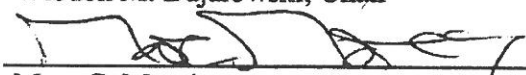The Dissertation Committee for Levani Zandarashvili Certifies that this is the approved version of the following Dissertation:

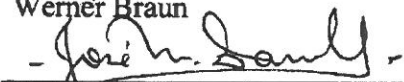**NMR Studies on Protein Dynamics and Its Relevance to Functional Kinetics**
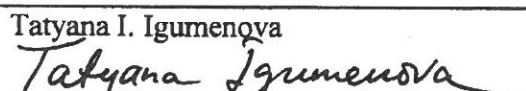
Committee:

Junji Iwahara, Supervisor

Wlodek M. Bujalowski, Chair

Marc C. Morais

Werner Braun

Jose M. Barral

Ken Fujise

Tatyana I. Igumenova

_____

Dean, Graduate School of Biomedical Sciences

NMR Studies on Protein Dynamics and Its Relevance to Functional Kinetics.

by

Levani Zandarashvili, B.S.

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 Philosophy

The University of Texas Medical Branch
August, 2013

Dedication

With love and gratitude to my parents, Tamazi Zandarashvili and Mariami Shamugia.

Acknowledgements

I would like to express my gratitude to my mentor, Dr. Junji Iwahara, for guiding me all these years. All this time, Dr. Iwahara was an example of an excellent scientist, mentor and a teacher, always ready to answer all of my questions no matter how long it would take to explain.

I would also like to thank my supervisory committee members, Dr. Wlodek Bujalowski, Dr. Werner Braun, Dr. Marc Morais, Dr. Jose Barral, Dr. Tatyana Igumenova and Dr. Ken Fujise for giving me their time and support over my PhD years.

Thanks to my current and previous lab members, Alexandre Esadze and Dr. Debashish Sahu, as well as my "lab neighbor" Dr. Michal Szymanski and Lisa Pipper for their friendship and valuable professional advices.

I want to especially thank Dr. James C. Lee, thanks to whom I came to UTMB. Even after I left his lab, he was always ready to encourage me and give me his advices.

Also, I want to thank Lisa Pipper and Aaron Brown for language editing of my final dissertation.

And finally, I want to thank my family and friends, who regardless of being over the ocean, were able to stand by my side and help me go through this long and difficult way.

NMR Studies on Protein Dynamics and Its Relevance to Functional Kinetics.

Publication No._____

Levani Zandarashvili, PhD

The University of Texas Medical Branch, 2013

Supervisor: Junji Iwahara

ABSTRACT

Existence of a strong relationship between protein's dynamics and its functional kinetics was suggested decades ago. Even though different groups investigated this relationship, technical difficulties and lack of appropriate methodology make full understanding of this relationship quite challenging. In this project we approach investigation of this phenomenon from several different points.

In the first part of the project, we demonstrate NMR based approach to elucidate lysine side-chain dynamics, as well as dynamics of hydrogen bonds in which they are involved. Such methodology provides us with detailed information on protein's behavior at atomic and pseudo-atomic resolutions. As amino acid side chains are involved in macromolecular interactions, knowledge on side-chain dynamics allows better understanding of the dynamic nature of such interactions. Presented data strongly suggests that hydrogen bonds involving lysine NH_3^+ groups have a very dynamic and transient nature.

In the second part, we investigated the role of domain dynamics in translocation kinetics of human transcription factor Egr-1. Nature has designed Egr-1 in such a way, at the same time with strong DNA binding affinity for high target specificity, that it is rapid enough in translocation on the human genome to satisfy requirements of its biological functions. Our data strongly suggests that Egr-1 is able to achieve rapidity and high specificity due to the dynamic nature of its domains in the nonspecific complex with DNA. We also demonstrated that nature's way of designing zinc-finger transcription factors can be used to optimize activity of artificial ones.

In the final part of this project, we demonstrated novel NMR based *in situ* method for elucidation of proteins' characteristics in various biological fluids. Using this method we determined oxidation speed of HMGB1 protein in human blood serum and saliva, as

well as in extracellular fluids. Different effect of these fluids on HMGB1's oxidation clearly showed the importance of *in situ* experiments for better understanding of protein function in its natural environment.

Studies presented here demonstrate that in order to understand the protein's dynamics/functional kinetics it is essential to simultaneously use different orthogonal approaches.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

HMGB1	High-mobility box 1
MD	Molecular dynamics
ND	Nuclease domain
NMR	Nuclear Magnetic Resonance
SOFAST-HMQC	band-selective optimized flip-angle short-transient heteronuclear multiple quantum coherence
ZF	Zinc finger
ZFN	Zinc finger nuclease

CHAPTER 1

Introduction

Macromolecular interactions define biological processes. In order to explain and predict the outcome of these complex processes, macromolecular interactions should be understood at atomic and molecular levels. Proteins are the major players in these interactions. They are very complex macromolecules, which take part in virtually all biological processes. For over a half century since the first high-resolution crystal structure of a protein was revealed in 1959, researchers in various biological fields pursued mechanisms of biological processes via determination of three-dimensional structures of proteins and their complexes with other molecules.[1] The efforts through X-ray crystallography, NMR spectroscopy, electron microscopy and other biophysical methods[2-27] have been remarkably successful, revealing the structural basis of various macromolecular functions.

While protein structures are essential information for our understanding of biological processes, it should be noted that, in reality, proteins are not rigid molecules with a single defined structure as they are shown in biology textbooks.[28-30] At any given moment they undergo dynamic changes. Proteins have multiple states in solution and each of these states might be essential for their biological function.[5-9] Flexible loops and side-chains often determine intra-molecular and inter-molecular interactions. There is an increasing number of evidences supporting the notion that conformational dynamics is the main determinant of proteins' ability to perform its functions, such as catalysis, allosterism and signal transduction.[6,17,28-30] For example, high importance of solution dynamic studies was demonstrated for the hyaluronan (HA) receptor CD44 protein's interactions with HA.[15,31-35] Initial studies of CD44's structure were quite controversial as

NMR solution experiments indicated that the HA-binding domain (HABD) of CD44 becomes unstructured upon HA binding[31-33], while the crystal structure showed that overall conformation of the HABD did not change much upon HA binding[35]. This controversy was resolved by additional solution NMR studies which showed that HA-free CD44 undergoes dynamic exchange between folded and partially unfolded states[15]. In the partially unfolded state recognition loop of CD44 exhibited dynamic motions. Populations of folded and partially unfolded states change from 86% for folded and 14% for partially folded states to 8% and 92% upon HA binding. This result explained the discrepancy between NMR and X-ray structures, indicating that only the folded state was observed in X-ray studies. CD44 is used by lymphocytes for rolling adhesion on endothelial cells, which plays an important role in the recruiting of the lymphocytes to inflammatory sites. Studies by Ogino S. *et al.* strongly suggested that ligand mediated dynamic exchange of CD44 between low-affinity (mainly folded) and high-affinity (mainly unfolded) states is crucial for cell rolling on ligand-coated surfaces. This observation was a direct evidence of the importance of CD44's recognition loop dynamics in protein function. Therefore, CD44 is an excellent example of a relationship between the protein's internal dynamics and its function. Recent works by D. Kern's group on adenylate kinase (Adk) is another great example demonstrating the role of protein dynamics in enzymatic function[7,22]. Adk catalyses conversion of ADP into ATP and AMP, therefore playing vital role in cell functioning. Though Adk has been studied extensively over the years, many important aspects of its function, such as the determinant(s) of the rate-limiting step in its catalysis, were poorly understood. In 2004, Wolf-Watz *et al.* demonstrated that Adk's nucleotide binding lid performs opening/closing dynamic motions on us-ms timescale[7]. These motions are required for the product release and are the rate-limiting step in protein's function. In their following work (Henzler-Wildman *et al.*, 2007) by combination of NMR relaxation, dynamics simulation and fluorescence studies, they also demonstrated that ps-ns motions in key

positions facilitate slower, large-amplitude domain motions.[22] This work suggests the importance of protein's dynamic motions in its functional kinetics and once again demonstrated the need for a deeper understanding of protein dynamics and its role in macromolecular interactions.

