetic matrix is obtained. The magnetization behavior is very similar to first-order reactions (as in Eq. 2.19 and 2.20) and hence NMR based methods can be used for reactions of higher order as long as only one spin system is involved.

CHEMICAL EXCHANGE FOR A TWO STATE SYSTEM

In order to understand the effects of chemical exchange on NMR spectroscopy let us consider a simple two-state (*A* and *B*) first order exchange process involving one spin system, where the exchange between the two different magnetic environments separated

with resonance frequency Δυ occurs with a rate constant of k_1 for forward and k_2 for backward reactions. In the following two-state reaction under equilibrium, the spin in species A is transferred to species B and vice versa due to chemical reaction. This could also involve other species (C_1 and C_2) that are not under observation.

$$A + C_1 \underset{k_2}{\overset{k_1}{\rightleftarrows}} B + C_2$$

For this system the modification of the Eq. 2.23 can be done to describe the population of individual species in terms of kinetics,

$$\frac{d}{dt}\begin{bmatrix}[A](t)\\ [B](t)\end{bmatrix} = \begin{bmatrix}-k_1[C_1](t) & k_2[C_2](t)\\ k_1[C_1](t) & -k_2[C_2](t)\end{bmatrix}\begin{bmatrix}[A](t)\\ [B](t)\end{bmatrix} \tag{2.28}$$

In these matrices the rate expressions for the species [C_1](t) and [C_2](t) were ignored as the spin of interest in not contained in these species. The extension of Eq. 2.26 and 2.27 can be done for this simple two state system to describe the evolution of both longitudinal and transverse magnetization (63).

$$\begin{aligned}\frac{d}{dt}\Delta M_z(t) &= -(R + K)\Delta M_z(t)\\ \frac{d}{dt}\Delta M^+(t) &= -(R + K - i\Omega)\Delta M^+(t)\end{aligned} \tag{2.29}$$

Here the rate matrix for describing the longitudinal magnetization is given by

$$R + K = \begin{bmatrix}R_{1,A} + k_1 & -k_2\\ -k_1 & R_{1,B} + k_2\end{bmatrix}, \tag{2.30}$$

whose eigenvalues can be calculated to be

$$\lambda_{\pm} = \frac{1}{2}\left((R_{1,A} + R_{1,B} + k_1 + k_2) \pm \sqrt{(R_{1,A} - R_{1,B} + k_1 - k_2)^2 + 4k_1k_2}\right) \tag{2.31}$$

The time course of the magnetization is given by

$$\begin{bmatrix}\Delta M_{z,A}(t)\\ \Delta M_{z,B}(t)\end{bmatrix} = \begin{bmatrix}a_{AA}(t) & a_{AB}(t)\\ a_{BA}(t) & a_{BB}(t)\end{bmatrix}\begin{bmatrix}\Delta M_{z,A}(t)\\ \Delta M_{z,B}(t)\end{bmatrix} \tag{2.32}$$

where components of rate matrix is analytically calculated (50)

$$a_{AA}(t)=\frac{1}{2}\left(\left(1-\frac{R_{1,A}-R_{1,B}+k_1-k_2}{(\lambda_+-\lambda_-)}\right)e^{-\lambda_- t}+\left(1+\frac{R_{1,A}-R_{1,B}+k_1-k_2}{(\lambda_+-\lambda_-)}\right)e^{-\lambda_+ t}\right)$$

$$a_{BB}(t)=\frac{1}{2}\left(\left(1+\frac{R_{1,A}-R_{1,B}+k_1-k_2}{(\lambda_+-\lambda_-)}\right)e^{-\lambda_- t}+\left(1-\frac{R_{1,A}-R_{1,B}+k_1-k_2}{(\lambda_+-\lambda_-)}\right)e^{-\lambda_+ t}\right) \quad (2.33)$$

$$a_{AB}(t)=\frac{k_2}{(\lambda_+-\lambda_-)}\left(e^{-\lambda_- t}-e^{-\lambda_+ t}\right)$$

$$a_{BA}(t)=\frac{k_1}{(\lambda_+-\lambda_-)}\left(e^{-\lambda_- t}-e^{-\lambda_+ t}\right)$$

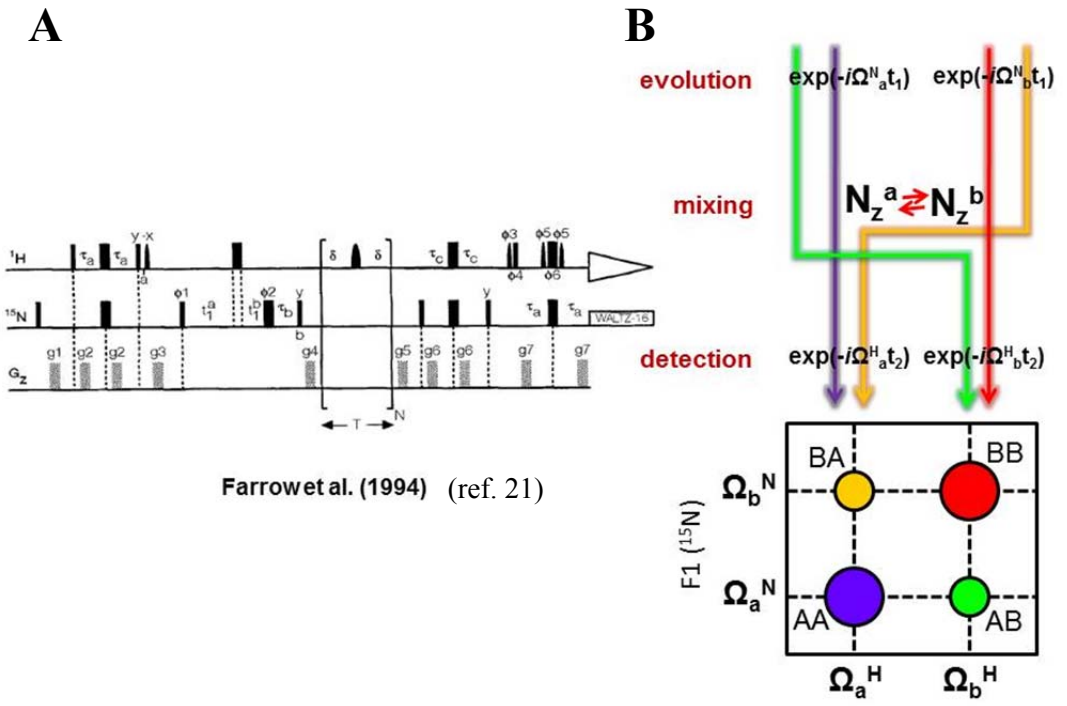

Figure 2.1: NMR technique to analyze slow exchanging systems (A) Pulse sequence of conventional z-exchange experiment to study systems under slow exchange (21) (B) Schematic diagram of the fate of magnetization during a z-exchange experiment

When the initial perturbation is nonselective, $\Delta M_{z,A} \propto p_A$ and $\Delta M_{z,B} \propto p_B$, where p_A and p_B are the population of individual states A and B. When $R_{1,A}=R_{1,B}=R_1$, the longitudinal magnetization time dependence is given by following equations

$$
\begin{aligned}
a_{AA}(t) &= \left(p_a + p_b e^{-2k_{ex}t}\right)e^{-R_1 t} \\
a_{BB}(t) &= \left(p_b + p_a e^{-2k_{ex}t}\right)e^{-R_1 t} \\
a_{AB}(t) &= p_a\left(1 - e^{-2k_{ex}t}\right)e^{-R_1 t} \\
a_{BA}(t) &= p_b\left(1 - e^{-2k_{ex}t}\right)e^{-R_1 t}
\end{aligned}
\tag{2.34}
$$

So when the exchange rate is slow ($k << 2\pi \Delta\upsilon$) we observe two distinct signals arising from individual states (A and B). The z-magnetization for a two-site exchanging system can be monitored using z-exchange experiment (21) whose pulse sequence is shown in figure 2.1(A). In this experiment the magnetization is initially transferred from the amide proton to the nitrogen during the t_1 period and then converted to N_z magnetization. Due to the dynamic exchange between the states A and B the magnetization is transferred between the two states during the 'T' mixing period. During this mixing period, if the nuclei experiences different magnetic environment due to exchange between the two or more states it gives rise to exchange cross-peaks. The magnetization is transferred back to proton for detection. Four peaks emerge from this experiment as shown in figure 2.2(B), two auto peaks AA (Ω_a^N, Ω_a^H) and BB (Ω_b^N, Ω_b^H) and two exchange cross-peaks AB (Ω_a^N, Ω_b^H) and BA (Ω_b^N, Ω_a^H). The intensities of these signals evolve as described in Eq. 2.33.

The rate matrix for describing the transverse magnetization is given by

$$
R + K - i\Omega = \begin{bmatrix} R_{2,A} + k_1 - i\Omega_A & -k_2 \\ -k_1 & R_{2,B} + k_2 - i\Omega_B \end{bmatrix}, \tag{2.34}
$$

whose eigenvalues can be calculated to be

$$
\lambda_{\pm} = \frac{1}{2}\left(\begin{array}{l} (R_{2,A} + R_{2,B} + k_1 + k_2 - i\Omega_A - i\Omega_B) \\ \qquad \pm \sqrt{(R_{2,A} - R_{2,B} + k_1 - k_2 + i\Omega_A - i\Omega_B)^2 + 4k_1k_2} \end{array} \right) \tag{2.35}
$$

The time course of the magnetization is given by

$$
\begin{bmatrix} M_A^+(t) \\ M_B^+(t) \end{bmatrix} = \begin{bmatrix} a_{AA}(t) & a_{AB}(t) \\ a_{BA}(t) & a_{BB}(t) \end{bmatrix} \begin{bmatrix} M_A^+(t) \\ M_B^+(t) \end{bmatrix} \tag{2.36}
$$

where (50)

$$
\begin{aligned}
a_{AA}(t) &= \frac{1}{2}\left(\begin{array}{l}\left(1-\dfrac{R_{2,A}-R_{2,B}+k_1-k_2-i\Omega_A+i\Omega_B}{(\lambda_+-\lambda_-)}\right)e^{-\lambda_- t} \\ \qquad +\left(1+\dfrac{R_{2,A}-R_{2,B}+k_1-k_2-i\Omega_A+i\Omega_B}{(\lambda_+-\lambda_-)}\right)e^{-\lambda_+ t}\end{array}\right) \\
a_{BB}(t) &= \frac{1}{2}\left(\begin{array}{l}\left(1+\dfrac{R_{2,A}-R_{2,B}+k_1-k_2-i\Omega_A+i\Omega_B}{(\lambda_+-\lambda_-)}\right)e^{-\lambda_- t} \\ \qquad +\left(1-\dfrac{R_{2,A}-R_{2,B}+k_1-k_2-i\Omega_A+i\Omega_B}{(\lambda_+-\lambda_-)}\right)e^{-\lambda_+ t}\end{array}\right) \\
a_{AB}(t) &= \frac{k_2}{(\lambda_+-\lambda_-)}\left(e^{-\lambda_- t}-e^{-\lambda_+ t}\right) \\
a_{BA}(t) &= \frac{k_1}{(\lambda_+-\lambda_-)}\left(e^{-\lambda_- t}-e^{-\lambda_+ t}\right)
\end{aligned}
\qquad (2.37)
$$

In the NMR experiment we observe the Fourier transform of $M_A^+(t)+M_B^+(t)$. The real part of the eigenvalue corresponds to the apparent transverse relaxation rate and the imaginary part corresponds to the apparent chemical shift. For simplicity, we assume $R_{2,A}=R_{2,B}=R_2$, then $\lambda_\pm = -(R_{2,app}+i\Omega_{app})$

$$
\begin{aligned}
R_{2,app} &= R_2+\frac{k_{ex}}{2}\pm\frac{1}{\sqrt{8}}\left(\sqrt{k_{ex}^2-\Delta\omega^2+\sqrt{\left(k_{ex}^2-\Delta\omega^2\right)^2-16p_Ap_B\Delta\omega^2k_{ex}^2}}\right) \\
\Omega_{app} &= \frac{\Omega_A+\Omega_B}{2}\pm\frac{1}{\sqrt{8}}\left(\sqrt{\Delta\omega^2-k_{ex}^2+\sqrt{\left(k_{ex}^2-\Delta\omega^2\right)^2-16p_Ap_B\Delta\omega^2k_{ex}^2}}\right)
\end{aligned}
\qquad (2.38)
$$

where $\Delta\omega = \Omega_A-\Omega_B$.

When the exchange rate is slow ($k_{ex} << 2\pi\,\Delta\upsilon$) the off diagonal terms in Eq 2.34 can be neglected and the magnetic components for each of the states evolve independent of each other. These magnetizations are given by (11, 64)

$$M_A^+(t) = M_A^+(0)\exp\left(-(R_{2,A} + k_1 - i\Omega_A)t\right)$$
$$M_B^+(t) = M_B^+(0)\exp\left(-(R_{2,B} + k_2 - i\Omega_B)t\right) \quad (2.39)$$

The observed magnetization $M_A^+(t) + M_B^+(t)$ thus gives rise to two peaks at positions Ω_A and Ω_B but the linewidths are no longer determined by just $R_{2,A}$ and $R_{2,B}$, but rather the chemical exchange broadens the peaks where the linewidths (using Eq 2.29) for states A and B are given by $R_{2,A}+k_1$ and $R_{2,B}+k_2$ respectively. This also suggests that the two peaks for states A and B broaden differently.

When the exchange rate is fast ($k_{ex} >> 2\pi\ \Delta\upsilon$) the observed magnetization $M_A^+(t) + M_B^+(t)$ is positioned at the population average precession frequency. The evolution of the magnetization (51) for this system is given by

$$M^+(t) = M^+(0)\exp\left(-(p_A R_{2,A} + p_B R_{2,B} + \frac{4\pi^2 p_A p_B \Delta\upsilon^2}{k_{ex}} - i(p_A\Omega_A + p_B\Omega_B))\,t\right) \quad (2.40)$$

In the case of a fast exchanging system we have one peak at the average chemical shift, which is due to the rapid inter-conversion of molecules between states A and B and the magnetization arising from all the individual molecules are not dephased significantly.

REAL-TIME KINETICS

NMR methods have also been successfully utilized in the analysis of the kinetics of chemical reactions happening in real time. Sometimes the complex processes like protein folding need to be mechanistically understood with microscopic precision as subtle changes might not be visible in other spectroscopic methods. Initially there had been efforts to observe the low populated intermediates by novel hydrogen-deuterium NMR experiments. This gave very useful information on the solvent accessible amide hydrogens of the proteins by successively recording NMR spectra. The time required to

measure 1D NMR spectra is very small and continuous rapid acquisition of 1D spectra could reveal the details of the intermediates formed. Even though we gain so much information on the kinetics of the system under observation, the occurrence of overlapping signals could complicate the analysis. This complexity of overlapping NMR signals in 1D measurements can be solved by heteronuclear labeling of proteins like ^{13}C and ^{15}N in order to add additional dimension and resolution, but acquisition of signals in additional dimension increases the recording time for each NMR experiment. This problem has been addressed over the recent years by the development of fast data acquisition methods to rapidly record two-dimensional heteronuclear correlation spectra.

When the chemical reactions measured using NMR is very slow the concentration of macromolecules can be assumed to be in steady-state where the concentrations are constant during the acquisition of NMR experimental data. This reduces the Eq. 2.26 and 2.27 in the same form as Eq. 2.19 and 2.20 by the formation of time-independent terms. So in a NMR based experiment on a system where the inter-conversion between the different states is very slow, the magnetization is transferred to each individual state populated during the period of the experiment. These magnetizations lead to observing signals arising from individual states assessed by the system during the NMR experiment. For a bimolecular reaction where two components react to produce the desired product, rapid acquisitions of NMR based methods like HSQC and TROSY can reveal the signals corresponding to the reactant and the product populations just like any other biophysical method used to monitor individual populations of reactants and products. A series of rapid 2D NMR experiments can be performed sequentially to not only obtain the kinetics of conversion of reactants into products, but also gain insights into the structural changes accompanying the reaction.

We now possess a wide variety of NMR tools to analyze biological processes happening in the microsecond (μs) to hour timescales along with acquiring structural information. Taking advantage of these NMR experiments, we can study a wide range of processes happening in the biological systems like side chain motions, domain motions, folding/unfolding of proteins, conformational change, ligand binding, enzyme kinetics, etc. Now that we have the biophysical tools to observe complex macromolecular processes, we here tackle challenging biological problems in nature that can give us insight into their functions and mechanisms.

*CHAPTER III

TRANSLOCATION OF HOMEODOMAIN BETWEEN TARGET SITES ON DIFFERENT DNA MOLECULES

INTRODUCTION

The fast and efficient target gene location is constantly being carried out by DNA binding proteins in an ocean of obstacles and non-specific sites, for the proper functioning of living organisms. It is known that the DNA binding proteins efficiently translocate and locate their cognate sites in a matter of seconds (65). The presence of some kinetic shortcuts has been proposed by theoreticians in the target location process, to account for such speeds of translocation. Complementing other biophysical approaches to study protein- DNA interactions, nuclear magnetic resonance (NMR) methods have been used to study the different modes of translocation taken by human transcription factor HoxD9 homeodomain (6, 12) during target search. The translocation of HoxD9 homeodomain was reported (13) to be in the slow exchanging regime when the protein translocates between two DNA molecules containing the target site. Even though the presence of direct transfer of protein from one complex with DNA to another was observed using conventional NMR exchange spectroscopy, the height of the energy barrier that needs to be crossed during protein translocation between two cognate DNA molecules was still unknown. Conventional NMR methods limit its use in the measurement of translocation kinetics under broad ranges of temperature, which is required for obtaining the activation energy for this process. Hence there is a need to improve upon the conventional NMR approaches to study the kinetic shortcuts taken my DNA binding proteins during target search process.

* This chapter is adapted from the paper published in Journal of American Chemical Society, (ref. 20) and the license of the use of content from the journal article is attached at the end of the thesis

NMR exchange spectroscopy (EXSY) is a powerful tool for studying dynamic processes in the slow exchange regime. The methodology permits the determination of rate constants for both forward and backward reactions at equilibrium. Two-dimensional ^{1}H-^{1}H EXSY experiments, first proposed by Jeener *et al.* in 1979 (50), have been extensively used for kinetic investigations of various chemical exchange processes involving small organic compounds (see Ref. (66) for a review). For macromolecules, ^{1}H-^{1}H EXSY experiments are generally less suited since exchange and NOE peaks are present and may be difficult to separate owing to extensive chemical shift overlap (67, 68). Exchange experiments involving heteronuclear longitudinal magnetization of product operator terms such as $2S_zI_z$ and S_z that only detect exchange processes were therefore proposed (21, 69, 70). *z*-exchange experiments based on heteronuclear correlation spectroscopy have been used for quantitative kinetic investigations of conformational exchange (69, 71-74), protein/RNA folding (21, 70, 75), and protein-ligand interactions (12, 58, 76).

In principle, activation energies can be obtained from the temperature dependence of the measured rate constants. For biological macromolecules, the available temperature range is limited owing to sample instability at high temperature and poorer quality spectra (i.e. extensive line-broadening) at low temperature owing to longer rotational correlation times. In this chapter, we demonstrate the utility of a ^{1}H-^{15}N TROSY-based *z*-exchange experiment for quantitative determination of rate constants, which is suitable for both low temperature measurements and large molecular weight systems. The extension of a ^{1}H-^{15}N TROSY-based pulse scheme to a *z*-exchange experiment is non-trivial since the simplistic incorporation of a *z*-mixing period following the t_1 evolution period results in undesired buildup of spurious peaks owing to different relaxation rates for the N_z and $2N_zH_z$ terms. We have solved this problem by incorporating a scheme to

convert the N_z term into $2N_zH_z$ and the $2N_zH_z$ term into $-N_z$ in the middle of the z-mixing period. This scheme has the additional benefit of simplifying the behavior of the magnetizations of these two terms in an exchanging system, allowing one to determine rate constants from the dependence of the auto- and exchange-peak intensities as a function of mixing time in the same way as for the conventional non-TROSY $^{15}N_z$-exchange experiment. Because of line-narrowing arising from the use of TROSY-principle (77), the quality of data obtained with the new pulse sequence at a low temperature is far superior to that for the conventional non-TROSY z-exchange experiment. The pulse sequence permits kinetic analysis over a wider range of temperature and is therefore suitable for the determination of activation energies for exchange processes involving biological macromolecules.

The utility of the method is demonstrated by determining the activation energy for the translocation of the HoxD9 homeodomain from its specific target site on one DNA molecule to another. This exchange reaction involves direct transfer of the protein between DNA molecules without going through the intermediary of free protein and plays an important role in the target search process whereby a transcription factor locates its specific DNA target site (6, 12, 13).

MATERIAL AND METHODS

NMR sample

^{2}H-/^{15}N-labeled HoxD9 homeodomain and double-stranded DNA duplexes (24 base pairs) were prepared as described previously (12, 77). The base sequence of one strand of DNA duplex *a* is 5'd-CACCTCT**C**TAATGGCTCACACCTG-3' (with the homeodomain binding site underlined). The equivalent strand for DNA duplex *b* is identical except that the C•G base pair indicated in bold is replaced by an A•T base pair. The affinities of

HoxD9 for DNA duplexes *a* and *b* DNA are virtually identical.[11] The NMR sample contained 0.7 mM ^{2}H-/^{15}N-labeled protein, 0.4 mM DNA duplex *a* and 0.6 mM DNA duplex *b* in a buffer comprising 10 mM sodium phosphate (pH 6.5), 40 mM NaCl, 0.4 mM NaF (as an anti-bacterial agent) and 93% H_2O/7% D_2O. Under these conditions, all the protein is bound to DNA (either duplex *a* or *b*) and the molar ratio of the two complexes *a* and *b* is 2:3. Although a Tris•HCl buffer was used for the previous investigation (12), we used phosphate buffer for the present study since the pH of Tris•HCl is highly dependent on temperature.

NMR spectroscopy

All NMR data were recorded on the Varian NMR system operated at a ^{1}H-frequency of 800 MHz. The TROSY-based *z*-exchange ^{1}H-^{15}N correlation experiments (Figure 1A) were carried out at 8, 15, 20, 30 and 35 ˚C. Data with eight different mixing times between 0.02 s and 0.65 s were acquired in an interleaved manner. Sixteen scans were accumulated per FID, and the maximum values of t_1 and t_2 were 67 ms and 54 ms, respectively, yielding a measurement time of about 23 hours. All other details of the NMR experiment are described in the caption to Figure 3.1. For comparison, $^{15}N_z$-exchange ^{1}H-^{15}N correlation spectra (21) were collected at 8 and 20 ˚C using the same number of scans, data points and spectral widths as those used for the TROSY-based *z*-exchange experiment. NMR data were processed using NMRPipe (78) and the spectra were analyzed using NMRView (79). The rate constants were obtained by best-fitting the intensities of auto- and exchange peaks as a function of mixing time by numerically integrating the McConnell equations (cf. Eq. 9) and optimizing the unknown parameters (rate constants, spin-latice relaxation rate and scale factors) using the program FACSIMILE (80) as described previously (12, 67, 68).

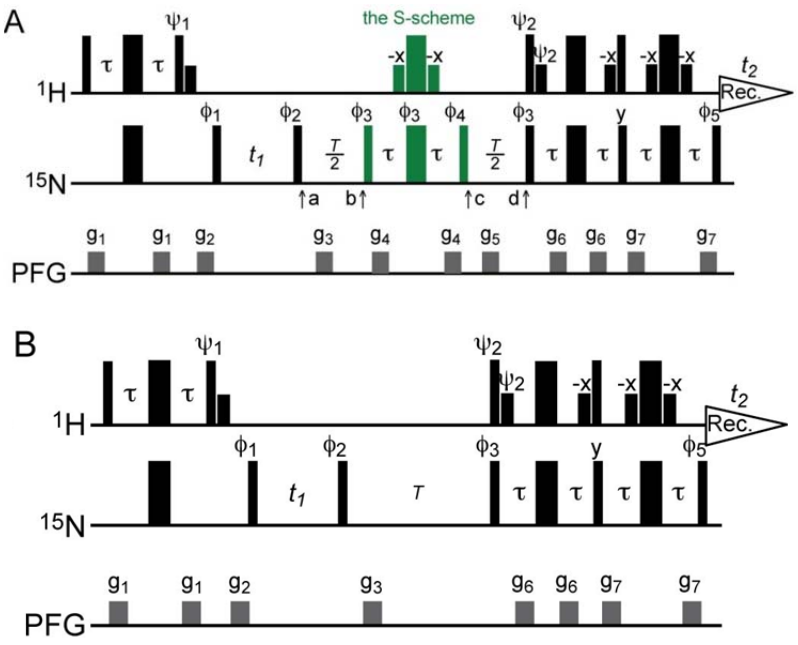

Figure 3.1. Pulse sequences for TROSY-based *z*-exchange spectroscopy. (A) Pulse sequence for the TROSY-based *z*-exchange 2D ^{1}H-^{15}N-correlation experiment incorporating the S-scheme to suppress the buildup of spurious semi-TROSY peaks during the mixing period. Thin and bold bars represent 90˚ and 180˚ pulses, respectively. Unless indicated otherwise, pulse phases are x. ^{1}H pulses represented by short bold bars are soft rectangular 90˚ pulses selective to water (1.2 ms). The delay τ is set to 2.7 ms. The S-scheme (colored in green) is applied in the middle of the *z*-mixing period *T* to convert N_z into $2H_zN_z$ and $2N_zH_z$ into $-N_z$ terms. For Varian spectrometers, the phase cycles are as follows: ϕ_1 = {x, -x, y, -y}; ψ_1 =-y; ϕ_2 = {2(x, -x, y, -y),2(-x, x, -y, y)}; ϕ_3 = {-x, x, -y, y}; ϕ_4 = {y, -y, -x, x}; ϕ_5 = {4x, 4(-x)}; ψ_2 = {4(-y), 4y}; rec. = {x, -x, y, -y, x, -x, -y, y, -x, x, -y, y, -x, x, y, -y}. For Bruker spectrometers, y and –y should be swapped for the ψ_1 and receiver phases. The difference in phases required for the two spectrometer systems is due to the fact that the manner used to shift phase for positive and negative gyro-magnetic nuclei is instrument dependent, as noted previously in the literature (79-81). Quadrature detection in the t_1 domain was achieved using States-TPPI, incrementing the phase ϕ_1. Amplitudes and lengths of pulse field-gradients were as follows: g_1, 8 G/cm, 1.0 ms; g_2, -19 G/cm, 1.7 ms; g_3, -19 G/cm, 1.0 ms; g_4, 8 G/cm, 0.6 ms; g_5, 12 G/cm, 0.7 ms; g_6, 12 G/cm, 0.6 ms; g_7, 15 G/cm, 0.7 ms. The rate constants for exchange were determined from the intensities of the exchange- and auto-peaks recorded in a series of 2D spectra with different values of *T*. (B) Simplistic version of a TROSY-based z-exchange experiment without incorporation of the S-scheme. Phase ψ_2 is {4(y), 4(-y)} for the TROSY-selection; other phases are as in (A). The pulse sequence of panel B was used for Figure 2C but in practice is not suitable for quantitative applications (see text). For each pulse sequence, the consequences of the phase cycles employed were analyzed with the program POMA (78) for both ^1H_z- and $^{15}N_z$-derived magnetizations.