Recent advances in molecular dynamics simulations give some promising results and often help interpretation of experimental observations[12,16,17]. But computational methods still remain premature for independent and reliable prediction of proteins' dynamics and functions, especially those occurring on a timescale slower than 10^{-7} s. Improvement of computational methods and models requires appropriate validation based on experimental data regarding macromolecular behavior at an atomic level.

NMR spectroscopy is arguably the most powerful and versatile methodology for studying macromolecular dynamics at an atomic level resolution[9,11,14,19,20]. Magnetic resonances of ^{1}H, ^{15}N and ^{13}C nuclei are extremely sensitive to their environment thus enabling detection of even subtle difference in structure and dynamics. Owing to this aspect, NMR spectroscopy permits analyzing protein structure and interaction with other macromolecules. Also, methodological advancements over the last two decades made it possible to analyze protein dynamics on an extremely wide range of timescales, from fast pico- to nano-second, up to slow seconds and hours timescale (Figure 1.1).[19,20] Literally thousands of different proteins and their complexes have been studied using NMR spectroscopy, providing information on their structure, ligand biding, folding, allostery, enzymatic activity and other properties. NMR provides detailed information on specific changes in protein at atomic and quasi-atomic resolution, and allows identification of specific binding sites, regions undergoing dynamic changes, functionally important amino acids and so on. Unlike X-ray crystallography, NMR experiments can be performed in solution allowing proteins to undergo dynamic changes between multiple
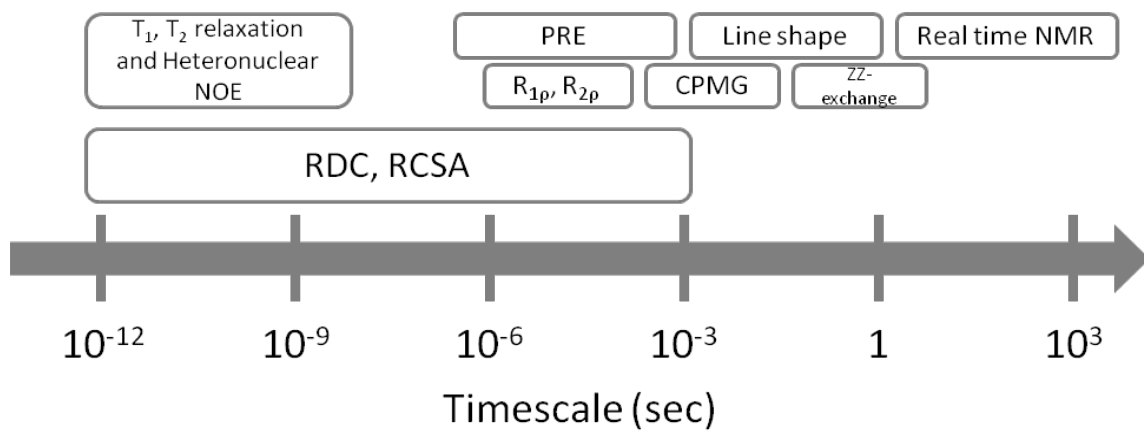
Figure 1.1:   NMR can be used to analyze protein dynamics on a very wide range of
timescales.

states. Some of them might be low population (<5%) states that are virtually invisible for conventional NMR as well as X-ray. But some new NMR methods, such as PRE and CPMG allow detection and analysis of such low populated states, even if their population is as low as 1%.[36-38] For many proteins such low population states are functionally important. Though NMR remains a major method for elucidating protein dynamics, it still cannot provide all tools necessary to fully analyze protein dynamics for various systems.

Although many significant steps have been made towards understanding the dynamics–function relationship, many aspects of this relationship remain to be addressed, mainly due to practical limitations of necessary experimental procedures. There are several barriers for establishing the link between protein's dynamics and its functional kinetics. One of the major bottlenecks is the lack of methods suitable for studying the dynamics of functionally important moieties of macromolecules.[4,39-43] The vast majority of NMR methods for investigating dynamics are applicable only to protein backbone or methyl groups. Although polar or ionizable moieties of side chains are crucial for protein function, available methodology to study the dynamics of the functionally important moieties is very limited. Another barrier for advancing the knowledge of proteins' dynamics/functional kinetics relationship is difficulty in modulating particular motions. Such engineering is possible only for well characterized molecules. Some research groups have suggested that protein dynamics can be a major determinant in the kinetics of protein functions, such as enzymatic activity, folding, target search kinetics and so on, underlying the importance of proteins' internal motions.[17,21,22,26,29] Unfortunately, due to lack of sufficient research data on these systems, the role of protein dynamics remains un-established, therefore more experimental data is necessary to fully understand the role of dynamics in functional kinetics. The third barrier is the lack of kinetic information on biological processes in actual biological fluids. Although nonspecific interactions in molecular crowding environment has been a major subject in biological sciences for

decades,[44-46] to date, there are only a few experimental methods that can provide pseudo-atomic or atomic resolution data on proteins in their physiological environment.[47,48] Lack of appropriate methods to study protein dynamics on atomic level resolution in physiological fluids leaves a shadow over reliability of data acquired using *in vitro* experiments. Overcoming these barriers is critical for establishing the dynamics-function relationship.

The **overall goal** of this project is to advance our understanding of the relationship between protein dynamics and functional kinetics (Figure 1.2). To achieve this goal, we should overcome the above-mentioned barriers. From this perspective we pursued the following three specific aims:

Aim 1: Development of new NMR methodology for hydrogen bond dynamics study involving side-chains. Hydrogen bonds play essential role in proteins' function. They take part in protein folding, interaction with other macromolecules, such as drugs, enzymes, DNA or other proteins. Due to the lack of suitable methods, dynamic properties of hydrogen bonds and their role in function remains poorly understood.[4,39-43] And it is even harder to analyze hydrogen bonds when they are formed using protein side-chains, as they are even more dynamic. Better understanding of protein side chain dynamics, as well as understanding dynamics of hydrogen bonds in which they are involved, will aid development of a wide range of biological fields. For example, different groups have suggested that proteins form transient dynamic states during their folding process stabilized by hydrogen bonds and side-chain interactions.[49] Therefore, deeper understanding the nature of hydrogen bonds, will enable us to better understand protein folding process. It has been shown that different drugs or inhibitors interact with proteins through hydrogen bonding networks. New knowledge will aid in the analysis of protein/drug complex behavior in solution, as well as improving development of docking softwares. In order to fill up the methodological gap preventing determination of protein
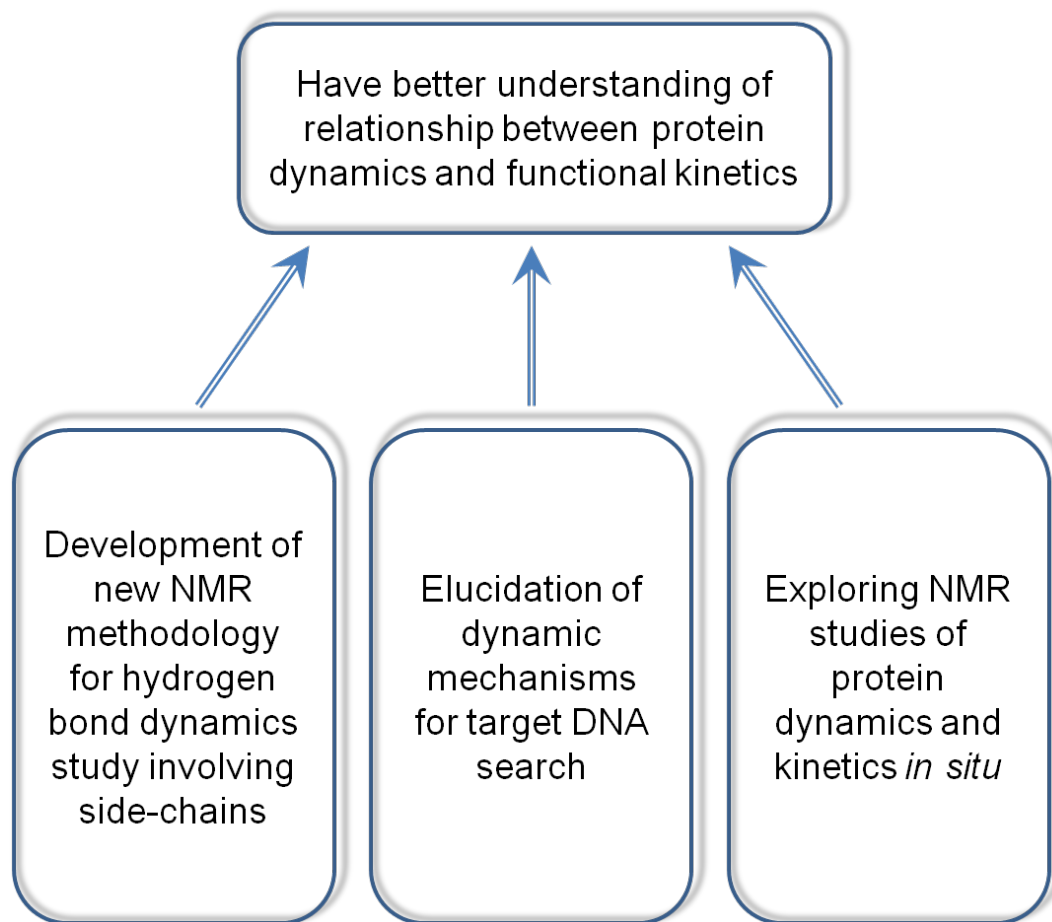
Figure 1.2:   Three specific aims pursued in this project.

side chain dynamics and their hydrogen bonds, we developed a new NMR method to characterize lysine side-chain dynamics (**CHAPTER 2**).[4] We demonstrate that this method can be effectively used to analyze dynamics of lysine side-chains, as well as to detect and characterize the hydrogen bonds involved in lysine NH_3^+ groups. In addition, our method can also be used to characterize dynamics of lysine side-chains through backbone nitrogen N-Cγ *J*-coupling, what presents very valuable information for benchmark and optimization of MD simulation force fields.

Aim 2: Elucidation of dynamic mechanisms for target DNA search. To have better understanding of relationship between protein dynamics and functional kinetics, we investigated the role of domain dynamics in proteins' translocation on DNA.[2] Some DNA binding proteins utilize dynamic nature of hydrogen bonds to form transient dynamic complexes with non-specific DNA sites.[50-52] Such dynamic complexes allow proteins to switch between two main conformations: search mode, and recognition mode.[50-52] In recognition mode protein forms number of ionic and hydrogen bonds with DNA trying to identify its target site. Due to their dynamic nature, hydrogen bonds are broken relatively easily, allowing proteins' fast translocation from one non-specific site to another, where it again switches to recognition mode, repeating cycle until it finds its target DNA sequence. To pursue this specific aim, I elucidated the domain dynamics of a human transcription factor Egr-1 and the role of its domain dynamics in translocation on DNA (**CHAPTER 3**). My data suggested the Egr-1's domain dynamics plays an important role in its translocation, allowing rapid activation of target genes on DNA. By pursuing this specific aim, I demonstrate the importance of protein dynamics in its functional activity.

Aim 3: Exploring NMR studies of protein dynamics and kinetics *in situ*. In order to investigate protein dynamics and function on molecular and atomic levels, we have to perform experiments in *in vitro* conditions, shielded from "noise" of non-specific interactions. Such simplification allows observation of only specific interactions, but

always raises questions: what happens in a cellular environment? what is the effect of crowding? how do non-specific interactions affect kinetics and dynamics of observed reactions? In specific aim 3 we developed unique NMR based method to investigate proteins' behavior in biological fluids and observe its properties at molecular, sub-molecular and pseudo-atomic levels (**CHAPTER 4**).[3] Here we studied oxidation of HMGB1 (high-mobility group box 1) protein in different extracellular fluids, as well as in human saliva and blood serum. By pursuing aim 3 we will be able to investigate protein dynamics *in situ* and compare results for *in vitro* experiments. This will allow us to better understand the effect of molecular crowding and non-specific interactions on protein dynamics and function.

This work is directed towards filling up the gap linking together molecular dynamics and functional kinetics. Only by detailed analysis of the system at cellular, molecular and atomic levels can one truly understand the role of protein dynamics in its function. By pursuing three specific aims we became one step closer to establishing a link between protein dynamics and function (Figure 1.2). By using NMR and fluorescence spectroscopy, along with other biophysical and biochemical approaches we a gained deeper understanding of proteins' dynamics at pseudo-atomic and sub-molecular levels. And only through this understanding we were able to control dynamics of Egr-1's zinc-finger domains, and consequently modulate its translocation kinetics. Our *in situ* method takes NMR to new step, allowing observation of the link between protein dynamics and function in more cell-like conditions. Overall, this project brings us one step closer to answering one of the most fundamental questions of biology, what is the connection between proteins' dynamics at atomic and molecular levels, and its function in biological processes.

CHAPTER 2:

Development of New NMR Methodology for Hydrogen Bond Dynamics Study Involving Side-Chains.

2.1 INTRODUCTION

Protein side chains play vital roles in molecular function, such as enzymatic catalysis and protein-protein interactions[53]. Although structural biology has provided important insights into the side-chain behavior of proteins and their complexes, the functional role of side-chain dynamics is not well understood. Powerful NMR techniques for the investigation of protein dynamics exist, but the vast majority are applicable only to the backbone or methyl groups.[19,20] The lack of specific experiments for characterizing the dynamics of side chains involved in hydrogen bonds and electrostatic interactions therefore represents a bottleneck in understanding structure/dynamics/function relationships in proteins. Recently, Iwahara *et al*. developed new NMR methods to characterize lysine side-chain amino groups, including the design of ^{15}N relaxation experiments for the extraction of side-chain S^2 order parameters.[40-43] For ubiquitin, the majority of the lysine NH_3^+ groups display notably low S^2 order parameters, reflecting a high degree of internal mobility for these functionally important entities, with correlation times in the subnanosecond range for the NH_3^+ bond rotations. However, the origin of the side-chain mobility could not be elucidated from experiment alone. On the basis of intra- and interresidual long-range ^{15}N-^{13}C scalar *J*-coupling constants, we have found in the

present work that the mobility about the χ_4 torsion angle of lysine side chains is generally high. Hydrogen bonds formed by the NH_3^+ groups therefore have a highly transient character. The experimental results have been quantitatively compared with an extended molecular dynamics (MD) simulation.