THEORETICAL CONSIDERATIONS

First, we consider the behavior of the components of *z*-magnetization during the mixing period for a non-exchanging ^{15}N-^{1}H spin system. At point a in the pulse scheme depicted in Figure 3.1A, two product operator terms $2N_zH_z$ and N_z are present as a result of *J*-evolution during the t_1 evolution period. Assuming that cross-relaxation between H_z and N_z terms is negligible, the behavior of the $2N_zH_z$ and N_z terms during the period between points a and b is given by: (85, 86)

$$\frac{d}{dt}\begin{pmatrix}\langle 2N_zH_z\rangle \\ \langle N_z\rangle\end{pmatrix} = -\begin{pmatrix}R_{zz} & \eta_z \\ \eta_z & R_z\end{pmatrix}\begin{pmatrix}\langle 2N_zH_z\rangle \\ \langle N_z\rangle\end{pmatrix} \qquad (3.1),$$

where R_{zz} and R_z are relaxation rates for the $2N_zH_z$ and N_z terms, respectively, and η_z is the rate for cross-correlation between these terms. R_{zz} is significantly larger than R_z owing to ^{1}H-^{1}H dipolar interactions. The magnetization of each term in Eq. 3.1 is the average for the phase cycle alternating signs of *z*-magnetization, and $\langle N_z\rangle(\infty) = 0$ instead of the Boltzmann magnetization (87). An analytical solution for Eq. 3.1 can readily be obtained using standard procedures (such as that given in Ref. (88)):

$$\begin{pmatrix}\langle 2N_zH_z\rangle(t) \\ \langle N_z\rangle(t)\end{pmatrix} = \mathbf{P}(t)\begin{pmatrix}\langle 2N_zH_z\rangle(0) \\ \langle N_z\rangle(0)\end{pmatrix} \qquad (3.2),$$

where the elements of $\mathbf{P}(t)$ are as follows:

$$p_{11} = \kappa\exp(-\lambda_- t) + \mu\exp(-\lambda_+ t) \qquad (3.3)$$

$$p_{22} = \mu\exp(-\lambda_- t) + \kappa\exp(-\lambda_+ t) \qquad (3.4)$$

$$p_{12} = p_{21} = \frac{-\eta_z}{\lambda_+ - \lambda_-}\{\exp(-\lambda_- t) - \exp(-\lambda_+ t)\} \qquad (3.5)$$

In Eqs. 3.3-3.5, the rates λ_- and λ_+ are given by $\lambda_{\pm} = 0.5\left\{R_{zz} + R_z \pm \sqrt{(R_{zz} - R_z)^2 + 4\eta_z^2}\right\}$, and the coefficients κ and μ are given by $\kappa = 0.5\{1 - (R_{zz} - R_z)/(\lambda_+ - \lambda_-)\}$ and $\mu = 0.5\{1 + (R_{zz} - R_z)/(\lambda_+ - \lambda_-)\}$, respectively.

The simplest way to incorporate a *z*-mixing period into the TROSY-based ^{1}H-^{15}N correlation experiment is shown Figure 1B but is problematic owing to the buildup of spurious semi-TROSY cross-peaks at (^{15}N, ^{1}H) = (Ω_N+π|J_{NH}|, Ω_H+π|J_{NH}|) arising from imbalance in the $2N_zH_z$ and N_z terms upon increasing the mixing time *T*. Considering the phase cycle for TROSY-selection, the intensities of the TROSY cross-peak at (Ω_N-π|J_{NH}|, Ω_H+π|J_{NH}|) and the spurious semi-TROSY cross-peak at (Ω_N+π|J_{NH}|, Ω_H+π|J_{NH}|) can be calculated as function of mixing time *T* using Eqs. 3.2-3.5 (Figure 3.2A). The simulation indicates that the two cross-peaks are of opposite sign and the signal at (Ω_N+π|J_{NH}|, Ω_H+π|J_{NH}|) can only be suppressed when $T = 0$ or $R_{zz} = R_z$ (which is not possible even for deuterated proteins). The buildup of the negative cross-peaks at (Ω_N+π|J_{NH}|, Ω_H+π|J_{NH}|) is clearly seen in the experimental spectra (Figure 3.2C) recorded using the pulse scheme of Figure 3.1B.

To solve this problem, we introduce a scheme, hereafter referred to as the S-scheme, that converts $2N_zH_z$ into $-N_z$ magnetization and N_z into $2N_zH_z$ magnetization in the middle of the *z*-mixing period and accordingly alter the phase cycle for the TROSY-selection (Figure 3.1A). In this case, the magnetization at the end of the *z*-mixing period (point d in Figure 3.1A) is given by:

$$\begin{pmatrix} \langle 2N_zH_z \rangle(T) \\ \langle N_z \rangle(T) \end{pmatrix} = f\mathbf{P}(\frac{T}{2})\mathbf{SP}(\frac{T}{2})\begin{pmatrix} \langle 2N_zH_z \rangle(0) \\ \langle N_z \rangle(0) \end{pmatrix} \qquad (3.6),$$

$$\mathbf{S}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.7}$$

where the coefficient *f* represents a scaling factor to account for relaxation during the S-scheme, and using Eqs. 3.3-3.5:

$$\mathbf{P}(\frac{T}{2})\mathbf{SP}(\frac{T}{2})=\begin{pmatrix} 0 & \exp\left(-\frac{R_{zz}+R_z}{2}T\right) \\ -\exp\left(-\frac{R_{zz}+R_z}{2}T\right) & 0 \end{pmatrix} \tag{3.8}$$

Note that the two terms decay in a single-exponential manner with identical relaxation rate $\overline{R}=(R_{zz}+R_z)/2$. Since imbalance between the N_z and $2N_zH_z$ terms does not occur in this case, the semi-TROSY component does not buildup even at long *z*-mixing times *T* (Figure 3.2B). In addition, the process is independent of the cross-correlation rate η_z.

Next, we consider for the S-scheme the effect of slow exchange between two states *a* and *b* with rate constants k_{ab} and k_{ba} for the $a\rightarrow b$ and $b\rightarrow a$ transitions, respectively. Since the signal decays in a single exponential manner (cf. Eq. 3.8), one would expect that the mixing time-dependence of the auto- and exchange-peaks can be described by the McConnell equations (63) for longitudinal magnetization:

$$\frac{d}{dt}\begin{pmatrix} M^a \\ M^b \end{pmatrix}=-\begin{pmatrix} \overline{R}^a+k_{ab} & -k_{ba} \\ -k_{ab} & \overline{R}^b+k_{ba} \end{pmatrix}\begin{pmatrix} M^a \\ M^b \end{pmatrix} \tag{3.9}$$

where *M* represents the signal intensities of the TROSY-components. Numerical calculations indicate that this is indeed correct as shown below. Strictly speaking, the overall behavior of the *z*-magnetization terms for the two-site exchange system is given by:

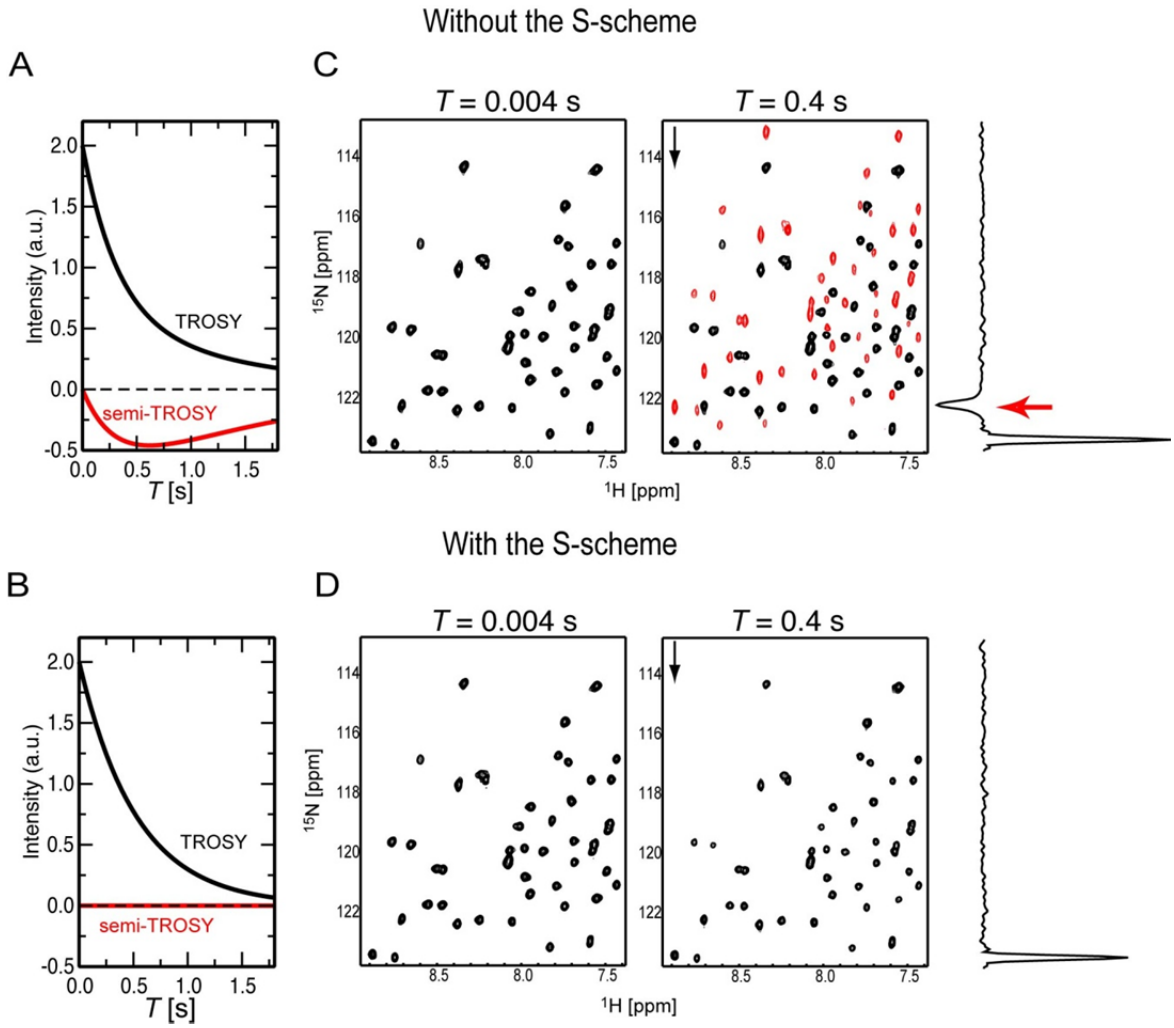

Figure 3.2. Suppression of spurious semi-TROSY peak buildup by the S-scheme placed in the middle of the *z*-mixing period. Simulated time courses (A) without and (B) with the S-scheme. Eqs. 3.2-3.5 are used for the former and Eqs. 3.6-3.8 for the latter and correspond to the pulse schemes shown in Figures 3.1B and A, respectively. Parameters used for these simulations were R_{zz}=4.8 s^{-1}, R_z=0.8 s^{-1} and η_z=0.5 s^{-1}. The phase cycles for TROSY-selection for each experiment were taken into consideration in the simulations. Experimental TROSY-based *z*-exchange ^1H-^{15}N correlation spectra measured at 20 °C on $^2H/^{15}N$-labeled HoxD9 homeo-domain complexed to the 24 bp DNA duplex *a* recorded (C) without and (D) with the S-scheme using the pulse sequences shown in Figures 3.1B and 3.1A, respectively. Positive and negative contours are displayed in black and red, respectively. Slices along the ^{15}N-dimension of the spectrum recorded with a mixing time T = 0.4 s at the positions indicated with black arrows are also shown. The appearance of spurious semi-TROSY peaks in the spectra recorded without the S-scheme is due to imbalance between the $2H_zN_z$ and N_z terms arising from their different relaxation rates.

$$\begin{pmatrix} \langle 2H_zN_z\rangle^a(T) \\ \langle N_z\rangle^a(T) \\ \langle 2H_zN_z\rangle^b(T) \\ \langle N_z\rangle^b(T) \end{pmatrix} = f\exp(-\mathbf{Q}\frac{T}{2})\mathbf{S}'\exp(-\mathbf{Q}\frac{T}{2})\begin{pmatrix} \langle 2H_zN_z\rangle^a(0) \\ \langle N_z\rangle^a(0) \\ \langle 2H_zN_z\rangle^b(0) \\ \langle N_z\rangle^b(0) \end{pmatrix} \quad (3.10),$$

$$\mathbf{Q} = \begin{pmatrix} R_{zz}^a + k_{ab} & \eta_z^a & -k_{ba} & 0 \\ \eta_z^a & R_z^a + k_{ab} & 0 & -k_{ba} \\ -k_{ab} & 0 & R_{zz}^b + k_{ba} & \eta_z^b \\ 0 & -k_{ab} & \eta_z^b & R_z^b + k_{ba} \end{pmatrix} \quad (3.11),$$

$$\mathbf{S}' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \quad (3.12).$$

Figure 3.3 shows the time-courses for the auto- and exchange-cross peaks simulated using either Eq. 3.9 (dashed lines) or Eqs. 3.10-3.12 (solid lines). Although Eq. 3.9 is much simpler, the results are identical and independent of the cross-correlation rates. Thus, the S-scheme pulse sequence shown in Figure 3.1A permits quantitative determination of rate constants using the same calculation approach as that used for the conventional non-TROSY $^{15}N_z$-exchange experiment.

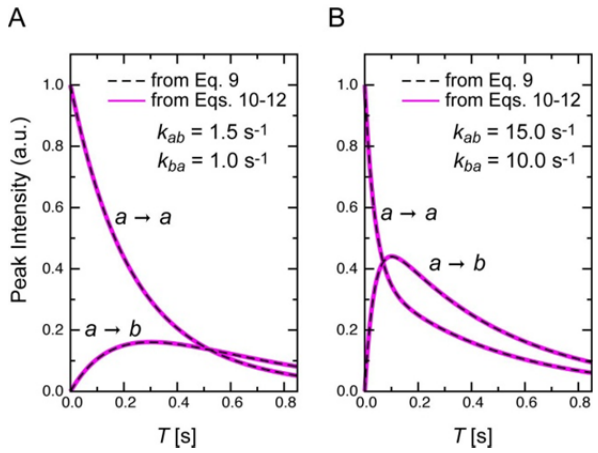

Figure 3.3. Simulations of time-courses of auto- and exchange-peak intensities for the TROSY-based *z*-exchange experiment with the S-scheme. Two panels show results of simulations employing different set of kinetic rate constants (A, k_{ab} = 1.5 s^{-1} and k_{ba} = 1.0 s^{-1}; B, k_{ab} = 15.0 s^{-1} and $_{kba}$ = 10.0 s^{-1}). The other parameters employed in each calculations are as follows: R_{zz}^{a}= 4.8 s^{-1}, R_{z}^{a}= 0.8 s^{-1}, η_{z}^{a}= 0.5 s^{-1}, R_{zz}^{b} = 3.0 s^{-1}, R_{z}^{b}= 0.5 s^{-1}, and η_{z}^{b}= 0.3 s^{-1}. The solid magenta lines are calculated with Eqs. 3.10-3.12 for (*a*→*a*) auto- and (*a*→*b*) exchange-peaks, whereas the dotted black lines are obtained using Eq. 3.9. For all examined values of the parameters R_{zz}, R_z, η_z, k_{ab} and k_{ba}, the results from Eqs. 3.10-3.12 were found to be identical with those with Eq. 3.9. Thus rate constants can be obtained in the same way as that for the conventional non-TROSY *z*-exchange experiment.

RESULTS AND DISCUSSION

We used the TROSY-based *z*-exchange experiment to study the temperature dependence of the kinetics of translocation of the homeodomain transcription factor HoxD9 from a specific target site on one DNA molecule to the specific target site on another DNA molecule using the 'mixture approach' employed in our previous studies (6, 12, 13). In this approach three macromolecular components are mixed together: ^{2}H/^{15}N-labeled HoxD9 homeodomain, and two 24-bp DNA duplexes *a* and *b* (Figure 3.4A). The two DNA duplexes *a* and *b* are identical except for a single base pair mutation immediately adjacent to the central 6 base-pair specific target site and the

affinities of HoxD9 for the two duplexes are virtually identical (12). For some residues, the chemical shifts of the ^{1}H-^{15}N correlation cross-peaks arising from the two complexes are slightly different owing to the difference in DNA sequence, thereby permitting us to study the kinetics of the exchange reaction in which the homeodomain is transferred from one DNA molecule to another. Previous studies have shown that at high concentrations of free DNA ($>10^{-6}$ M) intermolecular translocation of the HoxD9 homeodomain takes place predominantly through a direct transfer mechanism rather than via a two-step mechanism involving dissociation of the protein into free solution followed by re-association (12). Direct transfer, also known as 'inter-segment transfer', is a second-order reaction in which collision between the protein-DNA complex and free DNA mediates intermolecular translocation (35-37, 89, 90). As described previously (12), the behavior of longitudinal components of the magnetization for the direct transfer process can be described by Eq. 3.9, in which the apparent translocation rate is given by the product of the second-order rate constant for direct transfer and the concentration of free DNA. Under the present experimental conditions, translocation of HoxD9 between the two DNA duplexes *a* and *b* is in the slow exchange regime. Figure 3.4B shows the auto- and exchange-peaks arising from the backbone amide group of Arg-5 at 20 ˚C at three mixing times (20, 54 and 146 ms) using the TROSY-based *z*-exchange ^{1}H-^{15}N correlation experiment. The exchange peaks are apparent in the spectra recorded with the two longer mixing times. Using data collected at eight different mixing times, we were able to determine the translocation rates by non-linear least-squares fitting to the experimental time dependence of the exchange and auto-peak intensities.

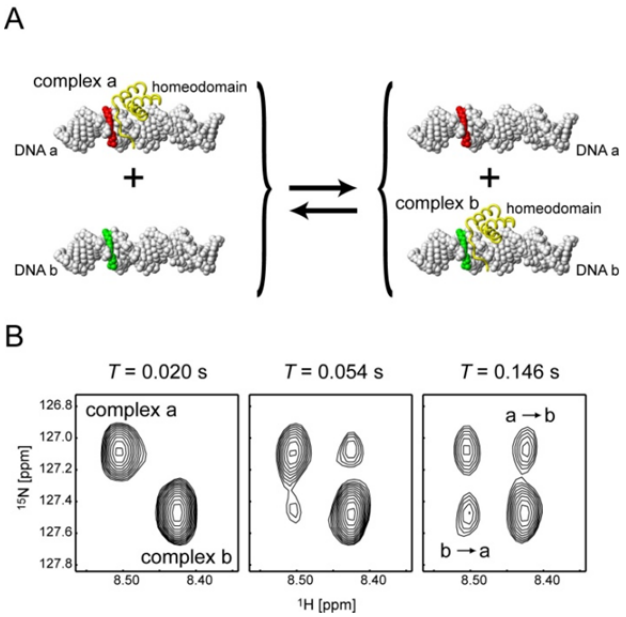

Figure 3.4. Intermolecular translocation of HoxD9 between specific sites located on different DNA duplexes. (A) The system studied: the NMR sample contains three macromolecular components: the $^{2}H/^{15}N$-labeled HOXD9 homeodomain, and two 24-bp DNA duplexes *a* and *b*. Red (C•G) and green (A•T) represent the base pair at position 8 that is different between DNA duplexes *a* and *b*. (B) Auto- and exchange peaks arising from the backbone amide group of Arg-5 observed in the TROSY-based *z*-exchange experiment recorded with three different mixing times *T*. The spectra were obtained at 20 °C.

Comparison of TROSY-based and non-TROSY N_z-exchange experiments.

Data on the same sample were collected at 20°C using the conventional non-TROSY $^{15}N_z$ exchange experiment (21) to compare the reliability of the rate constants derived from the non-TROSY and TROSY-based pulse sequences. As is evident from Figure 3.5, the time-courses for the auto- and exchange-peak intensities observed for the two experiments appear to be quite different. This is simply due to the fact that the apparent longitudinal relaxation rate during the mixing period is faster for the TROSY-based *z*-exchange experiment due to the 50% contribution from R_{zz} (see Eq. 3.8). However, the protein translocation rates, k_{ab} and k_{ba}, determined by least-squares analysis using Eq. 3.9

are identical within experimental error for the two experiments (Figure 3.5). This observation provides experimental confirmation that the TROSY-based and conventional non-TROSY *z*-exchange experiments can be analyzed in the same way.

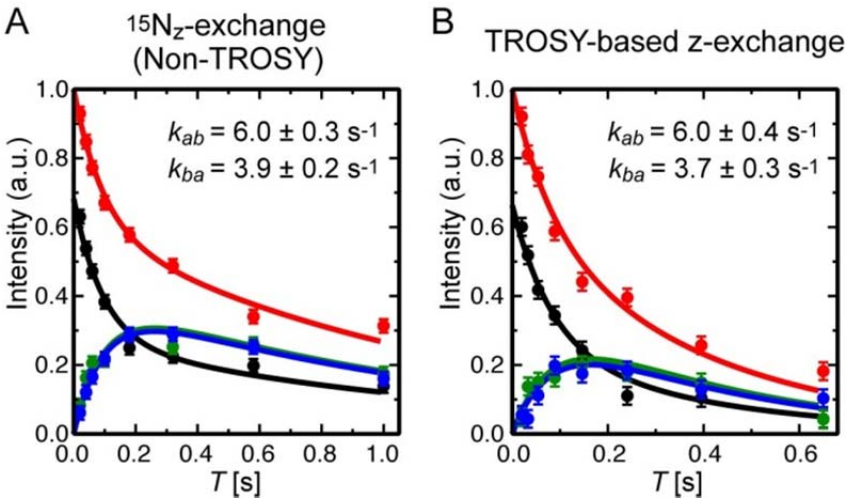

Figure 3.5. Comparison of (A) the conventional non-TROSY $^{15}N_z$-exchange experiment (21) and (B) the present TROSY-based *z*-exchange experiment for quantitative evaluation of rate constants. The experimental intensities for the auto- and exchange-peaks of Arg-5 (20 °C) as a function of mixing time together with the best-fit theoretical curves obtained by non-linear least-squares optimization are shown (black, a→a; red, b→b; green, a→b; blue, b→a). The values of k_{ab} and k_{ba} obtained are displayed in the figures and are identical within experimental error for the two experiments. The longitudinal relaxation rates for complexes *a* and *b* were assumed to be identical. The values of the relaxation rates were calculated to be $\overline{R}$ = 2.8 s^{-1} for the TROSY-based z-exchange experiment and R_z = 0.8 s^{-1} for the conventional non-TROSY $^{15}N_z$-exchange experiment.

The lineshapes of the auto- and exchange- peaks in the TROSY-based *z*-exchange ^{1}H-^{15}N correlation experiment are significantly narrower than those in the $^{15}N_z$-exchange experiment, which is especially advantageous at low temperature and high magnetic field. This is clearly illustrated in Figure 3.6 which provides a comparison of the two experiments for the auto- and exchange-peaks of the backbone amide group of Thr-9 measured at 8 °C and a ^{1}H-frequency of 800 MHz (with a mixing time of 0.38 s). In the case of Thr-9 the ^{1}H- and ^{15}N-chemical shift differences between complexes *a* and *b* are relatively small. As a result the auto- and exchange-peaks of Thr-9 are not resolved in the

conventional non-TROSY $^{15}N_z$-exchange (Figure 3.6A) but are clearly resolved in the TROSY-based *z*-exchange experiment due to the better lineshapes (Figure 3.6B). Thus, the TROSY-based *z*-exchange experiment is useful for kinetic analysis either at low temperature and/or for a large molecular weight systems.