2.2 MATERIAL AND METHODS

2.2.1 NMR samples and experiments

$^1H/^{13}C/^{15}N$ labeled ubiquitin was purchased from VLI research. NMR experiments were done in d_6-succinate · NaOH (pH 5.0), 40 mM NaCl and 0.4 mM NaF buffer, at 2 ºC. Thin and bold bars in black represent hard rectangular 90º and 180º pulses, respectively. Unless indicated otherwise, pulse phases are along x. Carrier positions: ^1H, the position of the water resonance; ^{15}N, 33 ppm; $^{13}C\gamma$, 23 ppm; $^{13}C\varepsilon$, 45 ppm; and $^{13}C=O$, 177 ppm. In (b), two data sets were recorded in an interleaved manner, with the two ^{13}C IBURP-2 pulses6 located at either positions a or b. Pulsed field gradients were optimized to minimize the water signal. Short bold bars represent water-selective soft-rectangular ^1H 90º pulses (1.2 ms). Shaped pulses: ^1H half-Gaussian 90º pulse (2.1 ms); ^{15}N r-SNOB 180º pulse[55] (1.03 ms); and ^{13}C IBURP-2 180º pulses (1.1 ms). Delays: τ_a = 2.7 ms, δ = 2.6 ms, T = 50 ms, and T_d = 104 ms for the $^3J_{N\zeta C\gamma}$ experiments; T = 106 ms and T_d = 212 ms for the $^{h3}J_{N\zeta C'}$ experiments. Phase cycles: φ_1 = {2x, 2(-x)}, φ_2 = {4x, 4(-x)}, φ_3 = {x, -x}, φ_4 = {y, -y}, φ_5 = {8x, 8(-x)}, φ_6 = {8x, 8y, 8(-x), 8(-y)}, and receiver={x, -x, -x, x, 2(-x, x, x, -x), x, -x, -x, x}. Quadrature detection in the t_1 domain was achieved using States-TPPI for φ_1. The experiments were carried out using a Varian NMR system operated at a ^1H frequency of 800 MHz. NMR data were processed and analyzed using the NMRPipe[56] and NMRView[57] programs.

2.2.2 Data fitting for Karplus parameters' optimization

Karplus parameters were optimized using non-least-square fitting in MATLAB software (MathWorks, Inc) using MATLAB's genetic algorithm.

2.3 RESULTS

2.3.1 Measurement of $^3J_{\text{N}\zeta\text{C}\gamma}$ for lysine NH_3^+ groups

In this project we report two different types of long-range ^{15}N-^{13}C J-coupling constants involving the lysine NH_3^+ groups: the intraresidue J-coupling between ^{15}Nζ and ^{13}Cγ nuclei ($^3J_{\text{N}\zeta\text{C}\gamma}$) and the J-coupling between ^{15}Nζ and carbonyl ^{13}C' nuclei across a Nζ-Hζ · · · O=C' hydrogen bond ($^{h3}J_{\text{N}\zeta\text{C}'}$). Figure 2.1 shows the lysine-specific NMR pulse sequences for measuring the absolute values of the J-coupling constants between the ^{15}Nζ and ^{13}C nuclei. Because of the unique ^{15}N chemical shifts of NH_3^+ groups ($\sim$33 ppm), these pulse sequences permit selective observation of lysine NH_3^+ groups in a uniformly ^{13}C/^{15}N-labeled protein. The remarkably slow relaxation of the in-phase single-quantum N_+ terms for NH_3^+ groups[40,43] permits the use of relatively long periods for the evolution of ^{15}N transverse magnetizations in order to measure the small ^{15}N-^{13}C J-coupling constants. Using these pulse sequences, we recorded the data for 1 mM ^{13}C/^{15}N-labeled ubiquitin at pH 5 and 2 ℃. The low pH and temperature were necessary to make the hydrogen exchange sufficiently slow that lysine NH_3^+ protons could be observed.[40]

The pulse sequence in Figure 2.1a corresponds to a two-dimensional heteronuclear correlation experiment to observe signals arising from coherence transfer via ^{15}N-^{13}C J-coupling for lysine NH_3^+ groups. The F_2 dimension corresponds to ^{1}Hζ chemical shifts of the NH_3^+ groups and the F_1 dimension to ^{13}C chemical shifts of nuclei coupled to ^{15}Nζ. In the spectrum recorded with this pulse sequence, (^{13}Cγ, ^{1}Hζ) cross-
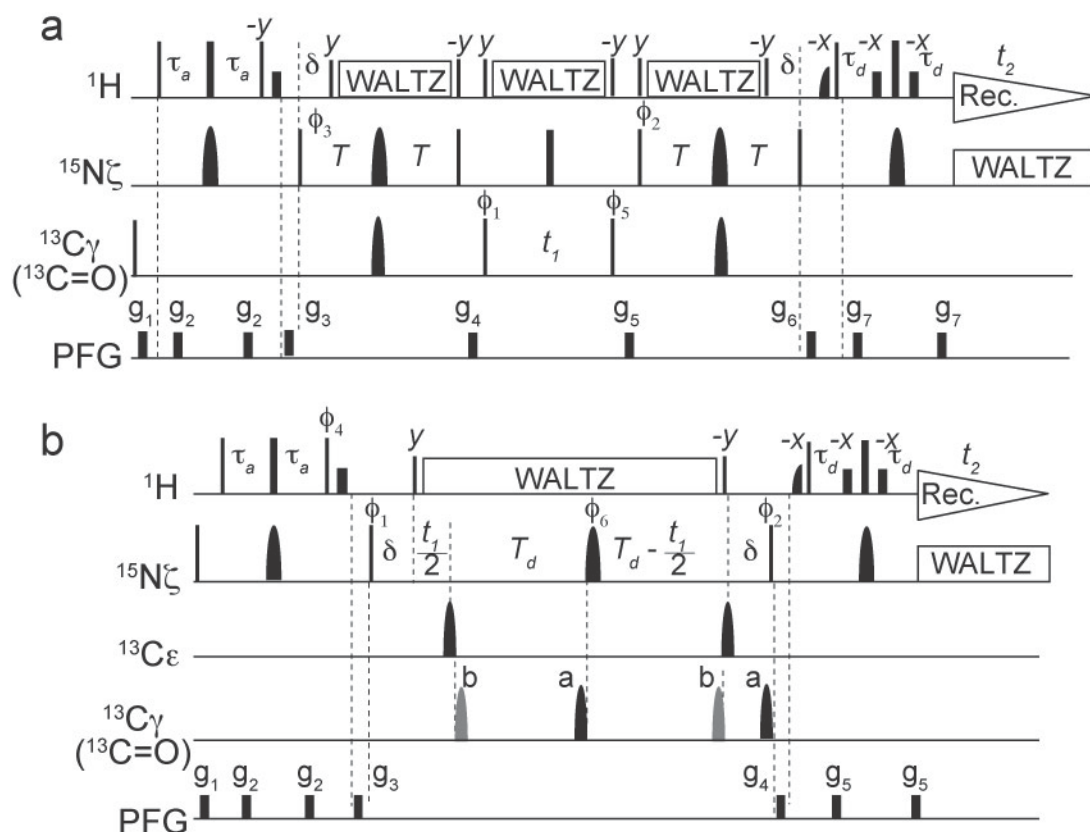
Figure 2.1. Pulse sequences for ^3J-coupling measurement. (a) Pulse sequence for observing heteronuclear correlation via long-range ^{15}N-^{13}C ^3J-coupling for lysine side-chain NH$_3^+$ groups. (b) Constant-time spin-echo difference ^{1}H-^{15}N correlation experiment to measure $^3J_{\text{N}\zeta\text{C}\gamma}$ and $^{h3}J_{\text{N}\zeta\text{C}'}$ for lysine NH$_3^+$ groups. Thin and bold bars in black represent hard rectangular 90° and 180° pulses, respectively.