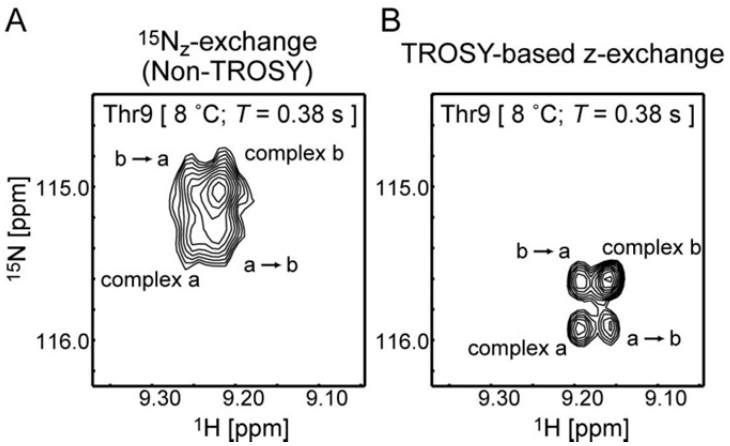

Figure 3.6. The TROSY-principle improves separation of auto- and exchange-peaks at low temperature. Signals from Thr-9 obtained at 8 ˚C with (A) the conventional non-TROSY $^{15}N_z$-exchange experiment[5] and (B) with the TROSY-based *z*-exchange experiment (Figure 1A) using the same mixing time (*T* = 0.38 s) are displayed. Data were measured using the same digital resolution and processed in an identical manner.

Determination of the activation energy for intermolecular translocation of HoxD9 between specific sites on different DNA molecules

Using the TROSY-based *z*-exchange experiment, we measured the translocation rates at 8, 15, 20, 30 and 35 °C. The translocation rates at 8 ˚C were ~15 fold slower than those at 35 °C. Figure 3.7A shows Eyring plots of the temperature-dependence of the translocation rates k_{ab} and k_{ba}. The activation enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) were determined using the Eyring equations ($\ln k / T = -\Delta H^{\ddagger} / RT + \ln k_B / h + \Delta S^{\ddagger} / R$, where

k_B is the Boltzmann constant; h, Planck's constant; and R, the gas constant), The obtained values were $\Delta H^{\ddagger}$ = 17.1 ± 1.0 kcal•mol^{-1} and $\Delta S^{\ddagger}$ = 3.7 ± 3.4 cal•mol^{-1}•K^{-1} for the translocation from DNA *a* to *b*; $\Delta H^{\ddagger}$ = 17.3 ± 1.6 kcal•mol^{-1} and $\Delta S^{\ddagger}$ = 3.9 ± 5.6 cal•mol^{-1}•K^{-1} for *b* to *a*. Since activation free energies at 298 K ($\Delta G^{\ddagger}_{298K}$) are calculated to be 16 kcal•mol^{-1}, the energy barrier for protein translocation between DNA molecules is primarily enthalpic in origin.

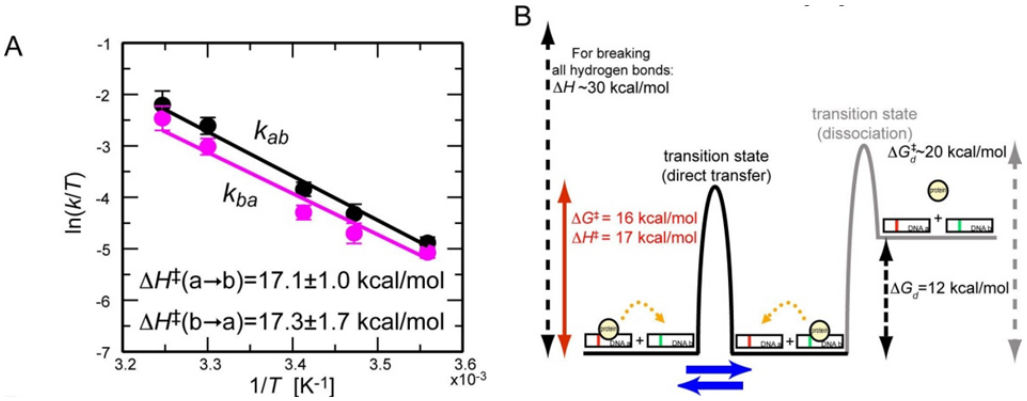

Figure 3.7. Energetics of translocation of HoxD9 between specific sites on different DNA molecules. (A) Eyring plots of the translocation rates measured at 8, 15, 20, 30 and 35 °C (k_{ab}, black; k_{ba}, magenta). The activation enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) were determined using the Eyring equation (see text). (B) Schematic diagram comparing the energetics of translocation (red) and dissociation (see text). Blue arrows represent the intermolecular translocation process characterized in the present study. The enthalpy for breaking all intermolecular hydrogen bonds (~30 kcal/mol) was estimated from the crystal structures of two highly homologous homeodomain-DNA complexes (PDB codes 1IG7 (89) and 9ANT (88)).

It is interesting to compare the energetics of translocation with those for the dissociation process from the bound state to the free state. The free energy difference between the two states (ΔG_d) can be calculated from the equilibrium dissociation constant. For the specific interaction between the HoxD9 homeodomain and the 24-bp DNA duplexes containing the specific target sequence, ΔG_d at 298 K is calculated to be 12 kcal/mol. The rate constant k_{off} for dissociation of HoxD9 from its specific DNA site,

determined from gel shift assays, is 0.01 s^{-1} (93). From this value and the equation $k_{off} = (k_B T / h)\exp\{-\Delta G_d^{\ddagger} / (RT)\}$, the activation free energy for the dissociation process ($\Delta G_d^{\ddagger}$) is estimated to be 20 kcal/mol. Thus, the energy barrier for translocation through the direct transfer mechanism (12) is significantly lower than that for dissociation. A schematic comparison of the energetics of translocation and dissociation is shown in Figure 3.7B.

CONCLUDING REMARKS

In this chapter, we have presented TROSY-based *z*-exchange spectroscopy for quantitative measurement of rate constants for a system in the slow exchange regime (20). The TROSY-based *z*-exchange experiment presented here permits quantitative determination of rate constants in the same manner as that employed for conventional non-TROSY exchange experiments. Because of the use of the TROSY principle, kinetic measurements are feasible over a wider range of temperatures and for larger molecular weight systems than was heretofore possible, thereby making feasible analysis of activation energies for biologically important processes involving macromolecules.

We now possess a very powerful NMR based tool to analyze the kinetics of biological processes in the slow exchanging regime. This adds to the arsenal of biophysical approaches to analyze the kinetics of biological processes happening in various time scales. Since these biophysical approaches have been extensively used to obtain relevant kinetic rate constants for biological processes, there is need to build a generalized N-state model to check for the accuracy of the macroscopic rate constants and the validity range of these NMR based approaches.

CHAPTER IV

COMPUTATIONAL VALIDATION OF NMR BASED APPROACHES FOR INVESTIGATING THE KINETICS OF PROTEIN TRANSLOCATION ON DNA

INTRODUCTION

The DNA binding proteins need to efficiently find their target site among an enormous ocean of nonspecific DNA efficiently in order to produce proper cellular response to external stimuli. The DNA binding proteins exploit some kinetic shortcuts during target search and this process of translocation happens very fast, in a matter of seconds (65) and. The process of target location of a protein translocating on DNA is greatly enhanced by mainly combination of these three processes, namely, sliding or one-dimensional diffusion, dissociation followed by re-association, and inter-segment transfer. In order to study the different modes of translocation, a few biophysical methods have been developed like nuclear magnetic resonance, Förster resonance energy transfer (FRET), fluorescence intensity/anisotropy and single molecule fluorescence based approaches.

Since the different modes of protein translocation differ from each other in terms of kinetics, careful kinetics experiments have been employed to delineate the role of each mode in facilitated target location. The direct observation of protein translocation on DNA has been made possible by some innovative single molecule experiments (reviewed in (8, 10)) that show the importance of non-specific sites on DNA and one-dimensional sliding process. Bulk measurements using fluorescence based methods have been used to study the dissociation followed by re-association processes which is limited by the slowest process i.e. dissociation. The role of intersegment transfer or direct transfer has

been very poorly understood until the development of new NMR based methods to directly observe this mechanism. This mechanism is very relevant in the facilitated target search process applicable to cases involving high concentrations of DNA as inside the cells, where protein is mostly bound to DNA and the transfer from one portion of DNA to another is feasible. In order to study the intersegment transfer of protein between DNA molecules the "mixture approach" was used. The NMR based mixture approach is performed in a sequence of three steps, the first two involve the individual mixture of protein with two different DNA molecules where some residues show slightly different chemical shifts owning to slight differences in hydrogen bonding with DNA. The third step involves mixture of protein with 1:1 mixture of both DNA molecules, which could lead to either of these scenarios. If the system is in a slow exchanging regime, we observe two sets of signals where z-exchange methods can be implemented to determine the translocation rate constants (12, 20). If the system is in the fast exchange regime, we observe one signal at the population average position where Lorentzian line-shape fitting can be employed to determine the translocation rate constants (6).

Our current work involves the validation of the recent structural and kinetic characterization of homeodomain translocating between two nonspecific DNA molecules (6) and also two cognate DNA molecules (12, 20). In the case of HoxD9 homeodomain translocation between two non-specific molecules, the experimental approach makes use of Lorentzian line-shape fitting (6) along with Reuben and Fiat approximation (51) valid only for two-state exchange process, to describe the complex translocation process. In order to build a theoretical model to describe these NMR based experiments we here developed a generalized N-state model using McConnell equations (63, 94) along with Kramers theory (95). The different translocation mechanisms were considered separately in the kinetic matrix (see Chapter 2 for details) used in the McConnell equations. When

the protein is bound to different sites on the same DNA molecule there could be the possibility of multiple microstates, which is not taken into account while using the simplistic Reuben and Fiat approach (6, 51). Hence, there is a need for the judicious assessment of this NMR based experiment to determine the rate constants for translocation between nonspecific DNA molecules.

In the case of HoxD9 homeodomain translocating between two cognate DNA molecules, the NMR based mixture approach lies in the slow-exchanging regime as discussed in the previous chapter. Just like the previous case microstates could exist on individual DNA molecules and this could have impact on the precision of the measured apparent macroscopic kinetic rate constant for intersegment transfer. Hence we also need to assess the accuracy of the intersegment transfer rate constant for HoxD9 homeodomain translocating between two cognate DNA molecules determined by the DNA concentration dependent z-exchange experiment (96, 97).

THEORETICAL CONSIDERATION

Kinetics of protein translocation between two DNA molecules

As has been stated earlier, the process of target location of a protein translocating on DNA is greatly enhanced by the combination of three individual processes (98, 99), namely, (i) sliding, (ii) dissociation followed by re-association, and (iii) inter-segment transfer (99, 100). Here we describe the theoretical consideration of this highly complex translocation process of protein between two DNA molecules.

Let us consider that there are n states present on DNA-a and DNA-b where protein can bind and translocate. The number of states "n" on each DNA of length n_{DNA} is defined by $2*(n_{DNA} - r_{sp} + 1)$, where r_{sp} is the recognition region occupied by protein on DNA and this also includes two different recognition orientations of the protein bound on

DNA. For this theoretical calculation, we generalize our assumption for a N-state system. We consider there are N states on each DNA molecule (N=2*n, DNA-a and DNA-b combined) and one free state. The first N/2 states correspond to the states of protein bound on DNA-a, next N/2 states correspond to protein bound to DNA-b, and N+1th state corresponds to the free state. Thus all of the matrices are of the size N+1 by N+1.

Here we make a safe assumption that our macromolecular translocation processes are governed by Kramers' theory (95, 101, 102) of Brownian motions over potential barriers, thus defining the free energies for the potential barriers corresponding to the sliding process (G_s^{TS}), intersegment transfer process (G_h^{TS}) and dissociation and re-association (G_{free}^{TS}) of protein with DNA. The reaction rates used for chemical reactions are mostly based on transition state theory that is mostly valid for systems with low viscosity, thus making the Kramers' theory more suitable for this case. Both theories depend on the boundaries dividing the initial and final states and only differ in their frequency factors. According to the Kramers' theory, the populations and kinetic rates of individual sites are governed by the difference in free energies between the initial state and the potential barrier (95, 101, 102). In order to create a diverse energy landscape for the nonspecific sites, the free energies (G_i) for the ith (where i<N) states are normally distributed. The higher affinity to specific sites can be easily modeled (102), by using lower values of free energy (G_{sp}) compared to G_i. Most of the proteins are bound to the DNA under our experimental conditions (6, 12, 20), but a very small population of protein exists in the free state dominated by the values of k_{on} and k_{off}.

This NMR based approach was used to determine the translocation kinetics of the facilitated target location of the protein on DNA, employed on a system under dynamic equilibrium (6, 96, 97). One of the direct consequences of using the Kramers' theory on systems under equilibrium is that the kinetic rate constants are solely dependent on the

difference in the free energies between the transition state and the initial state, thus taking into account the detailed balance of all the elements involved while creating the kinetic matrices.

Since the free energies are defined for individual states, we can obtain population of the i[th] state (102) by $P_i = \exp\left(\frac{-G(i)}{RT}\right)/Q$ and the population of protein in free state is given by $P_{free} = \exp\left(\frac{-G_{free}}{RT}\right)/Q$, where Q corresponds to the partition function for N states defined as $\sum_{h=1}^{N}\exp\left(\frac{-G(h)}{RT}\right)+\exp\left(\frac{-G_{free}}{RT}\right)$. Here N is 2*n for the mixture of two DNA and n for individual protein-DNA complexes.

The kinetic matrix is defined by the sum of matrices referring to different translocation mechanisms such as sliding (K_{intra}^{ab}), inter-molecular translocation (K_{inter}^{ab}), and dissociation/association process (K_{free}^{ab}).

$$K_{macro}^{ab} = K_{\text{intra}}^{\text{ab}} + K_{\text{inter}}^{\text{ab}} + K_{free}^{ab}$$

The process of intermolecular translocation enables the protein to move from one state on the DNA to another without the intermediate free-state, which is also sometimes referred to as “hopping” or “direct transfer”. Hence this matrix can be defined by the diagonal elements to be the outgoing components, which refers to the transfer from the i[th] state to all other N-1 states. Since the kinetic rate constant is dependent on the free energy of the initial state and the final state (G_h^{TS}), the sum of all outgoing terms will be $(N-1)*A*\exp\left(-\frac{G_h^{TS}-G_i}{RT}\right)$. All the off diagonal elements are the incoming component, referring to the term $A*\exp\left(-\frac{G_h^{TS}-G_i}{RT}\right)$. The last column and row refers to the free state of the protein that is dealt separately in the K_{free}^{ab} matrix.

$$K^{ab}_{inter}(i,j) = \begin{cases} (N-1) * k^{i}_{h}, & i = j \;\&\; i \le N \\ -k^{j}_{h}, & i \ne j \;\&\; i \le N \;\&\; j \le N \\ 0, & \text{otherwise} \end{cases}$$

In order to address the sliding of protein on a DNA, we consider the K^{ab}_{intra} matrix. Since the protein can slide from one state to its adjacent state, the diagonal and its adjacent elements are occupied. The protein is at the end of the DNA molecule at states 1, N/2, (N/2)+1, and N, hence the outgoing components are $A * \exp\left(-\frac{G_s^{TS} - G_i}{RT}\right)$ as the diagonal elements. Since the number of states also includes different orientation of the protein on DNA, the sliding from one orientation to another is impossible and thus similar considerations as above is adopted for positions N/4, (N/4)+1, 3N/4, and (3N/4)+1. For all other diagonal terms, the protein can slide to either side and hence the outgoing term is $2A * \exp\left(-\frac{G_s^{TS} - G_i}{RT}\right)$. Since the protein can only slide to its adjacent site(s) on the DNA, all other off diagonal elements except for the state corresponding to the adjacent state(s) are zero.

$$K^{ab}_{intra}(i,j) = \begin{cases} k_s^i, & i = j \;\&\; i \ne N+1 \;\&\; i \in \{1, N, N/2, (N/2)+1, N/4, (N/4)+1, 3N/4, (3N/4)+1\} \\ 2k_s^i, & i = j \;\&\; i \notin \{1, N, N/2, (N/2)+1, N/4, (N/4)+1, 3N/4, (3N/4)+1, N+1\} \\ -k_s^{i-1}, & i = j+1 \;\&\; j \le N \;\&\; i \notin \{N/4, N/2, 3N/4, N+1\} \\ -k_s^{i+1}, & i = j-1 \;\&\; j \le N \;\&\; i \notin \{(N/4)+1, (N/2)+1, (3N/4)+1, N+1\} \\ 0, & i > N \text{ or } j > N \\ 0, & \textit{otherwise} \end{cases}$$

Our experimental conditions are such that the DNA concentration is twice as much as protein concentrations; hence the population of the free protein is small, which is taken into account using the K^{ab}_{free} matrix. Since the rate of dissociation is limited by k_{off} and the rate of association is limited by k_{on}, all the diagonal elements (for i ≤ N) are k_{off}

and the N+1[th] column is k_{on}. The last row in this matrix contains elements corresponding to the outgoing terms for the free protein state, which takes care of the detailed balance.

$$K^{ab}_{free}(i,j)=\begin{cases} k^i_{off} & ,i=j \;\&\; i\le N \\ -k^i_{off} & ,j=N+1 \\ -k_{on} & ,i=N+1 \\ (2N-1)k_{on} & ,i=N+1 \;\&\; j=N+1 \\ 0 & ,\text{otherwise} \end{cases}$$

In order to simulate individual complexes of protein with DNA-a and with DNA-b, the kinetic matrices are modified such that the terms for the states on the other DNA are replaced by zeros. A sample kinetic matrix for the individual and 1:1 mixture of

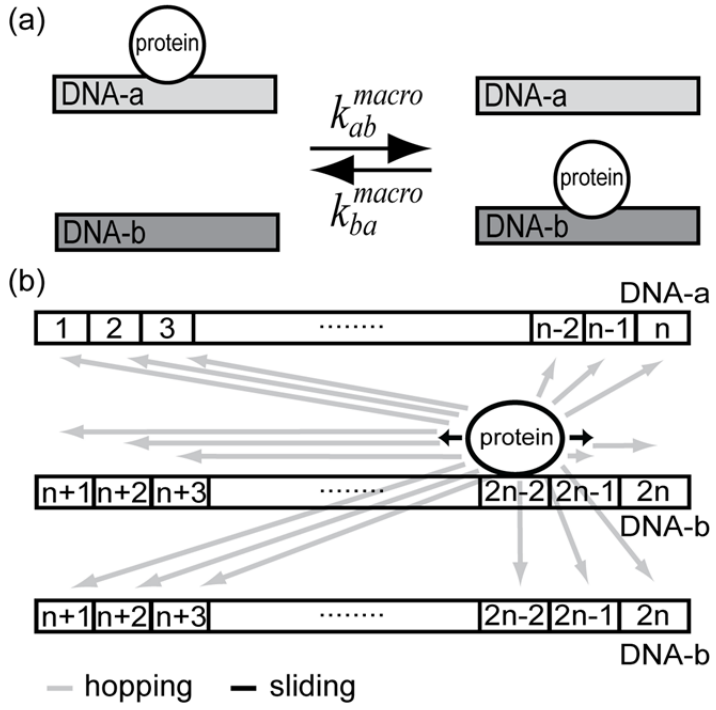

Figure 4.1: Macroscopic and microscopic events during protein translocation on DNA (a) Protein translocation between two DNA molecules (DNA-a and DNA-b) (b) Different translocation mechanisms of the protein to reach other 2n-1 sites on DNA-a and DNA-b

protein DNA complex is shown in detail in the appendix section for a system with two DNA molecules containing 3 sites, i.e. 6 states each.

Derivation of macroscopic rate constant for protein translocation between two DNA molecules

Let us consider a system exchanging between two major macroscopic states "a" and "b", comprising of n microstates each. This leads to describing 2*n microstates, where 1 to n constitute the macroscopic state "a" and n+1 to 2n constitute the macroscopic state "b". The rate of change of population of i^{th} state is in the form of the master equation,

$$\frac{dP_i}{dt} = -\sum_{\substack{m=1 \\ m\neq i}}^{2n} k_{im} P_i + \sum_{\substack{m=1 \\ m\neq i}}^{2n} k_{mi} P_m \tag{4.1}$$

On the macroscopic level, the change in the population of states "a" and "b" under equilibrium is given by the sum of all its microscopic state,

$$\frac{dP_A^{eq}}{dt} = \sum_{i=1}^{n} \frac{dP_i^{eq}}{dt} = -\sum_{i=1}^{n}\sum_{\substack{m=1 \\ m\neq i}}^{2n} k_{im} P_i^{eq} + \sum_{i=1}^{n}\sum_{\substack{m=1 \\ m\neq i}}^{2n} k_{mi} P_m^{eq} \tag{4.2}$$

$$\frac{dP_B^{eq}}{dt} = \sum_{i=n+1}^{2n} \frac{dP_i^{eq}}{dt} = -\sum_{i=n+1}^{2n}\sum_{\substack{m=1 \\ m\neq i}}^{2n} k_{im} P_i^{eq} + \sum_{i=n+1}^{2n}\sum_{\substack{m=1 \\ m\neq i}}^{2n} k_{mi} P_m^{eq} \tag{4.3}$$

Comparing this form to a two-state exchange process, the differential rate equations are given by,

$$\frac{dP_A}{dt} = -k_{AB}^{macro} P_A + k_{BA}^{macro} P_B \tag{4.4}$$

$$\frac{dP_B}{dt} = -k_{BA}^{macro} P_B + k_{AB}^{macro} P_A \tag{4.5}$$

The population of the i^{th} state under equilibrium is given by $P_i^{eq} = \exp\left(-\frac{G_i}{RT}\right) \Big/ \sum_{\gamma=1}^{2n} \exp\left(-\frac{G_\gamma}{RT}\right)$ and the kinetic rate constant for the transition from i^{th}

state to any other state is based on the Kramers' theory, can be defined by $k_{im} = A * \exp\left(-\frac{G^{TS} - G_i}{RT}\right)$. After rearrangement of terms in Eq. 4.2 and comparing with Eq. 4.4 we obtain the macroscopic rate constant,

$$k_{AB}^{macro} = \left(\sum_{i=1}^{n}\left\{P_i^{eq}\sum_{m=n+1}^{2n}k_{im}\right\}\right)\Big/P_A^{eq} \tag{4.6}$$

$$k_{AB}^{macro} = \frac{\sum_{i=1}^{n}\left\{\frac{\exp\left(-\frac{G_i}{RT}\right)}{\sum_{\gamma=1}^{2n}\exp\left(-\frac{G_\gamma}{RT}\right)}\cdot\frac{nA\exp\left(-\frac{G^{TS}}{RT}\right)}{\exp\left(-\frac{G_i}{RT}\right)}\right\}}{\frac{\sum_{i=1}^{n}\exp\left(-\frac{G_i}{RT}\right)}{\sum_{\gamma=1}^{2n}\exp\left(-\frac{G_\gamma}{RT}\right)}} \tag{4.7}$$

Further simplification of Eq. 4.7 yields, $k_{AB}^{macro} = n^2\left(\sum_{i=1}^{n}(k_{ib})^{-1}\right)^{-1}$. Similarly, upon rearrangement of terms in Eq. 4.3 and comparing with Eq. 4.5 we obtain $k_{BA}^{macro} = n^2\left(\sum_{i=n+1}^{2n}(k_{ia})^{-1}\right)^{-1}$. It is very interesting to note that this form is very similar to the formalism of apparent reaction rate of consecutive reactions. The rate constants obtained are in the macroscopic scale since they are obtained from a system under equilibrium where the protein translocates between DNA molecules using the different mechanisms described above.

Translocation between two non-specific DNA molecules

The mixture approach is described by considering different matrices for individual protein and DNA-a/DNA-b complexes, and protein with 1:1 mixture of DNA-a and DNA-b.

The transverse magnetization vector for this system is described below

$$\mathbf{m}_{tr} = \begin{pmatrix} M_{+,1} \\ \vdots \\ M_{+,N} \\ M_{+,\text{free}} \end{pmatrix}, \text{ where } M_+ = M_x + iM_y$$

The relaxation matrix is a diagonal matrix containing transverse relaxation rates for individual microstates,

$$\mathbf{R}_{tr}(i,j) = \begin{cases} R_i, & i = j \\ 0, & i \neq j \end{cases}$$

Since the protein has the same structure and conformation when bound to DNA, the transverse relaxation for the first N states on the DNA can be assumed to be the same. But the relaxation properties for the free state could be quite different, which is taken into account in the $R_{tr}(N+1,N+1)^{\text{th}}$ term.

The chemical shift matrix is described in the same way with chemical shifts for individual states as the diagonal elements.

$$\mathbf{\Omega}_{tr}(i,j) = \begin{cases} \Omega_i, & i = j \\ 0, & i \neq j \end{cases}$$

The first N/2 diagonal elements (for $i \leq N/2$) represent the chemical shifts of the protein bound to DNA-a and the next N/2 diagonal elements (for $N/2 < i \leq N$) represent

the chemical shifts of protein bound to DNA-b. The chemical shift corresponding to the free state (for i = N+1) of the protein could be quite different and Ω_{N+1} is set accordingly. As described above the kinetic matrix is constituted of the sum of three different matrices describing the different translocation processes. For this, we use the McConnell equation that relates NMR transverse magnetizations to the kinetics of the system under equilibrium in the following manner.

$$\frac{d}{dt}\mathbf{m}_{tr} = -(\mathbf{R}_{tr} + \mathbf{K} - i\boldsymbol{\Omega}_{tr})\mathbf{m}_{tr} \qquad (4.8)$$

Translocation between two specific DNA molecules

In order to determine the translocation rate constant for a protein translocating between two specific DNA molecules, the mixture approach was successfully implemented along with z-exchange (12) and TROSY based z-exchange spectroscopy (20). The sample used in this measurement is comprises of a protein along with 1:1 mixture of specific DNA molecules (spDNA-a and spDNA-b), both containing the cognate target site for the protein with a slight variation in the flanking residues around the cognate sites.

The simulations of the z-exchange experiments was performed in order to determine the translocation kinetics of protein between two cognate DNA molecules, where we monitor the longitudinal magnetization corresponding to the translocation of the protein from spDNA-a to spDNA-a, spDNA-a to spDNA-b, spDNA-b to spDNA-b, and spDNA-b to spDNA-a as a function of mixing time. We can use the matrices defined in above section for a N-state system. We consider there are N states on the specific DNA molecules (spDNA-a and spDNA-b combined) and one free-state. The first N/2 states correspond to the states of protein bound on spDNA-a, next of N/2 states correspond to

that on spDNA-b, and N+1th state corresponds to the free state. Thus all of the matrices are of the size N+1 by N+1. The specific site on spDNA-a ($i=sp^a$) is chosen near the center of spDNA-a with lower free energy G_{sp} as compared to other non-specific sites. Similarly the specific site on spDNA-b ($i=sp^b$) is also chosen near the center of spDNA-b.

The longitudinal magnetization vector for a N+1 state system can be described by

$$\Delta \mathbb{m}_{lo} = \begin{pmatrix} M_{z,1} - M_{\infty,1} \\ \vdots \\ M_{z,N} - M_{\infty,N} \\ M_{z,free} - M_{\infty,free} \end{pmatrix}$$

The longitudinal relaxation matrix for this system under equilibrium, is described by a diagonal matrix

$$\mathbb{R}_{lo}(i,j) = \begin{cases} R_{lo,i}, & i = j \\ 0 \ , & i \neq j \end{cases}$$

where the longitudinal relaxation of individual states constitute the diagonal terms. The longitudinal relaxation of the free state could be quite different, which is taken care of in $R_{lo}(N+1,N+1)^{th}$ term.

The kinetic matrix ($\mathbb{K}$) for this system can be carried over from the description for the complex of protein and 1:1 mixture of the two DNAs in the section above, with the exception of free energies for different states on the DNAs being replaced by the free energies of the cognate DNA molecules.

Here we use the McConnell's formalism to describe the relationship between longitudinal magnetization and kinetics by the following equation,

$$\frac{d}{dt}\Delta \mathbb{m}_{lo} = -(\mathbb{R}_{lo} + \mathbb{K})\Delta \mathbb{m}_{lo} \qquad (4.9)$$

MATERIALS AND METHODS

Simulations of NMR data based on McConnell equations

All of these simulations were coded to run on MATLAB programming environemtnts (The MathWorks, Natick, MA). A rugged landscape of free energies for the nonspecific sites on the DNA was defined by creating a random distribution of free energies with mean of zero and standard deviation of 500 cal/mol (103). To create specific sites on the DNA with high affinity towards protein, the free energies of those specific states were decreased by 4 kcal/mol.

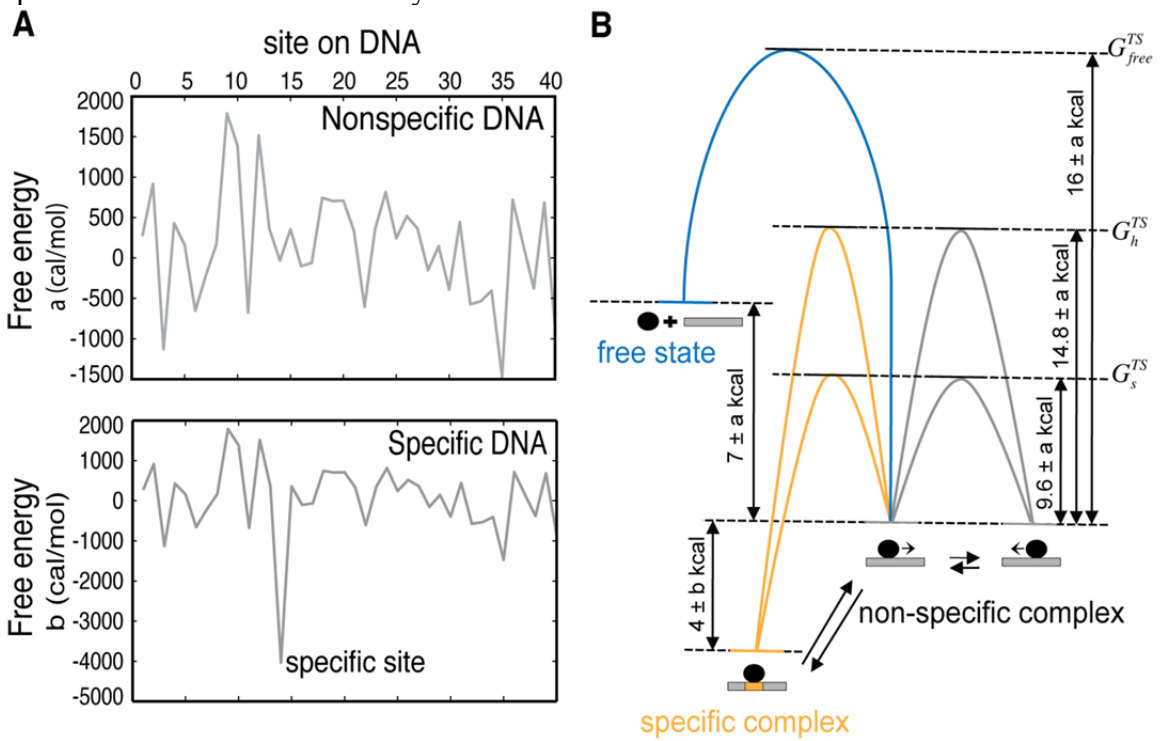

Figure 4.2: Free energies to describe protein translocation (A) Definition of free energies of protein bound to nonspecific (top) and specific DNA(bottom) with 40 sites each. (B) The free energy landscape of the various states defined in our simulations.

The free energy of the free state of the system is set using free DNA concentration and K_d ($G^{TS}_{free} = G_{ref} + RT \cdot \ln([\text{FreeDNA}]/K_d)$), in most cases the population of free state is

minimal due to high affinities of the protein towards DNA. The potential barriers for the three dynamic processes, sliding, intersegment transfer and dissociation followed by association were separately defined. To create the condition where the kinetic rate constant of sliding is far greater than the intersegment transfer process, the values of G_s^{TS}, G_h^{TS}, and G_{free}^{TS} were set at 9.8 kcal/mol, 14.8 kcal/mol, and 16 kcal/mol respectively.

Simulations of transverse magnetizations used for describing translocation between two nonspecific DNA molecules

The chemical shifts of the protein bound to DNA-a (Ω_i (for i≤N/2)) and bound to DNA-b (Ω_i (for N/2 < i ≤ N) share half of chemical shifts, since DNA-a and DNA-b have similar sites and the other half is chosen such that the difference in average chemical shifts of complex a ($\overline{\Omega}_a$) and complex b ($\overline{\Omega}_b$) is at least 50 Hz. The chemical shifts of complex a (Ω_a) and complex b (Ω_b) are shown in the Figure 4.3 (a) those represent both orientations of protein bound to DNA.

The initial conditions required to solve the partial differential Eq. 4.8 is defined by a column matrix ($\boldsymbol{m}_{tr}^{init}$) whose elements are governed by the equilibrium populations of corresponding states as described in the section above and is given by

$$\boldsymbol{m}_{tr}^{init}(m,1) = \begin{cases} P_m \ , \ m \le N \\ P_{free}, \ m = N+1 \end{cases}$$

After solving and Fourier transformation the Eq. 4.8 with matrices for individual protein and DNA complexes, and mixture of the two DNA and the protein, we get the frequency domain data. This is fitted to a Lorentzian line shape (6) which yields the apparent relaxation rate (R^{app}) and position of the peak (Ω^{app}).

In total, we now have the fitted transverse relaxation for the complex of protein and DNA-a and b (R_{ab}^{app}), complex of protein and DNA-a (R_a^{app}), and complex of protein

and DNA-b (R_b^{app}). Similarly, we also obtain the apparent chemical shifts of complexes of the protein and the mixture of two DNAs (Ω_{ab}^{app}), protein and DNA-a (Ω_a^{app}), and protein and DNA-b (Ω_b^{app}) from the position of the NMR signals. At this point we intend to compare the expected intermolecular translocation rate (k_{ex}^{macro}) with the macroscopic translocation rate (k_{ex}^{app}) obtained by using the Reuben and Fiat approximation (6, 51) described above for a two-state exchange process under equilibrium.

$$k_{ex}^{macro} = k_{AB}^{macro} + k_{BA}^{macro} = n^2\left(\left(\sum_{i=1}^{n}(k_{ib})^{-1}\right)^{-1} + \left(\sum_{i=n+1}^{2n}(k_{ia})^{-1}\right)^{-1}\right) \tag{4.10}$$

$$k_{ex}^{app} = \frac{4\pi^2 p_a p_b \left|\Omega_a^{app} - \Omega_b^{app}\right|}{R_{ab}^{app} - p_a R_a^{app} - p_b R_b^{app}} \tag{4.11}$$

where the population of the protein in complex with DNA-a (p_a) and with DNA-b (p_b) can be written as, $p_a = \dfrac{\left|\Omega_{ab}^{app} - \Omega_b^{app}\right|}{\left|\Omega_a^{app} - \Omega_b^{app}\right|}$ and $p_b = \dfrac{\left|\Omega_{ab}^{app} - \Omega_a^{app}\right|}{\left|\Omega_a^{app} - \Omega_b^{app}\right|}$ respectively.

Simulations of longitudinal magnetizations used for describing translocation between two specific DNA molecules

In order to obtain the state of the magnetization for a given mixing time (t), we need to solve the partial differential Eq. 4.9. To obtain the magnetization corresponding to the process of translocation of the protein from spDNA-a to spDNA-a and spDNA-a to spDNA-b, we solve the partial differential equation above with the following initial condition

$$\Delta m_{lo}^{init}(i,1) = \begin{cases} 1, & i = sp^a \\ 0, & otherwise \end{cases}$$

After solving the partial differential equation, the value of the magnetization at position sp^a corresponds to the magnetization of the spDNA-a to spDNA-a translocation

process and magnetization at position sp^b corresponds to the magnetization of the spDNA-a to spDNA-b translocation process.

For the following initial condition,

$$\Delta m_{lo}^{init}(i,1) = \begin{cases} 1, & i = sp^b \\ 0, & otherwise \end{cases}$$

and solving the partial differential equation described above, we obtain the magnetization corresponding to the spDNA-b to spDNA-b translocation process at position sp^b and the magnetization corresponding to spDNA-b to spDNA-a translocation process at position sp^a.

This calculation is repeated for various mixing times (t) and the longitudinal magnetization obtained for the spDNA-a to spDNA-a, spDNA-a to spDNA-b, spDNA-b to spDNA-b, and spDNA-b to spDNA-a processes are globally fitted numerically to obtain the macroscopic intermolecular rate constants k_{ab}^{app} and k_{ba}^{app}.

Simulations for different DNA concentrations

The simulation of longitudinal and transverse magnetizations on individual protein DNA complexes and mixture of protein and both DNA molecules were be performed at different DNA concentrations by increasing the intersegment transfer rate using Eq 4.12, where k^{it} is the second order rate constant for the direct transfer process.

$$k_h^{'} = k^{it} \cdot [\mathrm{DNA}] \qquad (4.12)$$

The sliding rate is remains unaffected by the DNA concentration and thus is unchanged during these simulations. The population of free protein and DNA vary with DNA concentration as the free energy of the free-state is governed using the following relation $G_{free}^{TS} = G_{ref} + RT \cdot \ln([\mathrm{FreeDNA}]/K_d)$.

RESULTS AND DISCUSSION

Assessment of NMR-based determination of rate constants for translocation between nonspecific DNA molecules

The simulations of individual protein-DNA complexes, and protein and mixture of two DNAs, along with Reuben and Fiat approximation yield us the macroscopic rate constant (k_{ex}^{app}) for translocation of protein between two nonspecific DNA molecules. The transition states are chosen such that the average sliding rate (~10^6 s^{-1}) was much larger than the average pseudo first-order intersegment transfer rate (~10 s^{-1}). For these settings different simulations were performed where the free energies of individual non-specific sites as well as chemical shifts of different sites are varied. After performing these simulations under conditions described above, the apparent rate constant (k_{ex}^{app}) obtained via the mixture approach is found to be very similar to the expected macroscopic rate constant (k_{ex}^{macro}). The results from one of the sample calculations are shown in figure 4.3 and the correlation of apparent and macroscopic rate constant for different sliding rates is shown in figure (figure 4.4 and 4.5). Thus, we obtain very similar values for apparent rate constant via the mixture approach k_{ex}^{app} and the expected macroscopic rate constant (k_{ex}^{macro}) from the derivation above (Eq. 4.10), and when the sliding rate is higher than pseudo first-order intersegment transfer at least by five orders of magnitude. To check the validity range of the Reuben and Fiat approximation, the transition states of sliding process was changed such that average sliding rates (~5000 s^{-1}) were of comparable magnitude to the pseudo first-order intersegment transfer rate (~10 s^{-1}). As per the figure (figure 4.4 and 4.5) for the conditions described above, the apparent rate constant via the mixture approach (k_{ex}^{app}) was significantly different from the expected macroscopic rate constant (k_{ex}^{macro}). When the sliding and hopping rates are comparable, we consistently observe larger value for k_{ex}^{app} as compared to expected k_{ex}^{macro}. These results suggest that

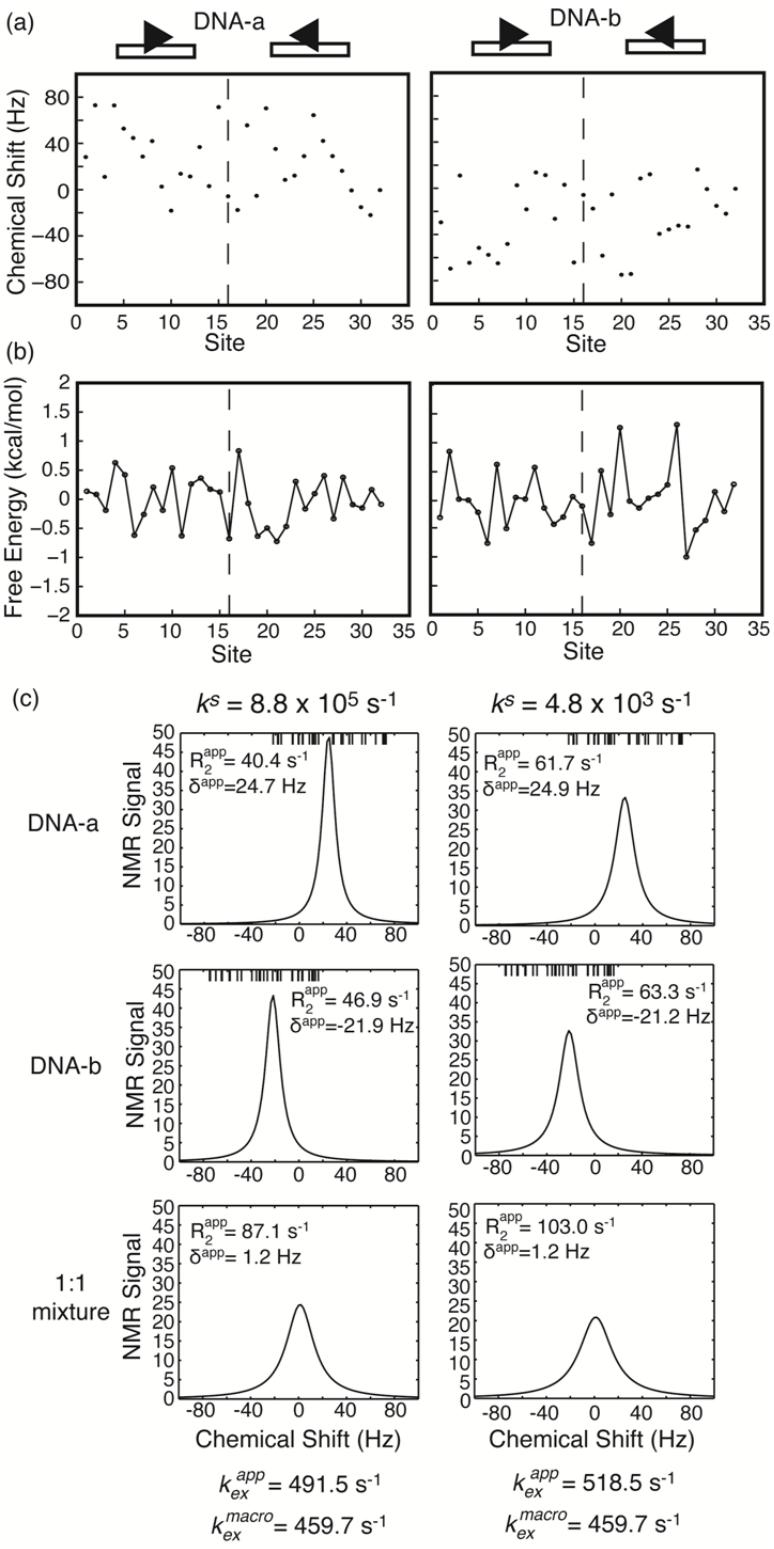
(a)
DNA-a
DNA-b
Chemical Shift (Hz)
Site
(b)
Free Energy (kcal/mol)
Site
(c)
k^s = 8.8 x 10^5 s^{-1}
k^s = 4.8 x 10^3 s^{-1}
DNA-a
DNA-b
1:1 mixture
NMR Signal
R_2^{app}= 40.4 s^{-1}
δ^{app}=24.7 Hz
R_2^{app}= 61.7 s^{-1}
δ^{app}=24.9 Hz
R_2^{app}= 46.9 s^{-1}
δ^{app}=-21.9 Hz
R_2^{app}= 63.3 s^{-1}
δ^{app}=-21.2 Hz
R_2^{app}= 87.1 s^{-1}
δ^{app}= 1.2 Hz
R_2^{app}= 103.0 s^{-1}
δ^{app}=1.2 Hz
Chemical Shift (Hz)
k_{ex}^{app}= 491.5 s^{-1}
k_{ex}^{macro}= 459.7 s^{-1}
k_{ex}^{app}= 518.5 s^{-1}
k_{ex}^{macro}= 459.7 s^{-1}

Figure 4.3: Description of individual sites on each nonspecific DNA molecule: (a) Chemical shifts and (b) free energies. (c) Results of simulation of mixture approach for two different sliding rate constants. The top two panels describe the protein bound to individual DNA molecules and the apparent rate R2 is obtained by Lorentzian line-shape fitting. The bottom panel describes the protein in the mixture the two nonspecific DNA molecules. When the sliding rate are fast and the intersegment transfer is predominant, the apparent kinetic rate constant from Reuben and Fiat approximation is very similar to macroscopic rate constant. On the other hand the apparent rate constant deviates largely from macroscopic rate constant when the sliding rates are comparable to pseudo first order rates.

that the NMR based experiments permit accurate measurement of macroscopic rate constants for protein translocation between two nonspecific DNA molecules only if sliding process is much faster than the direct transfer process.

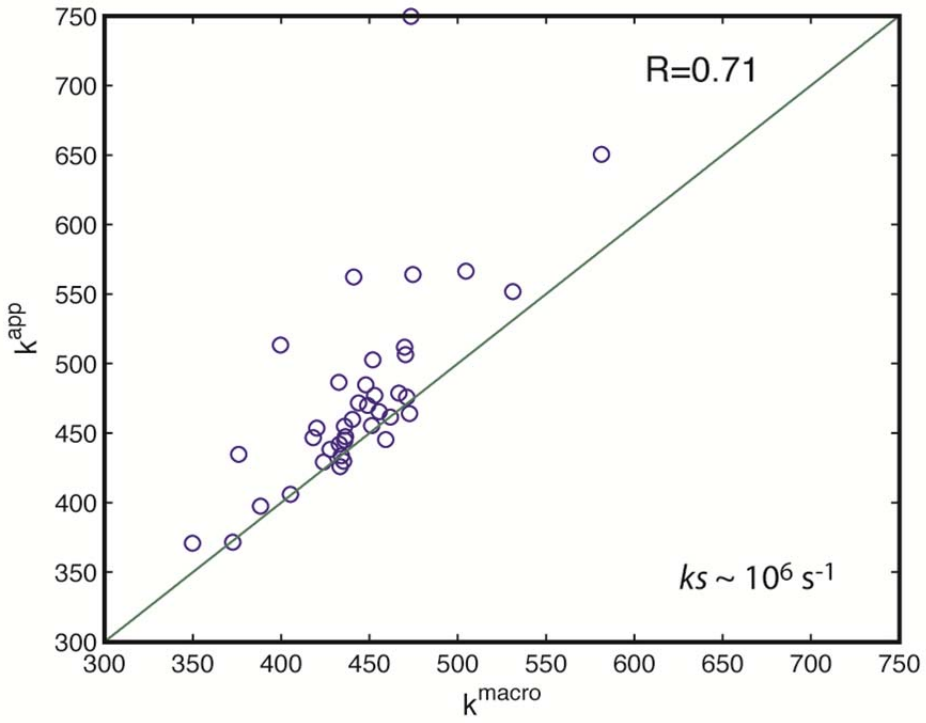

Figure 4.4: Correlation plot of k^{macro} calculated using Eq. 4.10 and the k^{app} obtained from the simulations where the sliding on nonspecific DNA is much faster (excess of 10^6 s^{-1}) than the pseudo first order rate for intersegment transfer (~10 s^{-1}).

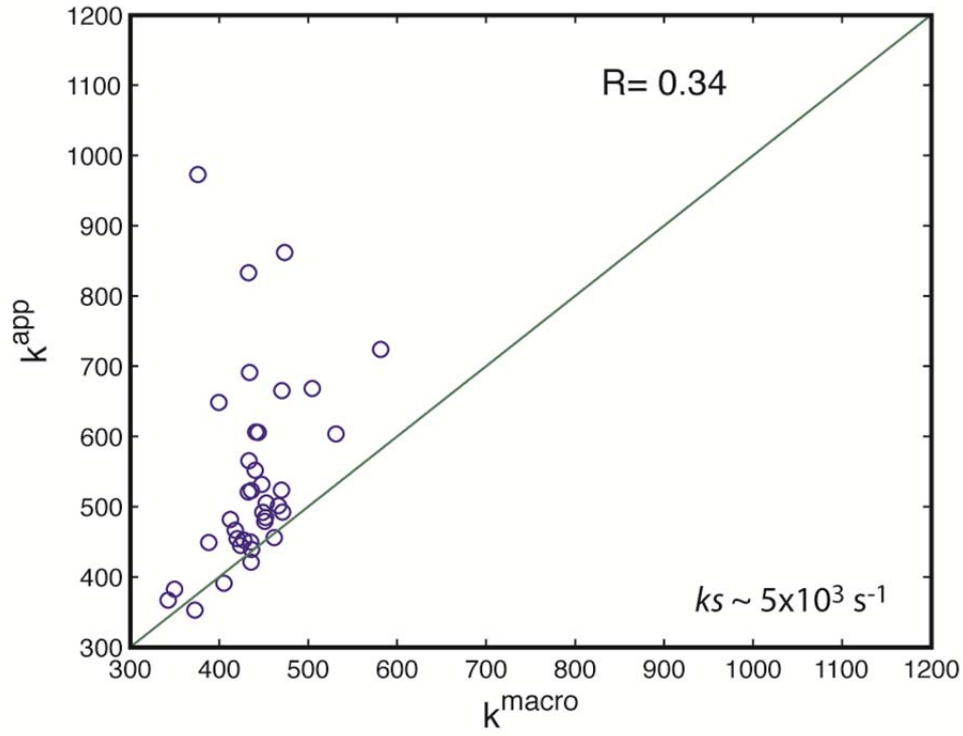

Figure 4.5: Correlation plot of k^{macro} calculated using Eq. 4.10 and the k^{app} obtained from the simulations where the sliding on nonspecific DNA is comparable (< 10^3 s^{-1}) to the pseudo first order rate for intersegment transfer (~10 s^{-1}).

When theoretical and experimental data were compared for relaxation rates for individual protein DNA complexes and mixture of the protein and both DNA molecules, similar trends of the apparent transverse relaxation rates as a function of DNA concentration were found if the sliding process is significantly faster than the direct transfer (Figure 4.6). There is an increased broadening of the NMR line-shape for the complex of protein and mixture of DNA-a and DNA-b, apparent from the increased relaxation values obtained using Lorentzian line shape analysis, which is due to the contribution by the intersegment transfer of the protein between the DNA molecules. Since the pseudo first order rate constant is varied, in our simulations, as a linear function of free DNA concentrations as described in Eq. 4.12, the apparent rate of translocation

between the two non-specific DNA molecules could vary with the DNA concentrations giving us the second order rate constant for intersegment transfer from its slope. This scenario was tested by performing 50 different simulations for different values of sliding rates, slow ($<10^3$ s^{-1}) and fast ($>10^6$ s^{-1}), where the accuracy of the second order intersegment transfer rate obtained using simulation of NMR experiments was checked. In addition to this, the each simulation had different values of free energies and chemical shifts associated with each site on the DNA. A correlation plot of the estimated and the apparent second order rate constant for the two different scenarios (figure 4.7) indicate that when rate of sliding and intersegment transfer are comparable the second order rate constant obtained from the NMR experiments could be inaccurate and conversely when the sliding rates are much faster than the pseudo first-order intersegment transfer rates, the slopes from the DNA dependence experiments give a very good estimate of the second order rate associated with intersegment transfer.