peaks were clearly observed for all lysine residues, indicating that the χ_4-related $^3J_{N\zeta C\gamma}$ values are sizable (Figure 2.2a). Although the ^{13}C IBURP-2 pulses covered the lysine $^{13}C\gamma/^{13}C\delta$ region (23-30 ppm), no $^{13}C\delta$ signals were observed in the spectrum. The absence of the $^{13}C\delta$ signals suggested that $^2J_{N\zeta C\delta} < 0.2$ Hz (as estimated from the noise level). For the quantitative measurement of $^3J_{N\zeta C\gamma}$ values, we used the spin-echo $^{15}N\{^{13}C\}$ difference experiment shown in Figure 2.1b. In this experiment, which is analogous to those developed for measuring χ_1-related $^3J_{NC\gamma}$-coupling constants,[58-60] two subspectra were recorded in an interleaved manner: one with the $^{13}C\gamma$-selective IBURP-2 pulses at positions a (subspectrum a) and the other at positions b (subspectrum b). The net evolution time for $^3J_{N\zeta C\gamma}$-coupling for ^{15}N transverse magnetizations is $2(T_d + \delta)$ for the former case and zero for the latter. The values of $^3J_{N\zeta C\gamma}$ can be calculated from the expression

$$\frac{I_a}{I_b} = \cos\left[2\pi J_{N\zeta C\gamma}(T_d + \delta)\right]$$

in which I_a and I_b represent the signal intensities in subspectra a and b, respectively. Table 2.1 shows the values of $^3J_{N\zeta C\gamma}$ obtained for lysine side chains in ubiquitin.

The precision of the measured coupling constants was high because of the very slow ^{15}N relaxation of the NH_3^+ groups, which allowed the use of a long period for J modulation (209 ms).

The $^3J_{N\zeta C\gamma}$-coupling constants were subsequently interpreted by means of Karplus equation,[61,62] which is expressed as:

$$^3J = A\cos^2\chi + B\cos\chi + C$$

where A, B and C are called Karplus parameters. Because empirical Karplus
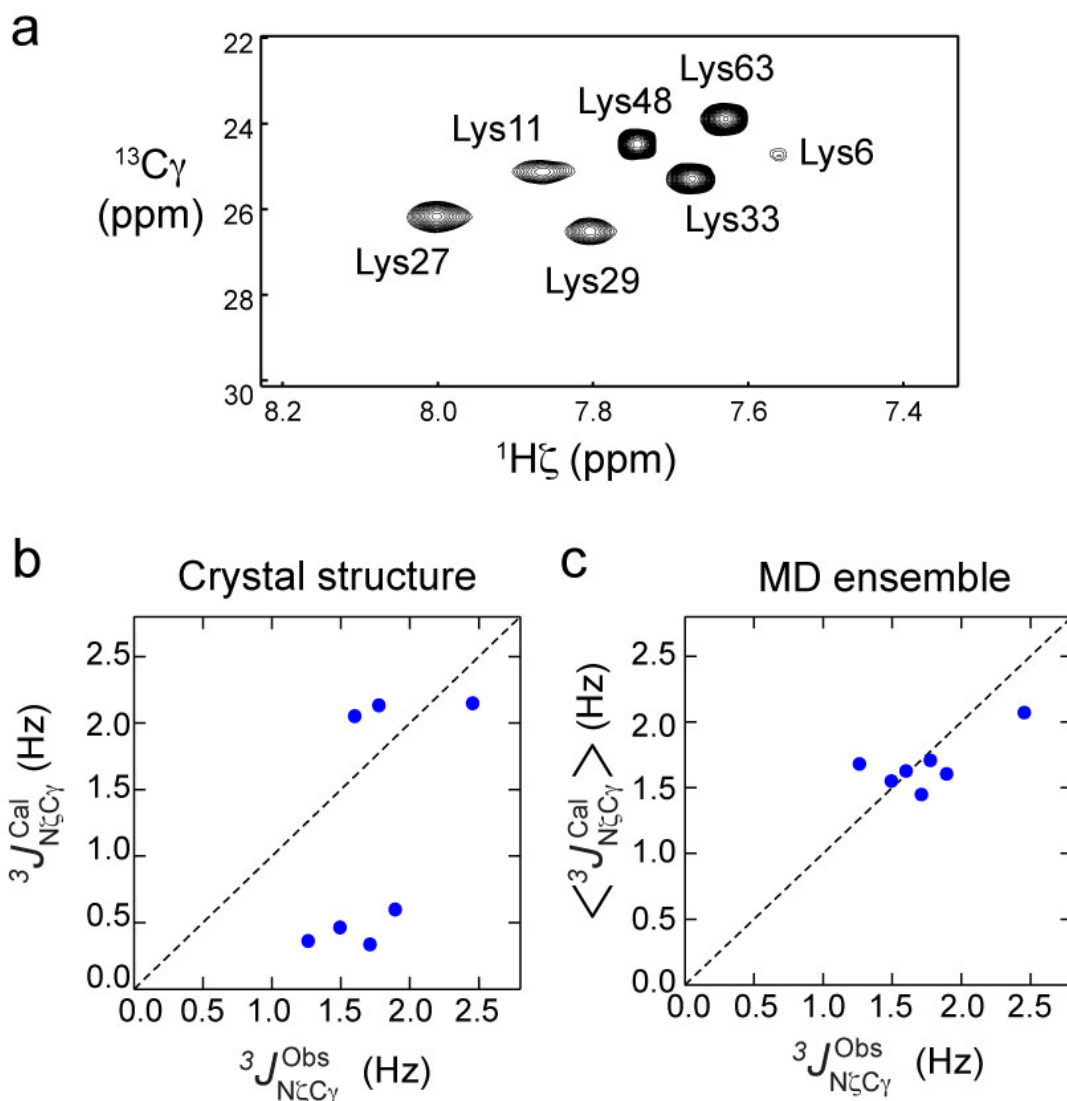
Figure 2.2 $^3J_{N\zeta C\gamma}$ analysis using NMR. (a) Long-range correlation spectra recorded with the pulse sequence shown in Figure 2.1a for lysine side-chain NH_3^+ groups in uniformly $^{13}C/^{15}N$-labeled ubiquitin. The ^{13}C IBURP-2 pulses were applied at 23 ppm. (b) Correlation between the observed $^3J_{N\zeta C\gamma}$ values and those calculated from the crystal structure (PDB entry 1UBQ). (c) Correlation between the observed $^3J_{N\zeta C\gamma}$ values and those calculated as ensemble averages for 1000 structures sampled every 1 ns from the 1 μs MD trajectory at 275 K.
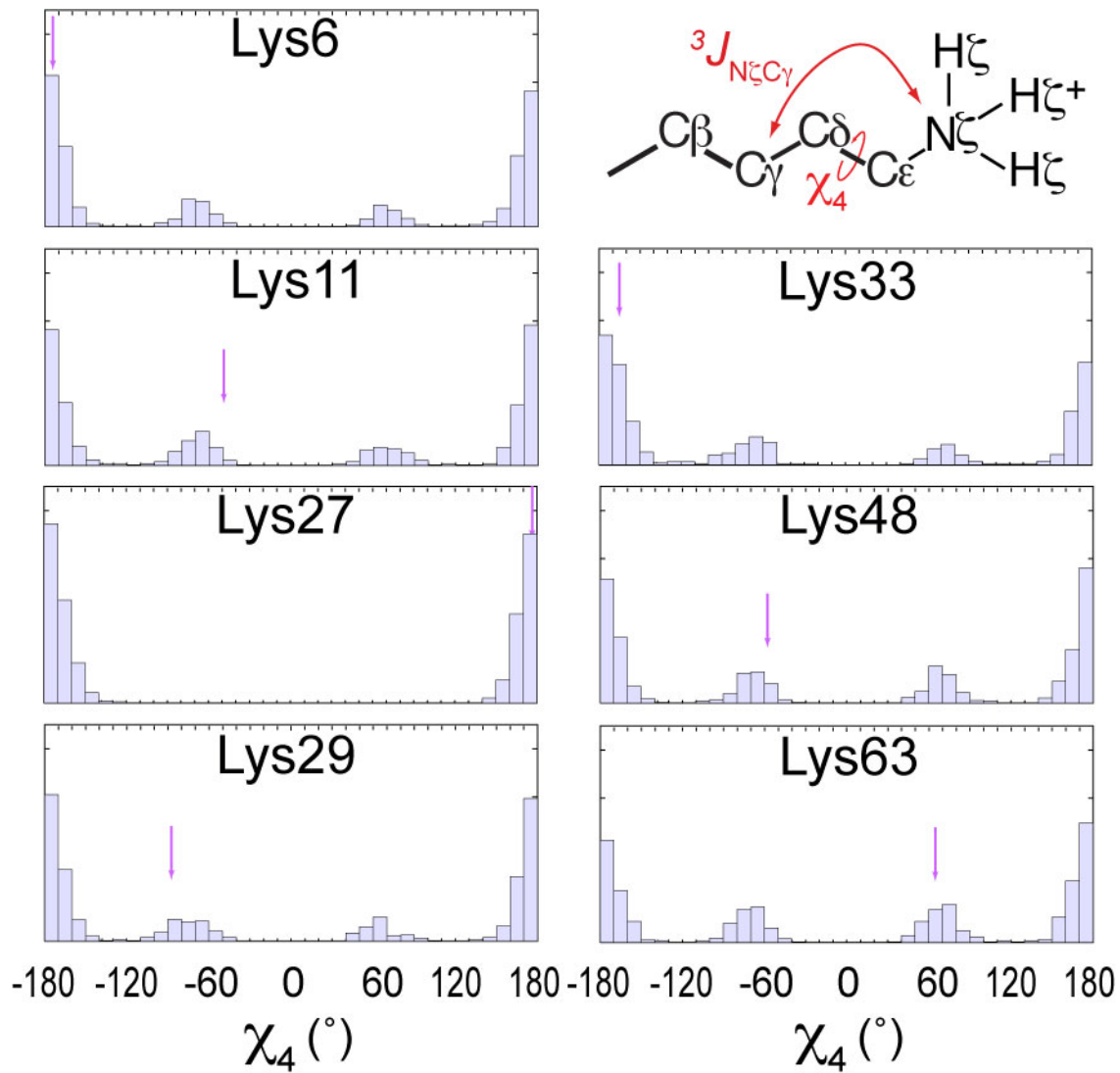
Figure 2.3   Histograms of the χ_4 torsion angles of individual lysine side chains of the MD ensemble. Purple arrows indicate the χ4 torsion angles in the crystal structure.

parameters are not available for $^3J_{N\zeta C\gamma}$ reporting on the torsion angle χ_4, here were used the experimental Karplus parameters for $^3J_{NC\gamma}$ belonging to χ_1 reported by Perez *et al.*,[63] according to whom (A = 1.29; B = -0.49; and C = 0.37). On the basis of these parameters, $^3J_{N\zeta C\gamma}$ values were back-calculated first from the χ_4 angles in the 1.8 Å resolution X-ray crystal structure of ubiquitin (PDB entry 1UBQ).[64] As shown in Figure 2.2b, a bimodal distribution around ~2 and ~0.5 Hz was found for the calculated values of $^3J_{N\zeta C\gamma}$ because the lysine χ_4 angles in the crystal structure are either trans or gauche. In contrast, the experimental $^3J_{N\zeta C\gamma}$ data do not exhibit such a bimodal distribution, and the agreement between the observed and calculated $^3J_{N\zeta C\gamma}$ data was poor, with a root-mean-square difference (rmsd) of 0.91 Hz (Figure 2.2b).

We also examined the ensemble averages $\langle\,^3J_{N\zeta C\gamma}\rangle$ calculated from the χ_4 angles sampled during a 1 μs MD simulation at 275 K using the GROMACS software package and the AMBER ff99SB force field. Simulations were performed by the group of Dr. Rafael Brüschweiler, who is our collaborator in this project.[40] To ensure a fair comparison, the simulation was run at the same temperature as in the experiment; otherwise, the MD simulation protocol was the same as described previously. The length of the MD simulation was expected to be adequate for this analysis, since the previous NMR relaxation study by Iwahara group showed that the reorientation of the symmetry axes of lysine NH_3^+ groups in ubiquitin occurs on the subnanosecond time scale. Values of $\langle\,^3J_{N\zeta C\gamma}\rangle$ were calculated from 1000 MD snapshots sampled every 1 ns. As shown in Figure 2.2c, the $\langle\,^3J_{N\zeta C\gamma}\rangle$ data from the MD ensemble exhibited substantially improved agreement with the experimental $^3J_{N\zeta C\gamma}$ data (rmsd = 0.26 Hz). The large difference between the $^3J_{N\zeta C\gamma}$ values calculated from the crystal structure and the $\langle\,^3J_{N\zeta C\gamma}\rangle$ values obtained from the MD ensemble is due to the wide distribution of χ_4 angles observed in the MD ensemble, which reflects extensive rotameric interconversions of this torsion angle (Figure 2.3). Hence, the agreement between the experimental $^3J_{N\zeta C\gamma}$ data and those
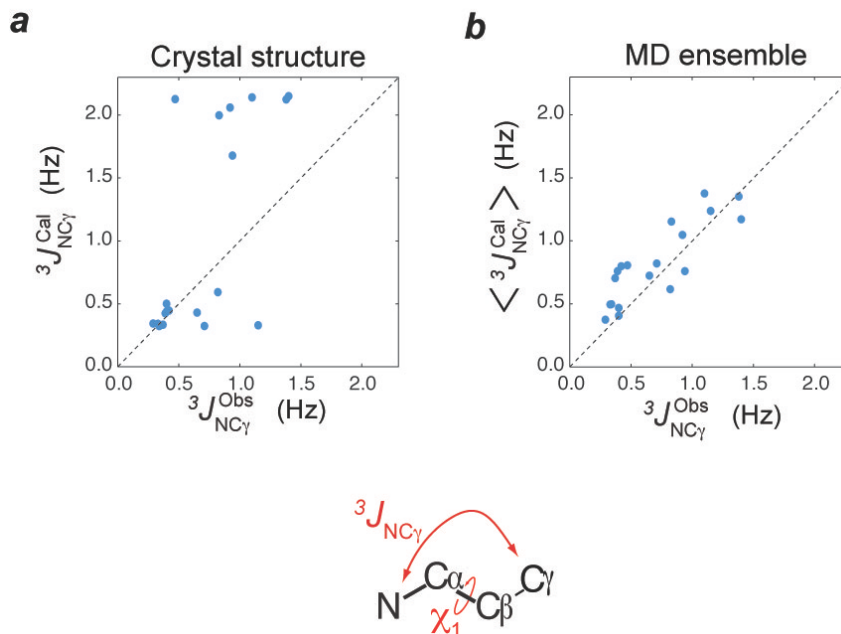
Figure 2.4. Impact of the dynamics on χ_1-related $^3J_{NC\gamma}$ coupling constants of Arg/Gln/Glu/Lys residues in ubiquitin. (a) Correlation between the experimental $^3J_{NC\gamma}$ data and those calculated from the crystal structure (PDB 1UBQ).[64] (b) Correlation between the experimental $^3J_{NC\gamma}$ data and those calculated as ensemble averages for the 1 μs MD trajectory at 300 K.