When the sliding process is fast enough, as shown in figure 4.6(A), the apparent transverse relaxation of individual protein-DNA complexes should remain unchanged and is clearly seen in these simulations. But when the sliding rates are comparable to the pseudo first-order intersegment rate constant, evident from Figure 4.6(B), we observe a dependence of apparent transverse relaxation for the individual protein-DNA complexes on the free DNA concentration. From the figure 4.6, corresponding to the intersegment transfer to be the predominant mechanism, simulations on the trend of apparent relaxation rates on free DNA concentrations provides a qualitative idea on the time scale of sliding process, i.e. sliding process is considerably faster than intersegment transfer process. Our simulations indicate that if the time scale of sliding between adjacent sites is as slow as millisecond, there is a stronger dependence of transverse relaxation of the individual protein-DNA complexes on the DNA concentration which is actually not

observed in the published (6) experimental data. These results suggest that the NMR based experiments with different DNA concentrations can give qualitative information on the timescale of sliding. The experimental data shown (6) in figure 4.6(C) imply that the sliding of HoxD9 homeodomain occurs in the time scale much faster than milliseconds.

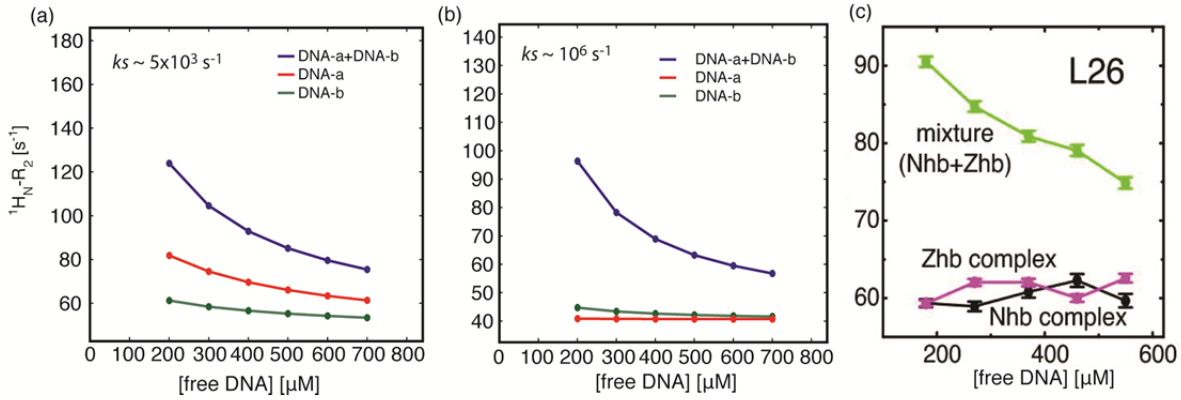

Figure 4.6: Apparent transverse relaxation of individual protein and DNA-a complex, protein and DNA-b complex, and mixture of protein, DNA-a and DNA-b complex as a function of DNA concentration. A sample calculation when sliding rate is (A) comparably faster than the pseudo first-order intersegment transfer rate, (B) comparable to the pseudo first-order intersegment transfer rate. (C) Apparent transverse relaxation of residue L26 of HoxD9 homeodomain measured using the mixture approach to observe its translocation between two nonspecific DNA molecules (6).

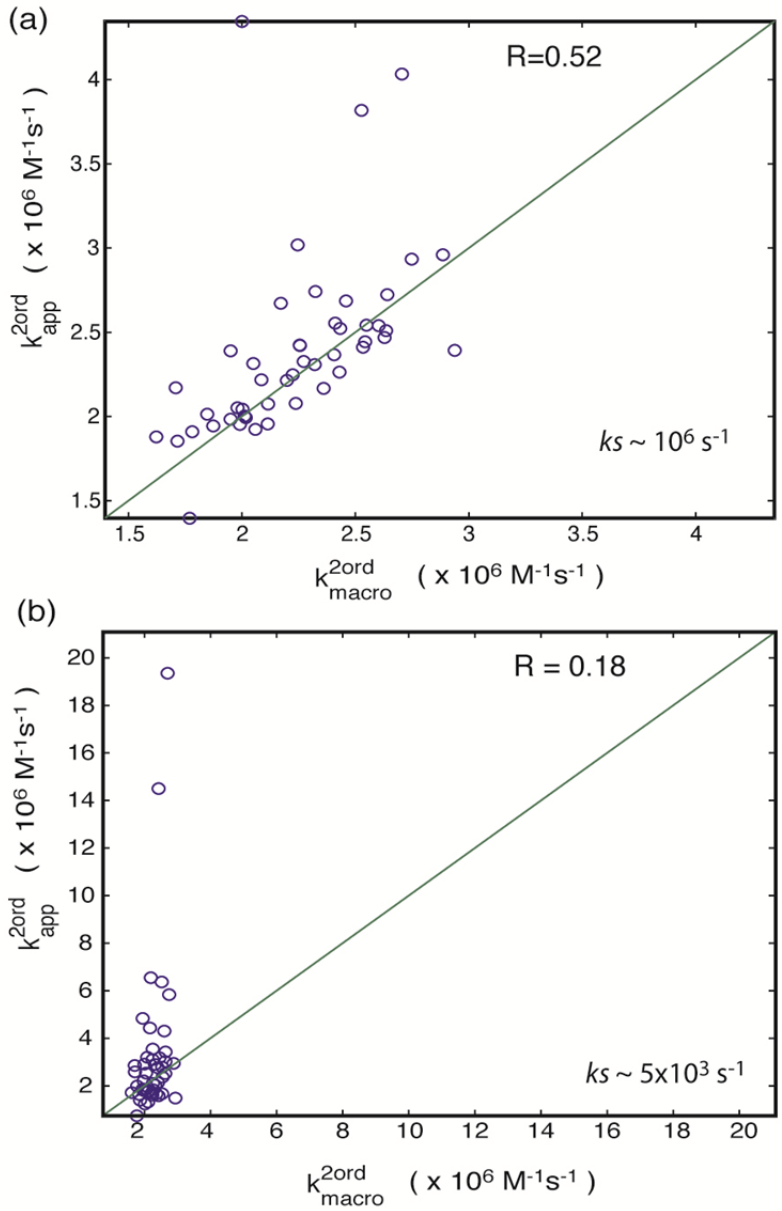

Figure 4.7: Correlation plot of 50 simulations where the DNA concentrations are varied to reveal the apparent second order rate constant for the inter-conversion between two nonspecific DNA molecules. (a) Correlation between calculated second order rate constant and the apparent rate obtained via Reuben and Fiat approximation when the sliding rate are faster(>10^6 s^{-1}) than the pseudo first-order rate constant for intersegment transfer. (b) Correlation between the calculated second order rate constant and the apparent rate obtained via Reuben and Fiat approximation when the sliding rate are comparable(<10^3 s^{-1}) as compared to pseudo first-order rate constant for intersegment transfer

Assessment of NMR-based determination of rate constants for translocation between specific DNA molecules

The intent of using the mixture approach and repeating the experiment at different DNA concentrations is to measure the microscopic translocation rate constant between the two nonspecific sites. The mixture approach gives us the macroscopic rate constant for translocation, which has contributions from sliding, hopping and dissociation-reassociation events to the nonspecific sites. The contribution of the state at which the protein is bound to the nonspecific sites might be significant and could cause a deviation from linearity of k^{app} values obtained from performing experiments at different DNA concentrations. If the affinity towards the nonspecific sites is comparable to that of specific sites, the protein might spend less time on cognate sites and slide away to adjacent nonspecific sites.

Interestingly, the simulations of z-exchange experiments done at different DNA concentrations yielded us pseudo first order rate constants $k_{ab}{}^{app}$ and $k_{ba}{}^{app}$, and when plotted against DNA concentration, it gives us the second order rate constants ($k^{it}{}_{ab}$ and $k^{it}{}_{ba}$) for the translocation between two cognate DNA molecules. These simulations were done using the McConnell equation for longitudinal magnetization as described in the section on materials and methods. The number of sites on each DNA used in these simulations were 40, since in the experimental method (12) n_{DNA} is 24, r_{sp} is 5 and the number of sites on DNA is given by $2*(n_{DNA} - r_{sp} + 1)$. Many simulations were performed for different free energies for nonspecific sites and the apparent second order rate-constant was obtained. These simulations were done at various values of sliding and hopping rates, to obtain the macroscopic rate constants for the protein translocation between the two DNA molecules. There is a larger deviation of the apparent rate constant

obtained from the simulation from the calculated macroscopic rate constants, when the sliding rates are comparable to the pseudo first-order intersegment transfer rate constant. From the simulations we learn that when the affinities towards the cognate sites are different from nonspecific sites by 4 kcal/mol, the protein spends most of the time bound to specific sites and the apparent pseudo first-order rate constant is given by $n\left(\sum_{i=1}^{n}\left(k_h^i\right)^{-1}\right)^{-1}$ where the major contribution is obtained from intersegment transfer between the two cognate sites. It is a reasonable assumption that the difference in free energies between the cognate and nonspecific sites is 4 kcal/mol, which is evident from affinities of HoxD9 towards cognate and nonspecific DNA (6, 96). Under the conditions above, it is possible that the sliding rates could be comparable to that of pseudo first-order intersegment transfer rates for larger protein systems; the apparent rate constant measured using this NMR based method will be predominantly dominated by intersegment transfer between the cognate sites due to the slow sliding rates. This was verified using 50 separate calculations to simulate the NMR based approach (12) where macromolecular concentrations and affinities were kept fixed and the free energies of non-specific sites were varied, to yield the apparent pseudo first order rate constant to be $1829 \pm 7\ s^{-1}$ and the macroscopic rate constant calculated from the above description was $1863 \pm 2\ s^{-1}$. On the other hand when the sliding rates is higher than pseudo first-order intersegment transfer rates, which are fixed by the transition barriers, the macroscopic rate constant derived from Eq. 4.10 and the apparent rate constant obtained from these simulations are comparable. This was again verified using similar calculations as above to yield the apparent pseudo first order rate constant as $(2.93 \pm 0.01) \times 10^4\ s^{-1}$ in excellent agreement with the macroscopic rate constant calculated from Eq 4.10 was $(2.93 \pm 0.01) \times 10^4\ s^{-1}$. he results of one of the sample calculations is shown in the figure 4.8 where the

apparent pseudo first-order rate constants were obtained as a function of the transition state describing the sliding process. So when the transition state has low free energy values, the sliding kinetic rate constant is large, i.e. the protein slides very rapidly and for large values of transition states the sliding rate constant is small, i.e. the protein slides slowly.

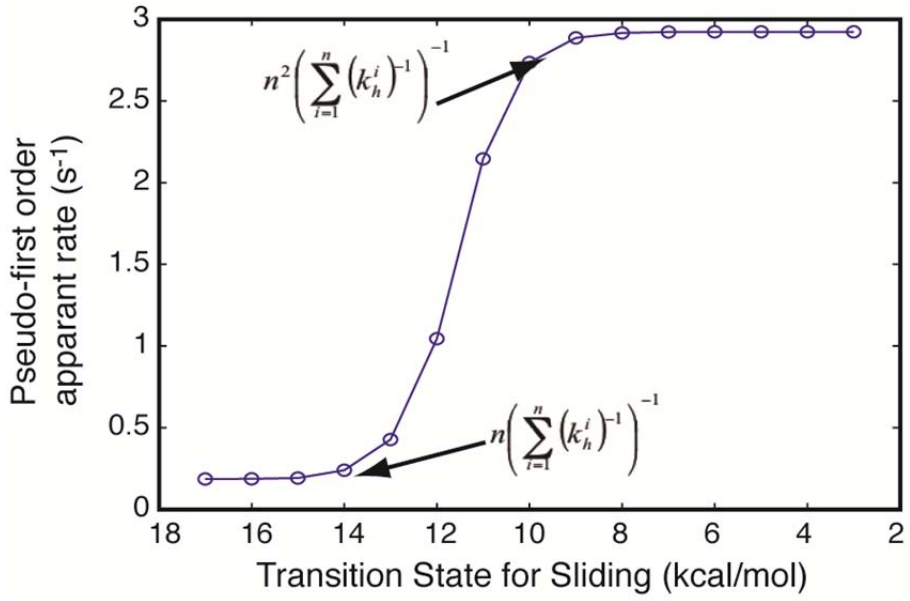

Figure 4.8: Apparent pseudo first-order rate constant obtained from a sample simulation as a function of transition state for sliding. The initial few points depict the condition when the sliding rate constants are small (k_s <10^3s^{-1}) and the last few points describe the condition where the sliding rates are large (k_s >10^6s^{-1}).

When we compare our trend of the macroscopic rate constant as a function of DNA concentration with that obtained from experimental data on translocation of HoxD9 between cognate DNA molecules (96), the plots as seen in Figure 4.9 show a linear trend as a function of DNA concentration. It is very interesting to note from the simulations that for greater sliding rates (for our case k_s >10^6s^{-1}) as compared to pseudo first-order intersegment rates, there is minimal contribution by the sliding process on the apparent

macroscopic rate constant measured using the z-exchange experiment. But when they become comparable there is a significant contribution to the measured apparent rate constant by the sliding process. Thus the macroscopic rate constant measured for HoxD9 homeodomain translocation (96) between two cognate sites is quite accurate for direct transfer mechanism assuming that the sliding process is much faster than the intersegment transfer process.

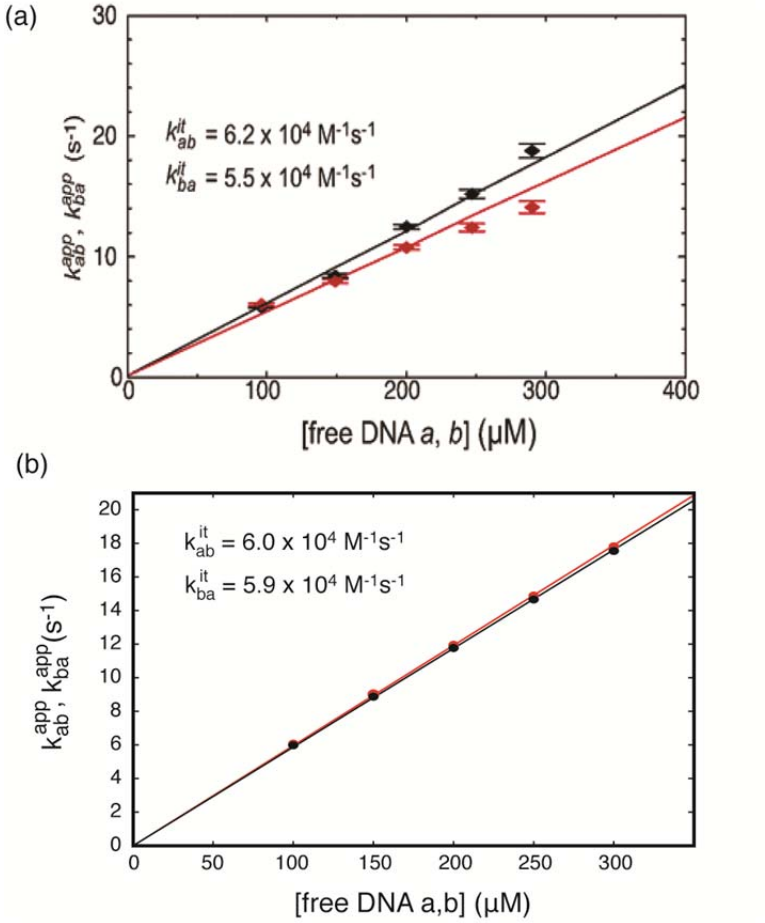

Figure 4.9: Results from NMR experiments and DNA concentration dependence simulations (A) The dependence of apparent rate constant of HoxD9 translocating between two cognate DNA molecules on free DNA concentration (5). The apparent rate constant was measured using z-exchange experiment and the amount of free DNA was varied. (B) The result from a sample simulation where $k_s > k_h$. Here the apparent second order rate constant obtained using NMR based methods is very similar to the actual value (5.85 *10^4 $M^{-1}s^{-1}$).

The kinetic rate constants measured as a function of DNA concentrations as described above, also give us some information on the dissociation process. The intercept of the plot of apparent rate constant as a function of free cognate DNA concentration is present because of the dissociation process that is governed by k_{off}. This intercept has been shown to be present in translocation of Oct1-HoxB1 between two cognate DNA molecules and interpreted as being $0.5*k_{off}$ (104). As the intercepts from the plot of apparent kinetic rate constant as a function of free DNA concentration can be obtained from the previous simulations, these intercepts were analyzed for different sliding rates. The trend of the intercept is shown as a function of transition free energy for sliding process in figure 4.10. When the sliding rates are slow (k_s <10^3 s^{-1}), the intercept actually reflects $0.5*k_{off}$. On the other hand when the sliding rates are fast (k_s >10^6 s^{-1}), the

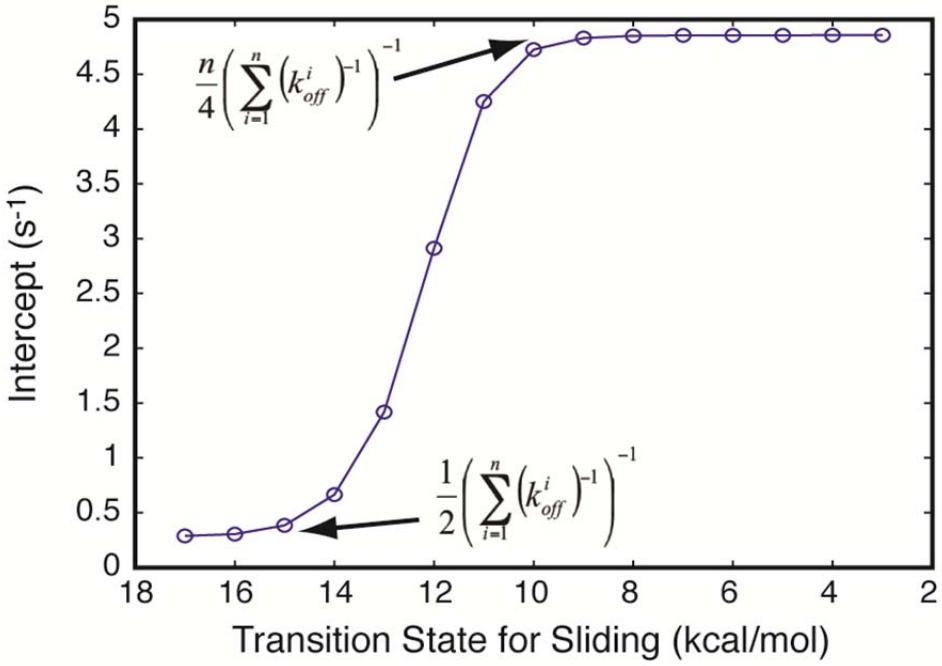

Figure 4.10: Intercept of the plot of apparent rate as a function of DNA concentrations as a function of transition state for sliding process. When the sliding rate is fast the intercept is dependent on the number of sites on the DNA.

intercept was analytically found to be (n/4)*k_{off}. Since the length of DNA (n_{DNA}) can be easily altered in the mixture approach, these NMR experiments can be repeated for different DNA lengths to get intercepts for apparent translocation rate as a function of free DNA concentration. As per our calculations, if the intercept changes with different DNA lengths we can estimate the time scale of the sliding process qualitatively.

These simulations also give us an idea about designing new experiments that can provide us an insight into the translocation mechanism of protein sliding. Since the contribution by the sliding rate to the translocation rate constant is a linear function of the number of non-specific sites on the DNA, we can get some insights into the contribution of sliding process. It is also possible for us to measure the macroscopic rate constants for systems where the DNA length is increased, to increase the number of nonspecific sites flanking the cognate sites. The increase in the number of nonspecific sites would contribute to increase in the macroscopic rate constant and thus a plot of macroscopic rate constant as a function of number of nonspecific sites could give a good understanding of the contribution due to sliding on DNA.

Potential applications to other biophysical problems

These simulations representing the generalized N-state system comprised of protein translocation on DNA molecules are defined by free energies of initial states, transition barriers, and final states. The use of kinetic matrices was possible because the kinetic mechanisms are pseudo first order or first order reactions. This generalized N-state system can be easily extended to other macromolecular processes like protein folding, unfolding, redox reactions, conformation dynamics, etc since the kinetic values can be defined based on free energies of different states. For other biophysical processes

and events, the suitable kinetic matrices relevant to the case in hand can be used for employing this formalism.

CONCLUSIONS

In these calculations, we have been able to develop a generalized N-state model to describe the predominant translocation modes during facilitated target location process. Its simplistic approach has made it possible to extend it to design suitable experiments and study other macromolecular events. The use of Kramers' theory to describe the kinetic rate constants was very advantageous in maintaining the detailed balance of the kinetic matrices. The NMR experiments were simulated by employing the McConnell equation that contains the kinetic matrix, which is constituted of three major translocation kinetic mechanisms present during target location by DNA binding proteins. These simulations most importantly reveal that the NMR based approaches used to obtain the macroscopic rate constants are pretty accurate when the sliding rates are much faster than pseudo first-order intersegment transfer rates. Furthermore we also obtain a qualitative idea on the time scales of different kinetic process governing the facilitated target location. These calculations also reveal the limits of the NMR based approaches and the reliability of the kinetic rate constants obtained.

These simulations have also given us ideas on designing new experiments to decipher the role of sliding events on nonspecific sites on DNA in the measured apparent translocation rate constants. Moreover these simulations can also be tweaked such that it reveals the validity range of NMR based mixture approaches. Thus this N-state generalized theoretical approach is very useful in validating the accuracy of the translocation rate constants obtained via powerful NMR based techniques.

CHAPTER V†

REDOX KINETICS FOR HMGB1 A-DOMAIN

INTRODUCTION

As shown in the previous chapters, we currently possess various NMR based tools to quantitatively analyze the kinetics of various biological processes. In certain cases we had to develop new methods and employ innovative concepts to analyze process happening in varying time scales. A very general N-state model was developed in the previous chapter to analyze the feasibility of various NMR based approaches and validity range of these precise biophysical methods. Since the NMR methods developed can give very accurate details of kinetic processes involving biomolecules, we here employ some of these powerful NMR based methods to understand the kinetics of reduction/oxidation of a biologically important DNA binding protein HMGB1.

Even though the primary function of HMGB1 is in the nucleus as a DNA binding protein, it also has a completely different function as a cytokine in extracellular environments. In the nucleus, this protein binds DNA in a non-specific manner and induces substantial distortion of DNA, which is involved in gene regulation (105). HMGB1 also plays several important roles as a cytokine; it is passively and actively released into the extracellular environment and interacts with receptors such as RAGE, TLR2, TLR4, and TLR9 to induce various cellular responses (106).

† This chapter is adapted from the paper published in FEBS Letters, (ref. 104) which allows authors to reproduce figures, tables or brief quotations from the text of articles published in FEBS Letters for non-commercial purposes. (Source: http://www.febsletters.org/content/authorinfo#copy)

Redox chemistry may be involved in regulation of the function HMGB1. This speculation has been reinforced further by a very recent work on the different redox states of HMGB1, that have revealed different affinities towards DNA (108). This protein contains three cysteine residues: Cys22 and Cys44 in the A-domain, and Cys105 in the B-domain. In the crystal structure of the HMGB1 A-domain•DNA complex (109), Cys22 and Cys44 are in close proximity with only 4.3 Å between the two Sγ atoms (Figure 5.1A), suggesting that a conformational change may allow them to form an intramolecular disulfide bond. Indeed, it has been found that Cys22 and Cys44 can form an intra-molecular disulfide bond within the A-domain, while Cys105 is redox-inactive and remains reduced (110).

Since the HMGB1 protein plays important roles in both reductive (nuclear) and oxidative (extracellular) environments, quantitative characterization of the redox reactions involving the Cys22–Cys44 pair is essential for understanding the functions of HMGB1. Since glutathione and thioredoxin are the predominant redox regulatory machineries inside the cells (53, 54), the kinetics of reduction/oxidation of HMGB1 by these redox machineries are quantitatively measured using NMR based methods. In this chapter, we analyzed the redox properties of the Cys22-Cys44 pair in the HMGB1 A-domain in terms of kinetics and thermodynamics.