Table 2.1   Long-Range ^{15}N-^{13}C J-Coupling Constants $^3J_{N\zeta C\gamma}$ and $^{h3}J_{N\zeta C'}$ Measured for Lysine Side-Chain NH$_3^+$ Groups in Human Ubiquitin[a]

NH$_3^+$	$^3J_{N\zeta C\gamma}$	$^3J_{N\zeta C'}$
Lys6	1.78 ± 0.25	–
Lys11	1.89 ± 0.03	–
Lys27	2.45 ± 0.03	–
Lys29	1.26 ± 0.03	0.23 ± 0.03 (Glu16 C=0)[d]
Lys33	1.60 ± 0.01	0.17 ± 0.02 (Thr14 C=0)[d]
Lys48	1.49 ± 0.01	–
Lys63	1.71 ± 0.01	–

[a] The J-coupling constants were measured using spin-echo ^{15}N{^{13}C} difference experiments (Figure 2.1b) at a ^{1}H frequency of 800 MHz. Only absolute values are reported because the measurements did not provide signs.
[b] Measured three times to improve the precision. Averages and standard deviations are shown.
[c] Measured once. Uncertainties were estimated from the noise levels in the spectra.
[d] The acceptor of the hydrogen bond, which was identified in the long-range correlation spectrum (Figure 2.8a), is shown in parentheses.
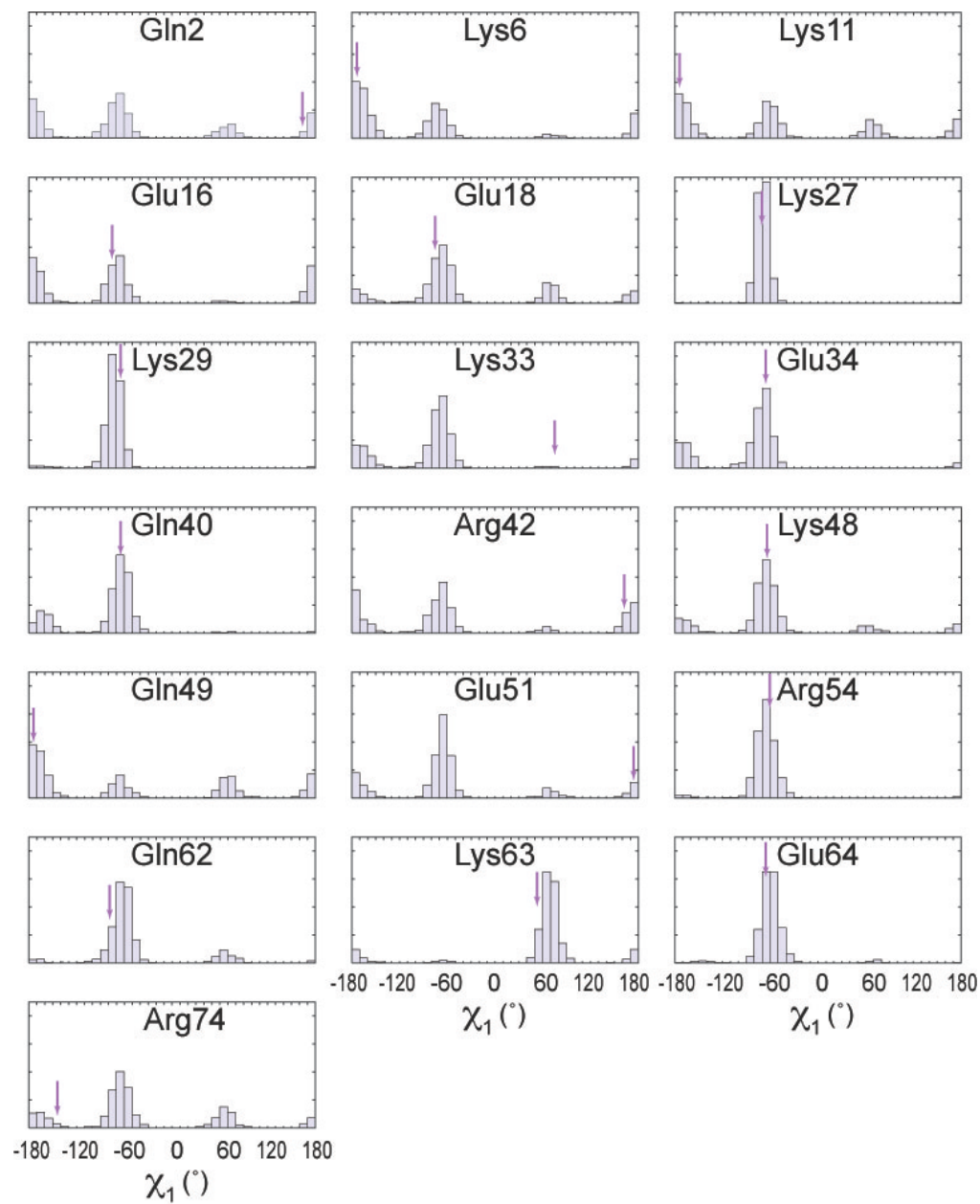
Figure 2.5   Histograms for the χ_1 angles of individual side-chains of the MD ensemble. The purple arrows indicate the values of χ_1 angles in the crystal structure.

predicted from the MD trajectory strongly indicates that the lysine χ_4 dihedral angles in ubiquitin are highly dynamic. This result is consistent with the low values of the order parameters found for most of the lysine C_ε-$N_\zeta H_3^+$ vectors in recent ^{15}N relaxation study from Iwahara Group.