MATERIALS AND METHODS

Preparation of the HMGB1 A-domain

The ^{15}N- or ^{13}C/^{15}N-labeled HMGB1 A-domain (human HMGB1 residues 1-84)

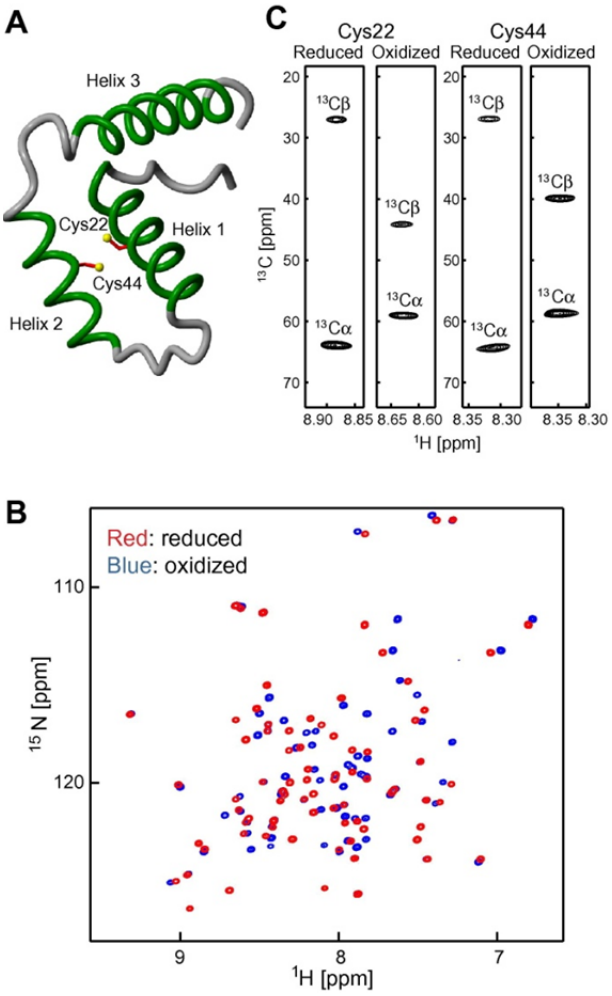

Figure 5.1: (A) Location of Cys22 and Cys44 in the crystal structure of the HMGB1 A-domain (PDB code 1CKT). (B) ^{1}H-^{15}N HSQC spectra recorded on oxidized and reduced forms of the HMGB1 A-domain at pH 5.5 and 20 ˚C. (C) Strips of CBCA(CO)NH spectra showing ^{13}Cβ/^{13}Cα resonances for Cys22 and Cys44. The observed ^{13}Cβ chemical shifts confirm the redox states of these cysteine residues (110).

was expressed in *E. coli* as described previously (112). The protein was purified by ammonium sulfate fractionation and phenyl-FF hydrophobic chromatography (50 mM Tris•HCl [pH 7.5], 1mM DTT, and 2000-0 mM ammonium sulfate). The protein solution

was then dialyzed against a buffer of 20 mM potassium phosphate (pH 6.0), and 100 mM NaCl. The intramolecular Cys22-Cys44 disulfide bond was spontaneously formed at this step. The oxidized A-domain was further purified by Mono-S cation exchange chromatography using 100-650 mM NaCl gradient in 20 mM potassium phosphate (pH 6.0), and Superdex-75 gel-filtration with a buffer of 20 mM Tris•HCl (pH 8.0) and 200 mM NaCl. The completely reduced form of the protein was obtained by adding 10 mM DTT.

NMR assignment

$^{1}H/^{13}C/^{15}N$ resonances for oxidized and reduced HMGB1 A-domains were assigned with 3D HNCA, HN(CO)CA, HNCO, HNCACB, CBCA(CO)NH, C(CO)NH spectra (113) recorded at 20 ˚C on 1.0 mM $^{13}C/^{15}N$-lableled proteins in a buffer containing 20 mM potassium phosphate (pH 5.5), 100 mM KCl, and 7% D_2O. For the reduced protein, 10 mM DTT was also present in the solution. The NMR data was processed and analyzed with NMRPipe and NMRView software.

Redox analysis of the A-domain in the glutathione system

All real-time kinetic experiments were carried out using a Varian 750-MHz NMR system. The reaction kinetics between ^{15}N-labeled HMGB1 A-domain and glutathione (reduced, GSH; oxidized, GSSG) were analyzed as follows. Initially, all of the A-domain was in the oxidized form and dissolved in a buffer of 40 mM Tris•HCl (pH 7.4), 2 mM GSSG, 100 mM KCl, and 7% D_2O. The reaction was initiated by mixing a solution of 100 mM GSH dissolved in the same buffer (pH was carefully adjusted) with the protein

solution. The initial concentrations of GSH, GSSG and ^{15}N-lableled A-domain in the reaction mixture (450 μl) were 20 mM, 2 mM and 0.2 mM, respectively. The reaction at 25 ˚C in an NMR tube was monitored with ^{1}H-^{15}N HSQC. The buffer was saturated with argon gas, which was also sealed in the NMR tube. Seven pairs of HSQC signals from backbone amides of K28, K29, K49, G57, and F59 and the side-chain NH_2 of N36, which are well isolated in spectra for both states, were used to determine populations of the oxidized and reduced forms. This kinetic experiment was performed four times. Kinetic rate constants and the standard redox potential for the A-domain (E_A^0) were determined as described in a later section. The standard redox potential E_G^0=-264 mV for glutathione at pH 7.4 and 25 ˚C (114) was used as a reference. The value of E_A^0 was also obtained from the equilibrium populations of the oxidized and reduced A-domain at six different $[GSH]^2/[GSSG]$ ratios using the Nernst equation.

Kinetic analysis of the reaction between A-domain and thioredoxin

The kinetics of the reaction between the HMGB1 A-domain and thioredoxin was analyzed at 25˚C using 500-μl solutions containing 0.2 mM ^{15}N-lableled A-domain, 50 mM Tris•HCl (pH7.4), 100 mM KCl, 7% D_2O, 0.19 μM rat thioredoxin reductase (Sigma-Aldrich), 0.5 mM NADPH, and human thioredoxin (Sigma-Aldrich) at four different concentrations (2.8, 5.7, 11.4 and 14.3 μM) under argon gas. A control experiment was also carried out in the absence of thioredoxin. Starting from the 100% oxidized state for the A-domain, the reactions were monitored with ^{1}H-^{15}N HSQC spectra until the reduced form predominated. Populations of reduced and oxidized A-domains at

each time point were calculated as described above. Kinetic rate constants were obtained by non-linear least-squares fitting.

CD analysis of thermal stability

CD experiments were performed on the oxidized and reduced A-domain with a JASCO J-720 spectropolarimeter. CD at 222nm were measured at temperatures from 10 to 90 °C, changing at a rate of 1 °C/min. Measured samples were 10 μM proteins in 40 mM potassium phosphate (pH 7.4) and 100 mM KCl. For the reduced form, 2 mM DTT was also present in the solution. Melting temperatures (T_m) and other thermodynamic parameters were calculated as described in literature (115).

RESULTS

NMR of reduced and oxidized HMGB1 A-domain

The reduced and oxidized forms of the HMGB1 A-domain exhibited quite different ^{1}H-^{15}N HSQC spectra, as shown in Figure 5.1B. The redox reactions of the Cys22-Cys44 pair were reversible, and the oxidized A-domain could be completely reduced with 10 mM DTT. Using $^{1}H/^{13}C/^{15}N$ triple resonance NMR spectroscopy, we assigned HSQC signals for both the reduced and oxidized forms individually. Formation of the intra-molecular Cys22-Cys44 disulfide bond in the oxidized form was confirmed by mass spectrometry (data not shown) and $^{13}C\beta$ chemical shifts of the two Cys residues (Figure 5.1C).

Redox reactions of HMGB1 A-domain with glutathiones

Reduced (GSH) and oxidized (GSSG) glutathione molecules are the major contributors to the redox environment in cells (114). Redox reactions between the HMGB1 A-domain and glutathione are represented by:

$$A_{oxi} + 2GSH \underset{k_2}{\overset{k_1}{\rightleftharpoons}} A_{red} + GSSG\ , \qquad (5.1)$$

where k_1 and k_2 are the kinetic rate constants for the forward and backward reactions, respectively. The time courses of the reactions are given by:

$$\frac{d}{dt}[A_{oxi}] = -k_1[A_{oxi}][GSH]^2 + k_2[A_{red}][GSSG], \qquad (5.2)$$

$$\frac{d}{dt}[A_{red}] = k_1[A_{oxi}][GSH]^2 - k_2[A_{red}][GSSG]. \qquad (5.3)$$

Under conditions where [GSH] and [GSSG] are much greater than [A] and virtually constant throughout the reactions, a pseudo-first-order approximation can be applied to the rate equations above, so their solutions for the reaction starting from 100% oxidized state become:

$$[A_{oxi}](t) = [A_{oxi}](0)\left[\frac{k_1'}{k_1' + k_2'}\exp\{-(k_1' + k_2')t\} + \frac{k_2'}{k_1' + k_2'}\right], \qquad (5.4)$$

$$[A_{red}](t) = [A_{oxi}](0)\frac{k_1'}{k_1' + k_2'}\left[1 - \exp\{-(k_1' + k_2')t\}\right], \qquad (5.5)$$

where k_1' and k_2' are pseudo-first-order rate constants given by $k_1[GSH]^2$ and $k_2[GSSG]$, respectively. Non-linear least–squares fitting against experimental time-course data

provides the kinetic rate constants k_1 and k_2. Based on the Nernst equation, the ratio of k_2 to k_1 is related to the standard redox potentials of the protein and glutathione:

$$\frac{k_2}{k_1} = \exp\left\{-(E_A^0 - E_G^0)\frac{2F}{RT}\right\} \quad , \qquad (5.6)$$

in which E_A^0 and E_G^0 are standard redox potentials for the HMGB1 A-domain and glutathione, respectively; F, the Faraday constant; R, gas constant; and T is temperature.

Using NMR, we analyzed kinetics of a reaction with the initial conditions of $[A_{oxi}]$ = 0.20 mM, $[A_{red}]$= 0.00 mM, [GSH] = 20 mM, and [GSSG] = 2 mM, for which the pseudo-first-order approximation is reasonably valid. ^{1}H-^{15}N HSQC spectra were recorded as a function of reaction time. Signals from the oxidized form of ^{15}N-labeled A-domain decreased and those from the reduced form increased, which represents the redox reactions occurring in the NMR sample (Figure 5.2A). The reaction reached equilibrium in 8 hours (Figure 5.2B). Strictly speaking, the reaction represented by Eq. 5.1 can involve intermediates corresponding to the S-glutathionylated proteins. However, the NMR spectra monitoring the reaction showed two distinct sets of signals from the oxidized and reduced forms of the protein (Figure 5.2A), whereas signals that may arise from the intermediate were not clearly observed. This means that the population of the intermediate is too low at least under the present experimental conditions, which justifies the analysis with the two-state model. The kinetic rate constants k_1 and k_2 were determined to be 0.51 ± 0.18 $M^{-2}s^{-1}$ and 0.06 ± 0.01 $M^{-1}s^{-1}$, respectively.

Interestingly, the final population of the oxidized A-domain in this experiment was as high as 40% despite the presence of 20 mM GSH, which indicates that the

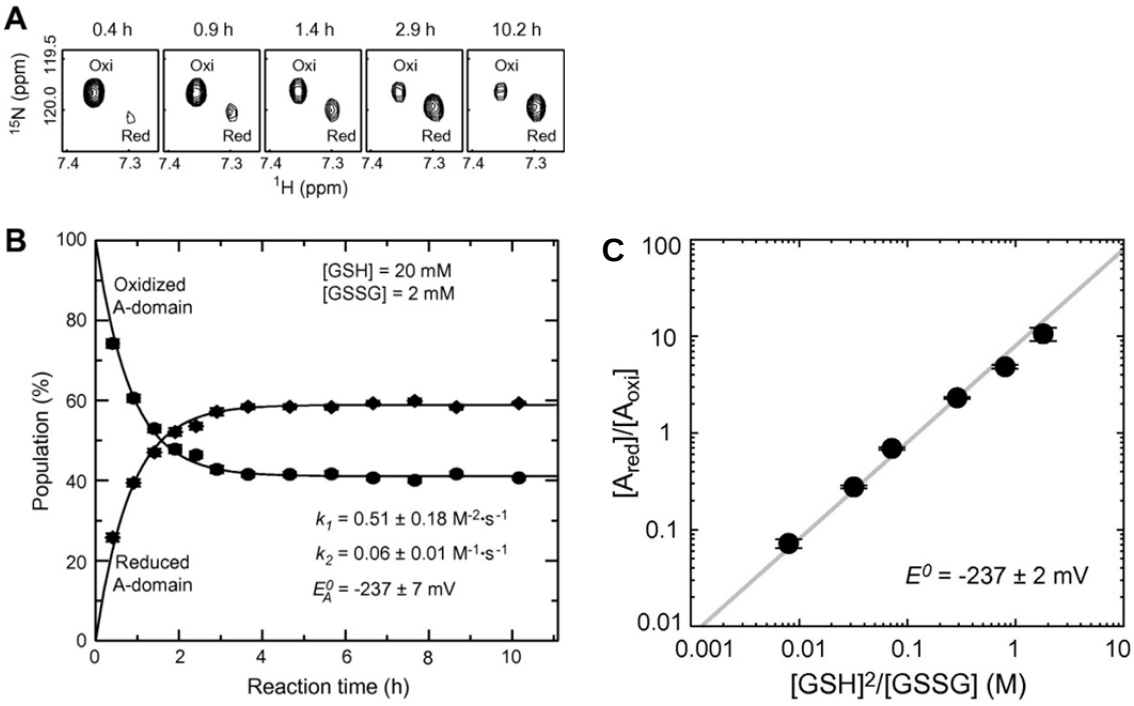

Figure 5.2. (A) Changes of ^{1}H-^{15}N HSQC signals due to the reaction between the ^{15}N-labeled HMGB1 A-domain and glutathiones. Shown signals are from the F59 amide group in the oxidized (Oxi) and reduced (Red) HMGB1 proteins. (B) The reaction time-course obtained from HSQC spectra recorded on the reaction mixture of 20 mM GSH, 2 mM GSSG, and 0.2 mM A-domain dissolved in a buffer of 40 mM Tris•HCl (pH7.4), 100 mM KCl, and 7% D_2O. Average values of k_1, k_2, and E_A^0 determined from four independent experiments are shown together with their standard deviations. (C) Determination of the standard redox potential E_A^0 from the equilibrium populations of the oxidized and reduced HMGB1 proteins at six different [GSH]^2/[GSSG] ratios.

standard redox potential for the Cys22-Cys44 pair is relatively low. Indeed, the value of E_A^0 was determined to be -237 ± 7 mV from k_1 and k_2 together with Eq. 5.6. An independent measurement of E_A^0 from the equilibrium populations at six different [GSH]^2/[GSSG] ratios gave a virtually identical value (-237 ± 2 mV) as shown in Figure 5.2C.

Kinetics of reduction of the Cys22-Cys44 disulfide bond by thioredoxin

Thioredoxin (Trx) is a 12-kDa protein that plays a major role in redox regulation by attacking and reducing oxidized proteins in the nucleus and cytoplasm (53). Thioredoxin is oxidized in the process, but its oxidized form is rapidly reduced by the enzyme thioredoxin reductase (TrxR) with NADPH as the cofactor. Although a previous study has shown that thioredoxin can reduce HMGB1 (110), those data are qualitative rather than quantitative. We thus performed a quantitative analysis of the kinetics of the reaction between thioredoxin and the HMGB1 A-domain.

Reaction kinetics was investigated for solutions containing thioredoxin, thioredoxin reductase, NADPH, and ^{15}N-labeled A-domain (Figure 5.3). The reaction occurring in an NMR tube was monitored by ^{1}H-^{15}N HSQC spectra. In these experiments, thioredoxin indeed reduced the oxidized A-domain, whereas the redox state of the A-domain remained unchanged within the experimental timeframe in the same reaction mixture without thioredoxin (Figure 5.3A). Since the standard redox potentials of the HMGB1 A-domain and thioredoxin are comparable and disulfide exchange is reversible, time-courses for the oxidized and reduced A-domains are given by the following rate equations:

$$\frac{d}{dt}[\mathrm{A_{oxi}}] = -k_{Trx}[\mathrm{A_{oxi}}][\mathrm{Trx_{red}}] + k_{back}[\mathrm{A_{red}}][\mathrm{Trx_{oxi}}] \quad , \qquad (5.7)$$

$$\frac{d}{dt}[\mathrm{A_{red}}] = k_{Trx}[\mathrm{A_{oxi}}][\mathrm{Trx_{red}}] - k_{back}[\mathrm{A_{red}}][\mathrm{Trx_{oxi}}] \quad , \qquad (5.8)$$

where k_{Trx} is the second-order rate constant for the reaction between the oxidized A-domain and the reduced Trx, and k_{back} is that for the backward reaction. Judging from the

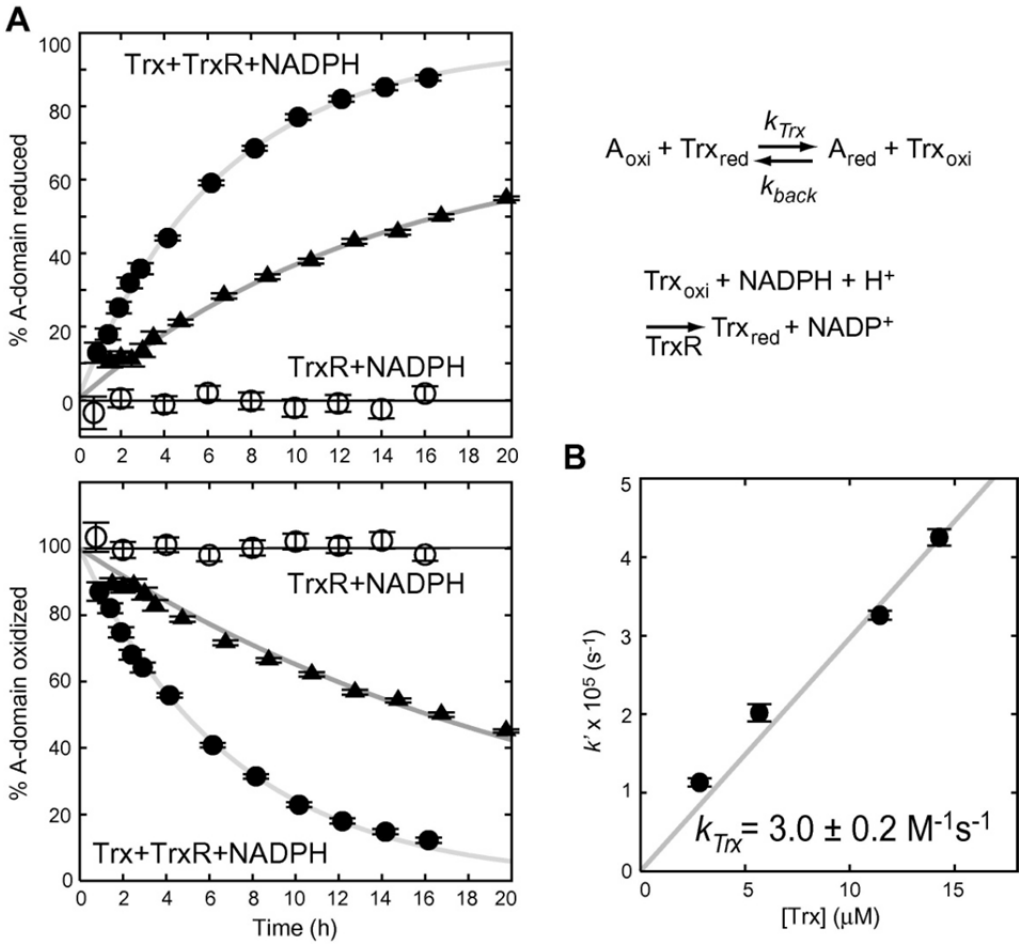

Figure 5.3. Kinetics of the reactions between Trx and the HMGB1 A-domain monitored with NMR. (A) Reaction time-course. Filled and open symbols represent the reactions in presence (circle, 14 μM; triangle, 6 μM) and absence of Trx, respectively. (B) Plot of pseudo-first-order rate constants measured at four different Trx concentrations.

enzymatic constants K_m and k_{cat} for rat thioredoxin reductase (116), the reduction of Trx by thioredoxin reductase is much faster than the reaction between Trx and the HMGB1 A-domain under the present experimental conditions. Hence contributions of the second terms with $[Trx_{oxi}]$ in Eqs 5.7 and 5.8 are negligible, and $[Trx_{red}]$ is virtually constant throughout the reaction. This leads to the following approximate solutions:

$$[A_{oxi}](t) = [A_{oxi}](0)\exp(-k't) \quad , \quad (5.9)$$

$$[A_{red}](t) = [A_{oxi}](0)\{1-(\exp(-k't)\} \quad , \quad (5.10)$$

in which *k'* is a pseudo first-order rate constant equal to $k_{Trx}[Trx_{red}]$. By using these equations, we obtained values of *k'* at four different Trx concentrations. As shown in Figure 5.3C, *k'* was indeed proportional to the Trx concentration, and k_{Trx} was determined to be 3.0 ± 0.2 $M^{-1}•s^{-1}$.

Thermal stability of the reduced and oxidized A-domain

We investigated the effect of Cys22-Cys44 disulfide bond formation on the thermal stability of the HMGB1 A-domain by using CD. Although CD spectra of the reduced and oxidized A-domains are very similar at 10 ˚C (Figure 5.4A), temperature

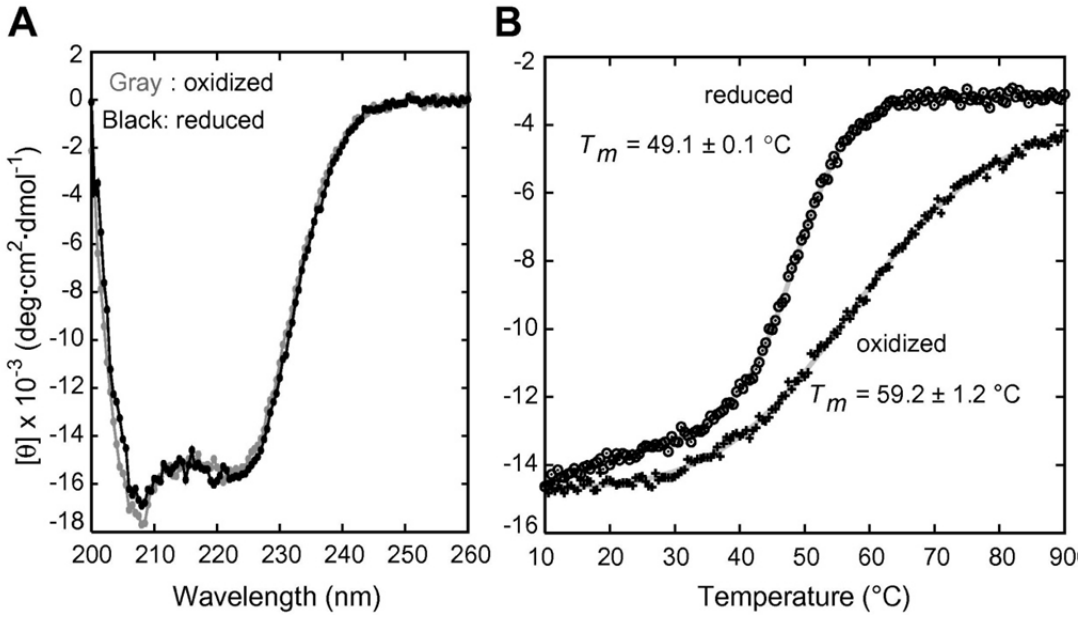

Figure 5.4. (A) CD spectra recorded at 10 ˚C on the oxidized and reduced forms of the HMGB1 A-domain. (B) CD at 222 nm as a function of temperature measured on the oxidized (+) and reduced (O) A-domains.

scans reveal that their thermodynamic properties are quite different (Figure 5.4B). Values of the melting temperature (T_m) for the reduced and oxidized forms were determined to be 49.1 ± 0.1 ºC and 59.2 ± 1.2 ºC, respectively. Thus, formation of the Cys22-Cys44 disulfide bond stabilizes the protein, increasing the T_m by 10 ˚C. We also determined the differences in enthalpy (ΔH_{ref}) and entropy (ΔS_{ref}) between the folded and unfolded states at T_m; their values were ΔH_{ref} = 18.8 ± 1.0 kcal•mol^{-1} and ΔS_{ref} = 56 ± 3 cal•K^{-1}•mol^{-1} for the oxidized A-domain and ΔH_{ref} = 49.8 ± 1.0 kcal/mol and ΔS_{ref} = 155 ± 3 cal•K^{-1}•mol^{-1} for the reduced form. The substantially smaller ΔS_{ref} for the oxidized form can be attributed to conformational restrictions by the disulfide bond on the unfolded state (117).

DISCUSSION

Populations of the oxidized and reduced HMGB1 proteins in various environments

Based on obtained k_1 and k_2 rate constants and standard redox potential E_A^0, we can estimate redox reactions of the HMGB1 proteins in various environments. The HMGB1 protein is present in nucleus, cytoplasm, and extracellular space (105, 106). The redox potential for GSH/GSSG in the cytoplasm is known to be ~-240 mV (114). For this redox environment, the equilibrium population of the oxidized HMGB1 is estimated to be as high as ~44%, based on the Nernst equation. Although the GSH concentration in nucleus is known to be ~8 mM and slightly higher than in cytoplasm (118), the GSH/GSSG redox potential in the nucleus is not available in literature as far as we know. Even though the potential is as low as -260 mV, the population of the oxidized HMGB1 is estimated to be ~14%. But the actual population of the oxidized form is likely to be

much lower than this estimate because reduction of the A-domain by thioredoxin is considerably faster than the oxidation by GSSG under physiological conditions, as mentioned in the next section. In fact, previous studies have found that the nuclear HMGB1 molecules are mostly in the reduced form (110, 119).

The redox potential in the endoplasmic reticulum (ER), where most extracellular proteins become oxidized before secretion, is -180 mV (114). If HMGB1 is secreted through the ER, 99% of the released HMGB1 should be in the oxidized form in equilibrium. However, it has been demonstrated that the HMGB1 protein bypasses ER and is secreted to the extracellular environment primarily via a non-classical, vesicle-mediated secretory pathway (120, 121). Unless the vesicle provides an environment as oxidative as the ER, both reduced and oxidized forms may exist when HMGB1 molecules are released to the extracellular space. The released HMGB1 in the reduced form should be a short-lived species because of the oxidative environment. These two forms of HMGB1 might have different roles in extracellular signaling.

It should be mentioned that post-transcriptional modifications such as acetylation of lysine residues could affect the redox properties of the protein, while the recombinant HMGB1 A-domain without the modifications was used in the present study. It is known that the acetylation of HMGB1 determines its relocation to the cytoplasm in cells (120). The lysine acetylation can change the electrostatic potential and affect pK_a for each thiol group of the Cys22-Cys44 pair, and therefore, the redox properties of HMGB1. However, we think that the impact is probably marginal, because no lysine amino group is present

within 8 Å from the thiol groups in the three-dimensional structure of the HMGB1 A-domain.

Thioredoxin's role to maintain the reduced form of HMGB1

The concentration of thioredoxin in eukaryotic cells is believed to be 1-20 μM (53). Our observed value of k_{Trx} (3.0 $M^{-1}•s^{-1}$) for the Cys22-Cys44 disulfide bond suggests that it takes as long as 6 hours for 10 μM thioredoxin to reduce 50% of the oxidized HMGB1. This reduction is far slower than other reducing reactions by thioredoxin. For example, the corresponding k_{Trx} value for the reaction between insulin and thioredoxin is 1.0×10^5 $M^{-1}•s^{-1}$ (122), giving a half-life time of only 0.7 s at the same Trx concentration (10 μM).

However, it seems that the slow reaction between HMGB1 and thioredoxin is still meaningful to keep HMGB1 reduced in cells, because the oxidation of the HMGB1 by GSSG is also slow as shown above. Under physiological conditions with 10 μM thioredoxin and 100 μM GSSG, the pseudo-first-order rate constants k_{Trx}[Trx] and k_2[GSSG] are calculated to be 3.0×10^{-5} s^{-1} and 0.60×10^{-5} s^{-1}, respectively, and thus, reduction by thioredoxin is five times faster. But when the GSSG concentration is over 500 μM due to oxidative stress, the oxidation of HMGB1 by GSSG can be faster than the reduction of HMGB1 by thioredoxin. Thus, the oxidized HMGB1 may accumulate even in cells under oxidative stress (The higher protein stability of the oxidized HMGB1 as demonstrated by the CD experiment might contribute to the accumulation). The increase

of the oxidized form could cause perturbations in various gene-regulations, if the oxidation alters DNA-binding properties of HMGB1.

CONCLUDING REMARKS

We have quantitatively characterized the Cys22-Cys44 pair in the HMGB1 A-domain using biophysical approaches. These two cysteine residues can rapidly form an intramolecular disulfide bond with the standard redox potential as low as -237 mV, which suggests that the cellular glutathione system alone is not enough to keep HMGB1 completely reduced in the cells. Our real-time kinetic data indicate that the reduction of the oxidized HMGB1 by thioredoxin is efficient enough to maintain the high population of the reduced HMGB1, while the reaction is far slower than other reducing reactions by thioredoxin. The kinetic characterization of the redox reactions involving HMGB1 and redox regulatory systems have given us the enough evidence to speculate that the different functions of HMGB1 could be attributed to the different redox states of the Cys22-Cys44 disulphide bonds. The low efficiency of the thioredoxin-HMGB1 reaction together with the protein stabilization by the Cys22-Cys44 disulfide bond might lead to accumulation of the oxidized form of the HMGB1 protein in cells under oxidative stress.

In this chapter, we have not only demonstrated the use of NMR based methods to give information beyond the structure, but also use them to provide biologically relevant information in terms of kinetics. The redox potential of the disulphide bonds present in HMGB1 has made it possible to calculate the relative population of the reduced and oxidized states inside the cells under different concentrations of redox regulatory systems

(107). It is important to note that the extensive kinetic characterization of reduction/oxidation of HMGB1 by redox regulatory systems reveal the speed of these reactions happening under physiological concentrations and conditions, thus leading us to attribute the distinct functions of HMGB1 to its redox environment.

CHAPTER VI

PERSPECTIVE

The overall objective in this thesis was to study the biologically relevant macromolecular interactions in terms of kinetics by systematically developing models and designing novel experiments. The biological function is not necessarily attained by bringing two macromolecules together, but the speed of the interaction play a vital role in the biological activities. While trying to achieve the overall goal in this study we were able to successfully push the envelope of the biophysical characterization of various macromolecular processes occurring in different time scales under physiological conditions.

Protein-DNA interactions are very important for the cells and the process of target location on DNA is done very efficiently by DNA binding proteins. The translocation of protein inside the nucleus could be very complex because of high affinities of DNA-binding proteins towards DNA along with high concentrations of DNA. Somehow these DNA binding proteins overcome these obstacles and carry out their respective functions. In order to address this biological problem we developed the TROSY based z-exchange experiment to study the translocation kinetics of HoxD9 transcription factor between two cognate DNA molecules using the mixture approach. The main idea about the improvements done to the conventional z-exchange experiment was to be able to incorporate the TROSY principle which would yield stronger and sharper NMR signals. The successful development of TROSY based z-exchange method now permits us to analyze larger biological systems undergoing exchange in the slow exchanging regime. In the case of HoxD9 homeodomain translocation between two different cognate DNA molecules, the TROSY based z-exchange permitted to successfully analyze kinetics of

protein translocation on DNA with wider temperature range to reveal the activation energetics. The thermodynamic characterization of slow exchanging systems was previously not feasible due to the narrow working temperature of the conventional z-exchange experiment. In terms of energetics, these aforementioned experiments revealed a very important proof for the existence of intersegment transfer mechanism under high DNA concentrations, while complementing previous studies (6, 12). Since the kinetic rate constants for translocation of DNA binding proteins on DNA can be successfully attained using our mixture approach along with TROSY based z-exchange experiment, a more biologically challenging problem can be tested is 'molecular crowding'. In order to understand the effects of molecular crowding inside the cells, kinetics of protein translocation of DNA can be measured in the presence of molecular crowding agents such as dextran, PEG, Ficoll, etc. Further extensions to this methodology would be to study the role of other cellular proteins that aid and/or inhibit transcription factors from doing their function, which is feasible by adding unlabeled cellular protein along with isotope-labeled transcription factor and unlabeled DNA. Due to increased signal to noise ratio gained using the TROSY based z-exchange experiments, this methodology can be easily adapted to other larger biological systems undergoing slow exchange between a major and minor state, that were previously not analyzable. The improvements over the conventional z-exchange experiments have resulted in the use of TROSY based z-exchange spectroscopy to study larger complex biological systems like Zif268 (123), Oct1-HoxB1 (104), etc.

The theoretical consideration of the NMR studies used to analyze the kinetics of protein translocation between different DNA molecules was necessary because of the presence of various microstates of protein bound to different regions of the DNA. The

NMR methods were successfully simulated using the McConnell equations and Kramers theory, where the kinetic rate constants obtained experimentally are indeed close to the expected rate constants obtained from the N-state model. One of the main implications from these simulations is that we now have a qualitative idea on the time scale of the sliding of protein on DNA. Our theoretical calculations give us some insights in to designing experiments, where the length of the DNA is varied, to test for the effect of sliding process on the apparent intersegment translocation kinetic rates. This would complement previous experiments on translocation of HoxD9 homeodomain between two nonspecific DNA molecules (6) where the simulation reveal the sliding rate constant is much faster than the pseudo first-order intersegment transfer rate constant. Similar simulations on the translocation of protein between cognate DNA molecules also reveal the validity range of these NMR based experiments. The N-state model used here is a generalized approach, i.e. it can be easily extended to be used in describing other biological processes like protein folding, ligand recognition, domain motions etc that access various states while performing their functions.

Another important biological problem was tackled here, in which the kinetics of redox reactions involving high mobility group box 1 (HMGB1) A-domain and redox regulatory machineries like glutathione and thioredoxin systems were analyzed. The redox state of HMGB1 might play a very critical role in its function as it is a bifunctional protein possessing different functions under different redox environments. The redox characterization of HMGB1 was very important step towards a broader understanding of the role of different redox states in different diseases involving HMGB1. One of the major functions of HMGB1 inside the nucleus is to facilitate other transcriptional factor to locate their cognate sites efficiently by bending the DNA upon binding. In the previous

chapter, owing to unique redox potential of HMGB1 along with redox kinetic experiments, it is can be inferred that the cellular redox machineries struggle a lot to keep HMGB1 reduced inside the cells. And under the oxidative stress environments, there could be an accumulation of oxidized HMGB1 inside the cells. These results shed light into the possible role of the different redox states in HMGB1's function. Inspired by our findings on HMGB1's redox properties, further research on the DNA binding activity of HMGB1 reveal that oxidized and reduced states of HMGB1 possess different affinities towards Cisplatin-modified bent DNA (108). It is also important to note that there are other redox machineries present inside the cell other than glutathione and thioredoxin system like nucleoredoxin, glutaredoxin, etc., that could affect the redox states of HMGB1. The real-time NMR based experiments were used to characterize these redox reactions of HMGB1 are not very specific to the biological problem being addressed here and can be extended to study other biologically problems. Like our aforementioned approach to study redox kinetics, careful biophysical experiments on other redox regulatory systems could lead to a better understanding of their role in keeping cellular environment reductive. The other function of HMGB1 is outside the cells where it acts as a cytokine and is mainly involved the cellular necrosis. Since elevated levels of HMGB1 are found in various disease states like cancer, inflammation, sepsis, arthritis, etc, HMGB1 has been a potential drug target to cure these diseases (124). Since it has already been shown that different redox states of HMGB1 possess varied DNA binding activity, it could be speculated that these different redox states of HMGB1 might play a pivotal role in its extracellular activity. The accumulation of oxidized HMGB1 could be detrimental to the cell if the enhanced DNA binding function is attributed to reduced state and the necrotic activity is attributed to the oxidized state. These thoughts need to be

tested by biologists to get a better understanding of the function of omnipresent protein HMGB1.

Overall, the biophysical characterization of various macromolecular processes occurring in different time scales was made possible by the use of various biophysical approaches, mainly NMR based methods. These experiments as well as simulations provide an insight into the different microscopic events that influence the function of these macromolecules. It is now possible to gain access into the different kinetic mechanisms explored by proteins to carry out their cognate function efficiently inside the cells as physiological concentrations of macromolecules can be mimicked in the NMR based methods. These NMR based biophysical experiments provide an advantage over conventional methods, in terms of experimental feasibility and development of assays for the kinetic measurements of macromolecular events.

APPENDIX

Translocation of protein between two DNA molecules with three sites each

The protein binds to three sites on DNA-a and DNA-b with two different orientations, thus there exists six bound states of protein to each DNA molecule. For this system (N=12), the states 1 to 6 describe the protein bound to DNA-a and states 7 to 12 describe the protein bound to DNA-b.

The kinetic matrix to describe the intermolecular translocation process for this system is given by,

$$K_{\text{inter}} = \begin{pmatrix} 11k_h^1 & -k_h^2 & -k_h^3 & -k_h^4 & -k_h^5 & -k_h^6 & -k_h^7 & -k_h^8 & -k_h^9 & -k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & 11k_h^2 & -k_h^3 & -k_h^4 & -k_h^5 & -k_h^6 & -k_h^7 & -k_h^8 & -k_h^9 & -k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & 11k_h^3 & -k_h^4 & -k_h^5 & -k_h^6 & -k_h^7 & -k_h^8 & -k_h^9 & -k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & -k_h^3 & 11k_h^4 & -k_h^5 & -k_h^6 & -k_h^7 & -k_h^8 & -k_h^9 & -k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & -k_h^3 & -k_h^4 & 11k_h^5 & -k_h^6 & -k_h^7 & -k_h^8 & -k_h^9 & -k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & -k_h^3 & -k_h^4 & -k_h^5 & 11k_h^6 & -k_h^7 & -k_h^8 & -k_h^9 & -k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & -k_h^3 & -k_h^4 & -k_h^5 & -k_h^6 & 11k_h^7 & -k_h^8 & -k_h^9 & -k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & -k_h^3 & -k_h^4 & -k_h^5 & -k_h^6 & -k_h^7 & 11k_h^8 & -k_h^9 & -k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & -k_h^3 & -k_h^4 & -k_h^5 & -k_h^6 & -k_h^7 & -k_h^8 & 11k_h^9 & -k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & -k_h^3 & -k_h^4 & -k_h^5 & -k_h^6 & -k_h^7 & -k_h^8 & -k_h^9 & 11k_h^{10} & -k_h^{11} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & -k_h^3 & -k_h^4 & -k_h^5 & -k_h^6 & -k_h^7 & -k_h^8 & -k_h^9 & -k_h^{10} & 11k_h^{10} & -k_h^{12} & 0 \\ -k_h^1 & -k_h^2 & -k_h^3 & -k_h^4 & -k_h^5 & -k_h^6 & -k_h^7 & -k_h^8 & -k_h^9 & -k_h^{10} & -k_h^{11} & 11k_h^{10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The kinetic matrix for intramolecular translocation of protein on DNA molecules is given by,

$$K_{\text{intra}} = \begin{pmatrix} k_s^1 & -k_s^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -k_s^1 & 2k_s^2 & -k_s^3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -k_s^2 & k_s^3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & k_s^4 & -k_s^5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -k_s^4 & 2k_s^5 & -k_s^6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -k_s^5 & k_s^6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & k_s^7 & -k_s^8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -k_s^7 & 2k_s^8 & -k_s^9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -k_s^8 & k_s^9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_s^{10} & -k_s^{11} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -k_s^{10} & 2k_s^{11} & -k_s^{12} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -k_s^{11} & k_s^{12} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The kinetic matrix to describe the process of protein disassociation from DNA molecule and re-associating with DNA molecules is given by,

$$K_{\text{free}} = \begin{pmatrix} k_{off}^1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -k_{on} \\ 0 & k_{off}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -k_{on} \\ 0 & 0 & k_{off}^3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -k_{on} \\ 0 & 0 & 0 & k_{off}^4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -k_{on} \\ 0 & 0 & 0 & 0 & k_{off}^5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -k_{on} \\ 0 & 0 & 0 & 0 & 0 & k_{off}^6 & 0 & 0 & 0 & 0 & 0 & 0 & -k_{on} \\ 0 & 0 & 0 & 0 & 0 & 0 & k_{off}^7 & 0 & 0 & 0 & 0 & 0 & -k_{on} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_{off}^8 & 0 & 0 & 0 & 0 & -k_{on} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_{off}^9 & 0 & 0 & 0 & -k_{on} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_{off}^{10} & 0 & 0 & -k_{on} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_{off}^{11} & 0 & -k_{on} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_{off}^{12} & -k_{on} \\ -k_{off}^1 & -k_{off}^2 & -k_{off}^3 & -k_{off}^4 & -k_{off}^5 & -k_{off}^6 & -k_{off}^7 & -k_{off}^8 & -k_{off}^9 & -k_{off}^{10} & -k_{off}^{11} & -k_{off}^{12} & 12k_{on} \end{pmatrix}$$

REFERENCES

1. Ellis, R. J., and A. P. Minton. 2003. Cell biology: Join the crowd. Nature 425:27-28.

2. Minton, A. P. 2001. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. The Journal of Biological Chemistry 276:10577-10580.

3. Nath, D., and S. Shadan. 2009. The ubiquitin system. Nature 458:421-421.

4. Bonifacino, J. S., E. C. Dell'Angelica, and T. A. Springer. 2001. Immunoprecipitation. Current Protocols in Molecular Biology 10.16.11-10.16.29.

5. Brymora, A., V. A. Valova, and P. J. Robinson. 2001. Protein-Protein Interactions Identified by Pull-Down Experiments and Mass Spectrometry. Current Protocols in Cell Biology 17.5.1–17.5.51.

6. Iwahara, J., M. Zweckstetter, and G. M. Clore. 2006. NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. Proceedings of the National Academy of Sciences U S A 103:15062-15067.

7. Anderson, B. J., C. Larkin, K. Guja, J. F. Schildbach, and L. J. Ludwig Brand and Michael. 2008. Chapter 12 Using Fluorophore-Labeled Oligonucleotides to Measure Affinities of Protein-DNA Interactions. In Methods in Enzymology. Academic Press, San Diego, CA. 253-272.

8. Gorman, J., and E. C. Greene. 2008. Visualizing one-dimensional diffusion of proteins along DNA. Nature Structural & Molecular Biology 15:768-774.

9. Tafvizi, A., F. Huang, A. R. Fersht, L. A. Mirny, and A. M. van Oijen. 2011. A single-molecule characterization of p53 search on DNA. Proceedings of the National Academy of Sciences U S A 108:563-568.

10. Wang, F., and E. C. Greene. 2011. Single-Molecule Studies of Transcription: From One RNA Polymerase at a Time to the Gene Expression Profile of a Cell. Journal of Molecular Biology In Press, Corrected Proof.

11. Cavanagh, J., J. W. Fairbrother, G. A. Palmer, M. Rance, and J. N. Skelton. 2007. Protein NMR Spectroscopy: Principles and Practice. Elsevier Academic Press, San Diego, CA.

12. Iwahara, J., and G. M. Clore. 2006. Direct Observation of Enhanced Translocation of a Homeodomain between DNA Cognate Sites by NMR Exchange Spectroscopy. Journal of American Chemical Society 128:404-405.

13. Iwahara, J., and G. M. Clore. 2006. Detecting transient intermediates in macromolecular binding by paramagnetic NMR. Nature 440:1227-1230.

14. Iwahara, J., Y. S. Jung, and G. M. Clore. 2007. Heteronuclear NMR Spectroscopy for Lysine NH_3 Groups in Proteins: Unique Effect of Water Exchange on ^{15}N Transverse Relaxation. Journal of American Chemical Society 129:2971-2980.

15. Iwahara, J., C. D. Schwieters, and G. M. Clore. 2004. Characterization of Nonspecific Protein-DNA Interactions by ^{1}H Paramagnetic Relaxation Enhancement. Journal of American Chemical Society 126:12800-12808.

16. Iwahara, J., C. Tang, and G. M. Clore. 2007. Practical aspects of ^{1}H transverse paramagnetic relaxation enhancement measurements on macromolecules. Journal of Magnetic Resonance 184:185-195.

17. Lewin, B. 2000. Genes VII. Oxford University Press: Oxford, UK. 545.

18. Isaac, V. E., L. Patel, T. Curran, and C. Abate-Shen. 1995. Use of Fluorescence Resonance Energy Transfer To Estimate Intramolecular Distances in the Msx-1 Homeodomain. Biochemistry 34:15276-15281.

19. Zhang, S., V. Metelev, D. Tabatadze, P. C. Zamecnik, and A. Bogdanov. 2008. Fluorescence resonance energy transfer in near-infrared fluorescent oligonucleotide probes for detecting protein-DNA interactions. Proceedings of the National Academy of Sciences U S A 105:4156-4161.

20. Sahu, D., G. M. Clore, and J. Iwahara. 2007. TROSY-Based z-Exchange Spectroscopy: Application to the Determination of the Activation Energy for Intermolecular Protein Translocation between Specific Sites on Different DNA Molecules. Journal of American Chemical Society 129:13232-13237.

21. Farrow, N. A., O. Zhang, J. D. Forman-Kay, and L. E. Kay. 1994. A heteronuclear correlation experiment for simultaneous determination of ^{15}N longitudinal decay and chemical exchange rates of systems in slow equilibrium. Journal of Biomolecular NMR 4:727-734.

22. Loria, J. P., M. Rance, and A. G. Palmer. 1999. A TROSY CPMG sequence for characterizing chemical exchange in large proteins. Journal of Biomolecular NMR 15:151-155.

23. Kay, L. E., L. K. Nicholson, F. Delaglio, A. Bax, and D. A. Torchia. 1992. Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T1 and T2 values in proteins. Journal of Magnetic Resonance 97:359-375.

24. Bax, A., and A. Grishaev. 2005. Weak alignment NMR: a hawk-eyed view of biomolecular structure. Current Opinion in Structural Biology 15:563-570.

25. Lipsitz, R. S., and N. Tjandra. 2004. Residual dipolar couplings in NMR structure analysis. Annual Review of Biophysics and Biomolecular Structure 33:387-413.

26. Battiste, J. L., and G. Wagner. 2000. Utilization of Site-Directed Spin Labeling and High-Resolution Heteronuclear Nuclear Magnetic Resonance for Global Fold Determination of Large Proteins with Limited Nuclear Overhauser Effect Data. Biochemistry 39:5355-5365.

27. Clore, G. M., and J. Iwahara. 2009. Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes. Chemical Reviews 109:4108-4139.

28. Gillespie, J. R., and D. Shortle. 1997. Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. Journal of Molecular Biology 268:170-184.

29. Liang, B., J. H. Bushweller, and L. K. Tamm. 2006. Site-Directed Parallel Spin-Labeling and Paramagnetic Relaxation Enhancement in Structure Determination of Membrane Proteins by Solution NMR Spectroscopy. Journal of the American Chemical Society 128:4389-4397.

30. Otting, G., Y. Q. Qian, M. Billeter, M. Muller, M. Affolter, W. J. Gehring, and K. Wuthrich. 1990. Protein DNA Contacts in the Structure of a Homeodomain DNA Complex Determined by Nuclear-Magnetic-Resonance Spectroscopy in Solution. Embo Journal 9:3085-3092.

31. Pabo, C. O., and R. T. Sauer. 1984. Protein-DNA Recognition. Annual Review of Biochemistry 53:293-321.

32. Jin, C., I. Marsden, X. Chen, and X. Liao. 1999. Dynamic DNA Contacts Observed in the NMR Structure of Winged Helix Protein-DNA Complex. Journal of Molecular Biology 289:683-690.

33. Misteli, T. 2001. Protein Dynamics: Implications for Nuclear Architecture and Gene Expression. Science 291:843-847.

34. Riggs, A. D., S. Bourgeois, and M. Cohn. 1970. The lac represser-operator interaction: III. Kinetic studies. Journal of Molecular Biology 53:401-417.

35. Berg, O. G., R. B. Winter, and P. H. Von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. Biochemistry 20:6929-6948.

36. Cho, S., and P. C. Wensink. 1997. DNA binding by the male and female doublesex proteins of Drosophila melanogaster. The Journal of Biological Chemistry 272:3185-3189.

37. Fried, M. G., and D. M. Crothers. 1984. Kinetics and mechanism in the reaction of gene regulatory proteins with DNA. Journal of Molecular Biology 172:263-282.

38. Halford, S. E., and J. F. Marko. 2004. How do site-specific DNA-binding proteins find their targets? Nucleic Acids Research 32:3040-3052.

39. Von Hippel, P. H., and O. G. Berg. 1989. Facilitated target location in biological systems. The Journal of Biological Chemistry 264:675-678.

40. Winter, R. B., O. G. Berg, and P. H. Von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor-operator interaction: kinetic measurements and conclusions. Biochemistry 20:6961-6977.

41. Winter, R. B., and P. H. von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 2. The Escherichia coli repressor--operator interaction: equilibrium measurements. Biochemistry 20:6948-6960.

42. Greene, E. C., and K. Mizuuchi. 2004. Visualizing the assembly and disassembly mechanisms of the MuB transposition targeting complex. The Journal of Biological Chemistry 279:16736-16743.

43. Prasad, T. K., C. C. Yeykal, and E. C. Greene. 2006. Visualizing the assembly of human Rad51 filaments on double-stranded DNA. Journal of Molecular Biology 363:713-728.

44. Yeykal, C. C., and E. C. Greene. 2006. Visualizing the behavior of human Rad51 at the single-molecule level. Cell Cycle 5:1033-1038.

45. Bustamante, C., Z. Bryant, and S. B. Smith. 2003. Ten years of tension: single-molecule DNA mechanics. Nature 421:423-427.

46. Myong, S., M. M. Bruno, A. M. Pyle, and T. Ha. 2007. Spring-Loaded Mechanism of DNA Unwinding by Hepatitis C Virus NS3 Helicase. Science 317:513-516.

47. Hu, T., and B. I. Shklovskii. 2007. How a protein searches for its specific site on DNA: The role of intersegment transfer. Physical Review. E, Statistical, Nonlinear, and Soft matter Physics 76:51908-51909.

48. Slutsky, M., and L. A. Mirny. 2004. Kinetics of Protein-DNA Interaction: Facilitated Target Location in Sequence-Dependent Potential. Biophysical Journal 87:4021-4035.

49. Gerland, U., J. D. Moroz, and T. Hwa. 2002. Physical constraints and functional characteristics of transcription factor-DNA interaction. Proceedings of the National Academy of Sciences U S A 99:12015-12020.

50. Jeener, J., B. H. Meier, P. Bachmann, and R. R. Ernst. 1979. Investigation of exchange processes by two-dimensional NMR spectroscopy. The Journal of Chemical Physics 71:4546-4553.

51. Reuben, J., and D. Fiat. 1969. Nuclear Magnetic Resonance Studies of Solutions of the Rare-Earth Ions and Their Complexes. IV. Concentration and Temperature Dependence of the Oxygen-17 Transverse Relaxation in Aqueous Solutions. The Journal of Chemical Physics 51:4918-4927.

52. Tafvizi, A., F. Huang, J. S. Leith, A. R. Fersht, L. A. Mirny, and A. M. van Oijen. 2008. Tumor suppressor p53 slides on DNA with low friction and high stability. Biophysical Journal 95:L01-03.

53. Holmgren, A. 2008. The thioredoxin system. In Redox Biochemistry. R. Banergee, editor. Wiley-Interscience, Hoboken, NJ. 68-74.

54. Schafer, F. Q., and G. R. Buettner. 2001. Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. Free Radical Biology and Medicine 30:1191-1212.

55. Hoppe, G., K. E. Talcott, S. K. Bhattacharya, J. W. Crabb, and J. E. Sears. 2006. Molecular basis for the redox control of nuclear transport of the structural chromatin protein Hmgb1. Experimental Cell Research 312:3526-3538.

56. Agresti, A., and M. E. Bianchi. 2003. HMGB proteins and gene expression. Current Opinion in Genetics & Development 13:170-178.

57. Bustin, M., and R. Reeves. 1996. High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. Progress in Nucleic Acid Research and Molecular Biology 54:35-100.

58. Cai, M., Y. Huang, J. Y. Suh, J. M. Louis, R. Ghirlando, R. Craigie, and G. M. Clore. 2007. Solution NMR structure of the barrier-to-autointegration factor-Emerin complex. The Journal of Biological Chemistry 282:14525-14535.

59. Thomas, J. O. 2001. HMG1 and 2: architectural DNA-binding proteins. Biochemical Society transactions 29:395-401.

60. Thomas, J. O., and A. A. Travers. 2001. HMG1 and 2, and related 'architectural' DNA-binding proteins. Trends in Biochemical Sciences 26:219-219.

61. Sims, G. P., D. C. Rowe, S. T. Rietdijk, R. Herbst, and A. J. Coyle. 2010. HMGB1 and RAGE in Inflammation and Cancer. Annual Review of Immunology 28:367-388.

62. Allerhand, A., and E. Thiele. 1966. Analysis of Carr-Purcell Spin-Echo NMR Experiments on Multiple-Spin Systems. II. The Effect of Chemical Exchange. The Journal of Chemical Physics 45:902-916.

63. McConnell, H. M. 1958. Reaction Rates by Nuclear Magnetic Resonance. The Journal of Chemical Physics 28:430-431.

64. Ernst, R. R., G. Bodenhausen, and A. Wokaun. 1987. Principles of Nuclear Magnetic Resonance in One and Two Dimensions. Oxford University Press, UK.

65. Misteli, T. 2001. Protein dynamics: implications for nuclear architecture and gene expression. Science 291:843-847.

66. Perrin, C. L., and T. J. Dwyer. 1990. Application of two-dimensional NMR to kinetics of chemical exchange. Chemical Reviews 90:935-967.

67. Clore, G. M., A. Bax, P. C. Driscoll, P. T. Wingfield, and A. M. Gronenborn. 1990. Assignment of the side-chain 1H and 13C resonances of interleukin-1 beta using double- and triple-resonance heteronuclear three-dimensional NMR spectroscopy. Biochemistry 29:8172-8184.

68. Clore, G. M., J. G. Omichinski, and A. M. Gronenborn. 1991. Slow Conformational Dynamics at the Metal Coordination Site of a Zinc Finger. Journal of the American Chemical Society 113:4350-4351.

69. Montelione, G. T., and G. Wagner. 1989. 2D Chemical exchange NMR spectroscopy by proton-detected heteronuclear correlation. Journal of American Chemical Society 111:3096-3098.

70. Wider, G., C. Weber, and K. Wuthrich. 1991. Proton Proton Overhauser Effects of Receptor-Bound Cyclosporine-a Observed with the Use of a Heteronuclear-Resolved Half-Filter Experiment. Journal of the American Chemical Society 113:4676-4678.

71. Bosco, D. A., E. Z. Eisenmesser, S. Pochapsky, W. I. Sundquist, and D. Kern. 2002. Catalysis of cis/trans isomerization in native HIV-1 capsid by human cyclophilin A. Proceedings of the National Academy of Sciences U S A 99:5247-5252.

72. Nieto, P. M., B. Birdsall, W. D. Morgan, T. A. Frenkiel, A. R. Gargaro, and J. Feeney. 1997. Correlated bond rotations in interactions of arginine residues with ligand carboxylate groups in protein ligand complexes. Febs Letters 405:16-20.

73. Otting, G., E. Liepinsh, and K. Wuthrich. 1993. Disulfide Bond Isomerization in Bpti and Bpti(G36s) - an Nmr-Study of Correlated Mobility in Proteins. Biochemistry 32:3571-3582.

74. Sprangers, R., A. Gribun, P. M. Hwang, W. A. Houry, and L. E. Kay. 2005. Quantitative NMR spectroscopy of supramolecular complexes: Dynamic side pores in ClpP are important for product release. Proceedings of the National Academy of Sciences of the United States of America 102:16678-16683.

75. Wenter, P., G. Bodenhausen, J. Dittmer, and S. Pitsch. 2006. Kinetics of RNA refolding in dynamic equilibrium by H-1-detected N-15 exchange NMR spectroscopy. Journal of the American Chemical Society 128:7579-7587.

76. John, M., M. J. Headlam, N. E. Dixon, and G. Otting. 2007. Assignment of paramagnetic (15)N-HSQC spectra by heteronuclear exchange spectroscopy. Journal of Biomolecular NMR 37:43-51.

77. Pervushin, K., R. Riek, G. Wider, and K. Wuthrich. 1997. Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proceedings of the National Academy of Sciences U S A 94:12366-12371.

78. Delaglio, F., S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. Journal of Biomolecular NMR 6:277-293.

79. Johnson, B. A., and R. A. Blevins. 1994. Nmr View - a Computer-Program for the Visualization and Analysis of Nmr Data. Journal of Biomolecular Nmr 4:603-614.

80. Chance, E. M., A. R. Curtis, I. P. Jones, and C. R. Kirby. 1979. FACSIMILE: a computer program for flow and chemistry simulation and general initial value problems. In Atomic Energy Research Establishment Report R8775. Harwell, H. M. Stationary Office, London.

81. Guntert, P., N. Schaefer, G. Otting, and K. Wuthrich. 1993. Poma - a Complete Mathematica Implementation of the Nmr Product-Operator Formalism. Journal of Magnetic Resonance Series A 101:103-105.

82. Levitt, M. H. 1997. The signs of frequencies and phases in NMR. Journal of Magnetic Resonance 126:164-182.

83. Rance, M., J. P. Loria, and A. G. Palmer. 1999. Sensitivity improvement of transverse relaxation-optimized spectroscopy. Journal of Magnetic Resonance 136:92-101.

84. Schulte-Herbrüggen, T., and O. W. Sørensen. 2000. Clean TROSY: Compensation for Relaxation-Induced Artifacts. Journal of Magnetic Resonance 144:123-128.

85. Boyd, J., U. Hommel, and I. D. Campbell. 1990. Influence of cross-correlation between dipolar and anisotropic chemical shift relaxation mechanisms upon longitudinal relaxation rates of 15N in macromolecules. Chemical Physics Letters 175:477-482.

86. Kroenke, C. D., J. P. Loria, L. K. Lee, M. Rance, and A. G. Palmer. 1998. Longitudinal and transverse H-1-N-15 dipolar N-15 chemical shift anisotropy relaxation interference: Unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules. Journal of the American Chemical Society 120:7905-7915.

87. Sklenar, V., D. Torchia, and A. Bax. 1987. Measurement of C-13 Longitudinal Relaxation Using H-1 Detection. Journal of Magnetic Resonance 73:375-379.

88. Cavanagh, J., J. W. Fairbrother, G. A. Palmer, M. Rance, and J. N. Skelton. 2007. Relaxation and Dynamic Processes. In Protein NMR Spectroscopy: Principles and Practice. Elsevier Academic Press, San Diego, CA. 334-404.

89. Lieberman, B. A., and S. K. Nordeen. 1997. DNA intersegment transfer, how steroid receptors search for a target site. The Journal of Biological Chemistry 272:1061-1068.

90. Ruusala, T., and D. M. Crothers. 1992. Sliding and Intermolecular Transfer of the lac Repressor: Kinetic Perturbation of a Reaction Intermediate by a Distant DNA Sequence. Proceedings of the National Academy of Sciences U S A 89:4903-4907.

91. Fraenkel, E., and C. O. Pabo. 1998. Comparison of X-ray and NMR structures for the Antennapedia homeodomain-DNA complex. Nature Structural Biology 5:692-697.

92. Hovde, S., C. Abate-Shen, and J. H. Geiger. 2001. Crystal Structure of the Msx-1 Homeodomain/DNA Complex. Biochemistry 40:12013-12021.

93. Catron, K. M., N. Iler, and C. Abate. 1993. Nucleotides flanking a conserved TAAT core dictate the DNA binding specificity of three murine homeodomain proteins. Molecular Cell Biology 13:2354-2365.

94. Allerhan.A, and E. Thiele. 1966. Analysis of Car-Purcell Spin-Echo Nmr Experiments on Multiple-Spin Systems .2. Effect of Chemical Exchange. Journal of Chemical Physics 45:902-916.

95. Kramers, H. A. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. Physica 7:284-304.

96. Iwahara, J., and G. M. Clore. 2006. Direct Observation of Enhanced Translocation of a Homeodomain between DNA Cognate Sites by NMR Exchange Spectroscopy. Journal of the American Chemical Society 128:404-405.

97. Sahu, D., G. M. Clore, and J. Iwahara. 2007. TROSY-Based z-Exchange Spectroscopy: Application to the Determination of the Activation Energy for Intermolecular Protein Translocation between Specific Sites on Different DNA Molecules. Journal of the American Chemical Society 129:13232-13237.

98. Halford, S. E., and J. F. Marko. 2004. How do site-specific DNA-binding proteins find their targets? Nucleic Acids Research 32:3040-3052.

99. Von Hippel, P. H., and O. G. Berg. 1989. Facilitated target location in biological systems. The Journal of Biological Chemistry 264:675-678.

100. Berg, O. G., R. B. Winter, and P. H. von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. Biochemistry 20:6929-6948.

101. Zwanzig, R. 1997. Two-state models of protein folding kinetics. Proceedings of the National Academy of Sciences U S A 94:148-150.

102. Gerland, U., J. D. Moroz, and T. Hwa. 2002. Physical constraints and functional characteristics of transcription factor-DNA interaction. Proceedings of the National Academy of Sciences U S A 99:12015-12020.

103. Tafvizi, A., F. Huang, J. S. Leith, A. R. Fersht, L. A. Mirny, and A. M. van Oijen. 2008. Tumor suppressor p53 slides on DNA with low friction and high stability. Biophysical Journal 95:L01-03.

104. Doucleff, M., and G. M. Clore. 2008. Global jumping and domain-specific intersegment transfer between DNA cognate sites of the multidomain transcription factor Oct-1. Proceedings of the National Academy of Sciences U S A 105:13871-13876.

105. Agresti, A., and M. E. Bianchi. 2003. HMGB proteins and gene expression. Current Opinion in Genetics & Development 13:170-178.

106. Bianchi, M. E., and A. A. Manfredi. 2007. High-mobility group box 1 (HMGB1) protein at the crossroads between innate and adaptive immunity. Immunological Reviews 220:35-46.

107. Sahu, D., P. Debnath, Y. Takayama, and J. Iwahara. 2008. Redox properties of the A-domain of the HMGB1 protein. FEBS Letters 582:3973-3978.

108. Park, S., and S. J. Lippard. 2011. Redox State-Dependent Interaction of HMGB1 and Cisplatin-Modified DNA. Biochemistry 50:2567-2574.

109. Ohndorf, U. M., M. A. Rould, Q. He, C. O. Pabo, and S. J. Lippard. 1999. Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. Nature 399:708-712.

110. Hoppe, G., K. E. Talcott, S. K. Bhattacharya, J. W. Crabb, and J. E. Sears. 2006. Molecular basis for the redox control of nuclear transport of the structural chromatin protein Hmgb1. Experimental Cell Research 312:3526-3538.

111. Sharma, D., and K. Rajarathnam. 2000. C-13 NMR chemical shifts can predict disulfide bond formation. Journal of Biomolecular NMR 18:165-171.

112. Iwahara, J., C. D. Schwieters, and G. M. Clore. 2004. Characterization of nonspecific protein-DNA interactions by ^{1}H paramagnetic relaxation enhancement. Journal of the American Chemical Society 126:12800-12808.

113. Clore, G. M., and A. M. Gronenborn. 1998. Determining the structures of large proteins and protein complexes by NMR. Trends in Biotechnology 16:22-34.

114. Schafer, F. Q., and G. R. Buettner. 2001. Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. Free Radical Biology and Medicine 30:1191-1212.

115. Greenfield, N. J. 2006. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nature Protocols 1:2527-2535.

116. Holmgren, A., and C. Lyckeborg. 1980. Enzymatic reduction of alloxan by thioredoxin and NADPH-thioredoxin reductase. Proceedings of the National Academy of Sciences U S A 77:5149-5152.

117. Creighton, T. E. 1988. Disulphide bonds and protein stability. Bioessays 8:57-63.

118. Soboll, S., S. Grundel, J. Harris, V. Kolb-Bachofen, B. Ketterer, and H. Sies. 1995. The content of glutathione and glutathione S-transferases and the glutathione peroxidase activity in rat liver nuclei determined by a non-aqueous technique of cell fractionation. Biochemical Journal 311 (Pt 3):889-894.

119. Kohlstaedt, L. A., D. S. King, and R. D. Cole. 1986. Native state of high mobility group chromosomal proteins 1 and 2 is rapidly lost by oxidation of sulfhydryl groups during storage. Biochemistry 25:4562-4565.

120. Bonaldi, T., F. Talamo, P. Scaffidi, D. Ferrera, A. Porto, A. Bachi, A. Rubartelli, A. Agresti, and M. E. Bianchi. 2003. Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. EMBO J 22:5551-5560.

121. Gardella, S., C. Andrei, D. Ferrera, L. V. Lotti, M. R. Torrisi, M. E. Bianchi, and A. Rubartelli. 2002. The nuclear protein HMGB1 is secreted by monocytes via a non-classical, vesicle-mediated secretory pathway. EMBO Rep 3:995-1001.

122. Holmgren, A. 1979. Reduction of disulfides by thioredoxin. Journal of Biological Chemistry 254:9113-9119.

123. Takayama, Y., D. Sahu, and J. Iwahara. 2010. NMR Studies of Translocation of the Zif268 Protein between Its Target DNA Sites. Biochemistry 49:7998-8005.

124. Ellerman, J. E., C. K. Brown, M. de Vera, H. J. Zeh, T. Billiar, A. Rubartelli, and M. T. Lotze. 2007. Masquerader: high mobility group box-1 and cancer. Clinical Cancer Research 13:2836-2848.

AMERICAN CHEMICAL SOCIETY LICENSE TERMS AND CONDITIONS

Mar 22, 2011

This is a License Agreement between Debashish Sahu ("You") and American Chemical Society ("American Chemical Society") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by American Chemical Society, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2634280420810
License Date	Mar 22, 2011
Licensed content publisher	American Chemical Society
Licensed content publication	Journal of the American Chemical Society
Licensed content title	TROSY-Based z-Exchange Spectroscopy: Application to the Determination of the Activation Energy for Intermolecular Protein Translocation between Specific Sites on Different DNA Molecules
Licensed content author	Debashish Sahu et al.
Licensed content date	Oct 1, 2007
Volume number	129
Issue number	43
Type of Use	Thesis/Dissertation
Requestor type	Not specified
Format	Print
Portion	Full article
Author of this ACS article	Yes
Order reference number	
Title of the thesis / dissertation	NMR studies on kinetics of processes involving DNA-binding proteins
Expected completion date	May 2011
Estimated size(pages)	140
Billing Type	Invoice
Billing Address	711 Holiday Drive APT 40 Galveston, TX 77550 United States

Customer reference info

Total 0.00 USD

Terms and Conditions

Thesis/Dissertation

ACS / RIGHTSLINK TERMS & CONDITIONS
THESIS/DISSERTATION

INTRODUCTION
The publisher for this copyrighted material is the American Chemical Society. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

LIMITED LICENSE
Publisher hereby grants to you a non-exclusive license to use this material. Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process.

GEOGRAPHIC RIGHTS: SCOPE
Licenses may be exercised anywhere in the world.

RESERVATION OF RIGHTS
Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

PORTION RIGHTS STATEMENT: DISCLAIMER
If you seek to reuse a portion from an ACS publication, it is your responsibility to examine each portion as published to determine whether a credit to, or copyright notice of, a third party owner was published adjacent to the item. You may only obtain permission via Rightslink to use material owned by ACS. Permission to use any material published in an ACS publication, journal, or article which is reprinted with permission of a third party must be obtained from the third party owner. ACS disclaims any responsibility for any use you make of items owned by third parties without their permission.

REVOCATION
The American Chemical Society reserves the right to revoke a license for any reason, including but not limited to advertising and promotional uses of ACS content, third party usage, and incorrect figure source attribution.

LICENSE CONTINGENT ON PAYMENT

While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

COPYRIGHT NOTICE: DISCLAIMER
You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Reprinted (“Adapted” or “in part”) with permission from REFERENCE CITATION. Copyright YEAR American Chemical Society."

WARRANTIES: NONE
Publisher makes no representations or warranties with respect to the licensed material.

INDEMNITY
You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

NO TRANSFER OF LICENSE
This license is personal to you or your publisher and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

NO AMENDMENT EXCEPT IN WRITING
This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

OBJECTION TO CONTRARY TERMS
Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

JURISDICTION

This license transaction shall be governed by and construed in accordance with the laws of the District of Columbia. You hereby agree to submit to the jurisdiction of the courts located in the District of Columbia for purposes of resolving any disputes that may arise in connection with this licensing transaction.

THESES/DISSERTATION TERMS
Regarding your request for permission to include **your** paper(s) or portions of text from **your** paper(s) in your thesis/dissertation, permission is now automatically granted; please pay special attention to the **implications** paragraph below. The Copyright Subcommittee of the Joint Board/Council Committees on Publications approved the following:

Copyright permission for published and submitted material from theses and dissertations
ACS extends blanket permission to students to include in their theses and dissertations their own articles, or portions thereof, that have been published in ACS journals or submitted to ACS journals for publication, provided that the ACS copyright credit line is noted on the appropriate page(s).

Publishing **implications** of electronic publication of theses and dissertation material
Students and their mentors should be aware that posting of theses and dissertation material on the Web prior to submission of material from that thesis or dissertation to an ACS journal may affect publication in that journal. Whether Web posting is considered prior publication may be evaluated on a case-by-case basis by the journal's editor. If an ACS journal editor considers Web posting to be "prior publication", the paper will not be accepted for publication in that journal. If you intend to submit your unpublished paper to ACS for publication, check with the appropriate editor prior to posting your manuscript electronically.

Reuse/Republication of the Entire Work in Theses or Collections: Authors may reuse all or part of the Submitted, Accepted or Published Work in a thesis or dissertation that the author writes and is required to submit to satisfy the criteria of degree-granting institutions. Such reuse is permitted subject to the ACS' "Ethical Guidelines to Publication of Chemical Research" (http://pubs.acs.org/page/policy/ethics/index.html); the author should secure written confirmation (via letter or email) from the respective ACS journal editor(s) to avoid potential conflicts with journal prior publication*/embargo policies. Appropriate citation of the Published Work must be made. If the thesis or dissertation to be published is in electronic format, a direct link to the Published Work must also be included using the ACS Articles on Request author-directed link - see http://pubs.acs.org/page/policy/articlesonrequest/index.html

* Prior publication policies of ACS journals are posted on the ACS website at http://pubs.acs.org/page/policy/prior/index.html

If your paper has not yet been published by ACS, please print the following credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [JOURNAL NAME], in press (or 'submitted for publication'). Unpublished work copyright [CURRENT YEAR] American Chemical Society." Include appropriate information.

If your paper has already been published by ACS and you want to include the text or portions of the text in your thesis/dissertation in **print or microfilm formats**, please print the ACS copyright credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [FULL REFERENCE CITATION.] Copyright [YEAR] American Chemical Society." Include appropriate information.

Submission to a Dissertation Distributor: If you plan to submit your thesis to UMI or to another dissertation distributor, you should not include the unpublished ACS paper in your thesis if the thesis will be disseminated electronically, until ACS has published your paper. After publication of the paper by ACS, you may release the entire thesis (**not the individual ACS article by itself**) for electronic dissemination through the distributor; ACS's copyright credit line should be printed on the first page of the ACS paper.

v1.2

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK10954917.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact Rightslink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Vita

Debashish Sahu was born in Kalpakkam, India on May 10th 1985, to the proud parents Hrushi Kesh and Jyotirmayee Sahu. Debashish earned a Bachelors of Technology with honors, in Industrial Biotechnology from Alagappa College of Technology, Anna University in May of 2006. After his undergraduate education, he joined the Ph. D. program in Molecular Biophysics Education Track at University of Texas Medical Branch. In 2007, Debashish began his thesis work in the laboratory of Dr. Junji Iwahara in the department of biochemistry and molecular biology. While at UTMB, Debashish received many awards like Curtis W. Lambert scholarship, Robert A. Welch award in chemistry and Finn Wold travel award. Debashish was nominated as one of the Who's who among students in American universities in 2011. During his tenure at UTMB, he has actively served in many committees for the biochemistry and molecular biology program recruitment, the biological chemistry student organization and the international student organization. In addition, Debashish has been the teaching assistant for the graduate level course, Molecular Biophysics I and also taught a portion of Biochemistry course for the first year graduate students.

Debashish can be contacted at 711 Holiday Drive APT 40, Galveston, Texas, 77550.

EDUCATION

B. Tech, May 2006, Alagappa College of Technology, Anna University, Chennai, India

PUBLICATIONS

A. Articles in Peer-Reviewed Journals:

Takayama, Y., **Sahu, D.** and Iwahara, J., NMR studies of translocation of the Zif268 protein between its target DNA sites, Biochemistry, 49(37), 7998-8005, 2010.

Sahu, D., Takayama, Y. and Iwahara, J., Redox properties of the A-domain of the HMGB1 protein, FEBS Letters, 582(29), 3973-3978, 2008.

Takayama, Y., **Sahu, D.** and Iwahara, J., Observing In-phase Single-Quantum 15N Multiplets for NH2/NH3+ Groups with Two-dimensional Heteronuclear Correlation Spectroscopy, Journal of Magnetic Resonance, 194(2), 313-316, 2008.

Sahu, D., Clore, G.M., and Iwahara, J., TROSY-Based z-Exchange Spectroscopy: Application to the Determination of the Activation Energy for Intermolecular Protein Translocation between Specific Sites on Different DNA Molecules, Journal of American Chemical Society, 129(43): 13232 - 13237, 2007.

Jayaram, B., Bhushan, K., Thukral, L., Shenoy, S.R., Narang, P., Bose, S., Agrawal, P., **Sahu, D.** and Pandey, V., Bhageerath: an energy based web enabled computer software suite for limiting the search space of tertiary structures of small globular proteins, Nucleic Acids Research, 34(21): 6195–6204, 2006.

B. Abstracts in Proceedings of Conferences/Symposia:

Sahu, D. and Iwahara, J., Kinetics of protein translocation between two DNA duplexes at equilibrium: Validation of NMR studies, Protein Society Meeting 2010 Aug 1-5.

Sahu, D. and Iwahara, J., Kinetics of protein translocation between two DNA duplexes at equilibrium: Validation of NMR studies, John S. Dunn, Sr. GCC Conference 2010 Jul 22.

Sahu, D., Takayama, Y. and Iwahara, J., Theoretical And Experimental Approaches To Analyze Facilitated Target Location Process, Sealy Center for Structural Biology Symposium, UTMB 2010 March 19.

Sahu, D. and Iwahara, J., Theoretical consideration on NMR approach to analyze kinetics of intermolecular protein translocation between nonspecific DNA molecules, Keystone Symposium 2009 Feb 15-20.

Sahu, D. and Iwahara, J., Theoretical consideration on NMR approach to analyze kinetics of intermolecular protein translocation between nonspecific DNA molecules, John S. Dunn, Sr. GCC Conference 2009 Feb 6-7.

Sahu, D. and Iwahara, J., Determination of the Activation Energy for Intermolecular Protein Translocation between Specific Sites on Different DNA Molecules, Poster in Protein Society Meeting 2008 July 18-23.

Sahu, D. and Iwahara, J., Characterization of Redox Properties of HMGB1 A-domain using NMR Spectroscopy, Poster presented in Sealy Center for Structural Biology Symposium, UTMB 2008 May 17.

Sahu, D., Clore, G.M., and Iwahara, J., TROSY-Based z-Exchange Spectroscopy: Application to the Determination of the Activation Energy for Intermolecular Protein Translocation between Specific Sites on Different DNA Molecules, Poster at Biochemistry and Molecular Biology new student orientation, UTMB 2007 Sep 26. (Won the Best Poster Award)

Sahu, D. and Iwahara, J., Development of New NMR Methods to Investigate Side-chain Dynamics, Kinetics, and Energetics of Protein-DNA Interactions, Poster presented in Sealy Center for Structural Biology Symposium, UTMB 2007 May 18-19.

Jayaram, B., Bhushan, K., Thukral, L., Shenoy, S.R., Narang, P., Bose, S., Agrawal, P., **Sahu, D.** and Pandey, V., Bhageerath: An Energy Based Protein Tertiary Structure Prediction Server for Small Globular Proteins, Abstracts in CASP7, 2006 Nov 26.