2.3.2 Effects on the dynamics of the Cα-Cβ bond rotations on χ_1-related $^3J_{NC\gamma}$ couplings

While it has been well-established that internal motions changing the bond torsion angles directly affect the relevant vicinal ^3J-coupling constants, the derivation of quantitative dynamics information from ^3J-coupling data is not straightforward.[63,65-78] By far the easiest interpretation of ^3J-coupling data entails the extraction of structural information on relatively static portions of molecules. As demonstrated above, the combined use of a long MD simulation and ^3J-coupling data permits a direct consistency analysis of the torsion-angle dynamics of protein side chains. Using trajectory from a 1 μs MD simulation at 300 K from previous publication of Iwahara group,[40] we also examined the impact of the dynamics on the χ_1-related $^3J_{NC\gamma}$-coupling constants (through N-Cα-Cβ-Cγ) for ubiquitin using the experimental data at 303 K reported by Hu and Bax.[59]

As shown in the Figure 2.4, the experimental $^3J_{NC\gamma}$ data for arginine, lysine, glutamate, and glutamine residues exhibited far better correlation with those calculated from the MD ensemble than with those calculated from the crystal structure. This suggests that the χ_1 torsion angles of these solvent-exposed hydrophilic side chains are highly dynamic, as seen in the MD simulation (Figure 2.5). It should be noted that the $\langle\, ^3J\rangle$ values from the MD ensemble depend on the molecular mechanics force field used for the simulation. Hence, the degree of the agreement between the $\langle\, ^3J\rangle$ values from the MD simulation and the experimental ^3J-coupling constants is expected to be directly

useful for benchmarking and improving future MD force fields using NMR data for full-length proteins in their native environments.[79-82]

2.3.3 Direct fitting of the Karplus parameters to the MD ensemble

In chapters 2.3.1 and 2.3.2, the $^3J_{N\zeta C\gamma}$- and $^3J_{NC\gamma}$-coupling constants were calculated using the Karplus parameters of Perez *et al*. While the empirical parameters seem reasonably accurate (see below), the parameters can also be optimized by fitting to the MD ensemble as demonstrated by Markwick *et al*.[83] Here we show results of such direct fitting of the Karplus parameters for χ_1-related $^3J_{NC\gamma}$ and χ_4-related $^3J_{N\zeta C\gamma}$ coupling constants. Using an optimization routine of the MATLAB software (MathWorks, Inc), we optimized the Karplus parameters *A*, *B*, and *C* via minimization of the following target function:

$$ F = \sum_j \left\{ A\langle\cos^2\chi_j\rangle + B\langle\cos\chi_j\rangle + C - J_{obs,j}\right\}^2 / w_j^2 $$

where $\langle\rangle$ represents an ensemble average; J_{obs}, an observed *J*-coupling constant; and *w*, an inverse weight based on the experimental uncertainties. Figure 2.6 shows correlations between the experimental *J*-coupling constants and those calculated from the MD ensemble using the optimized Karplus parameters for χ_1-related $^3J_{NC\gamma}$ (panel a) and χ_4-related $^3J_{N\zeta C\gamma}$ (panel b) data. Compared to correlations shown in Figures 2.2c and 2.4b, the direct fitting procedures gave only marginal improvement of rmsd (0.22→ 0.18 Hz for $^3J_{NC\gamma}$; 0.26 → 0.25 Hz for $^3J_{N\zeta C\gamma}$), because, as shown in Figure 2.7, $\langle ^3J\rangle$ calculated with the optimized parameters agreed well with those calculated with the parameters of Peréz *et al*. These results suggest that the Karplus parameters determined by Peréz *et al*. are reasonably accurate for both $^3J_{NC\gamma}$ and $^3J_{N\zeta C\gamma}$.
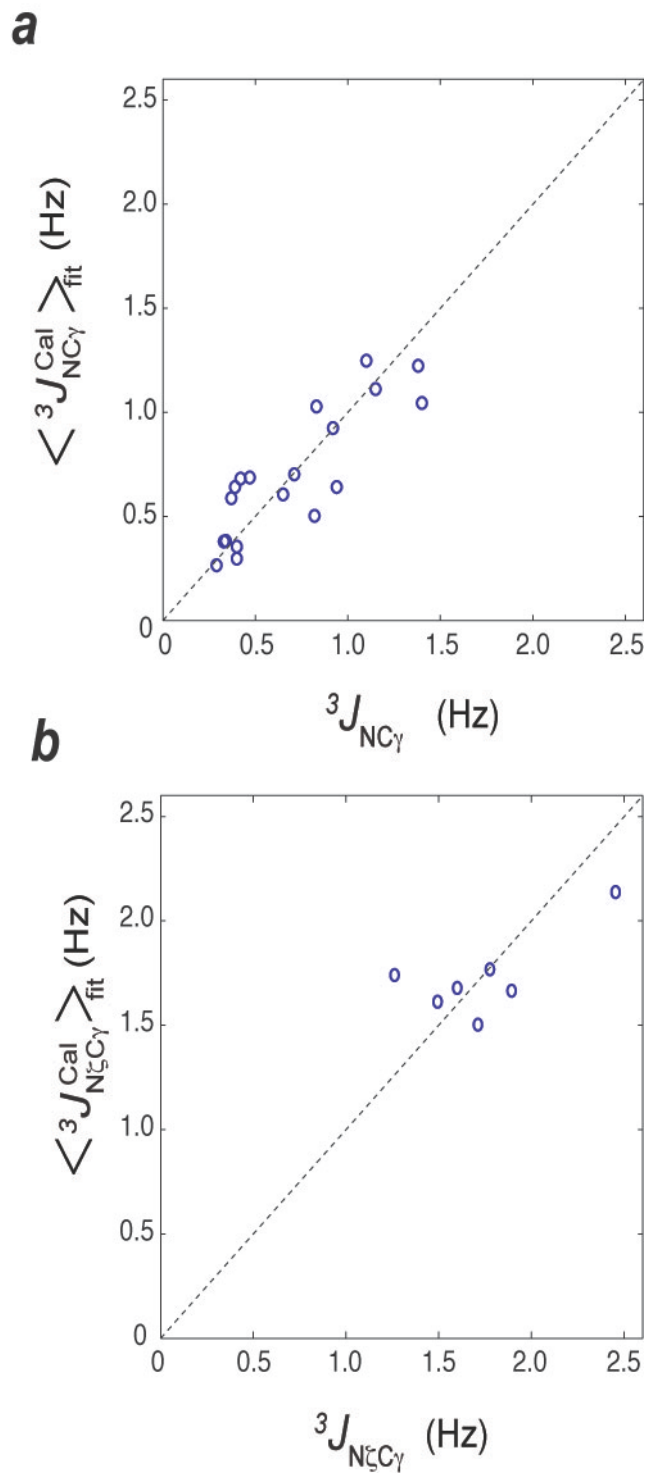
Figure 2.6 Correlation between experimental ^3J-coupling constants and ensemble averages $<^3J>$ calculated by direct fitting of Karplus parameters to the MD ensemble (a, $^3J_{NC\gamma}$; b, $^3J_{N\zeta C\gamma}$). The Karplus parameters were optimized to be A = 1.24, B = -0.49, and C = 0.27 for $^3J_{NC\gamma}$; A = 1.83, B = 0.09, and C = 0.29 for $^3J_{N\zeta C\gamma}$.
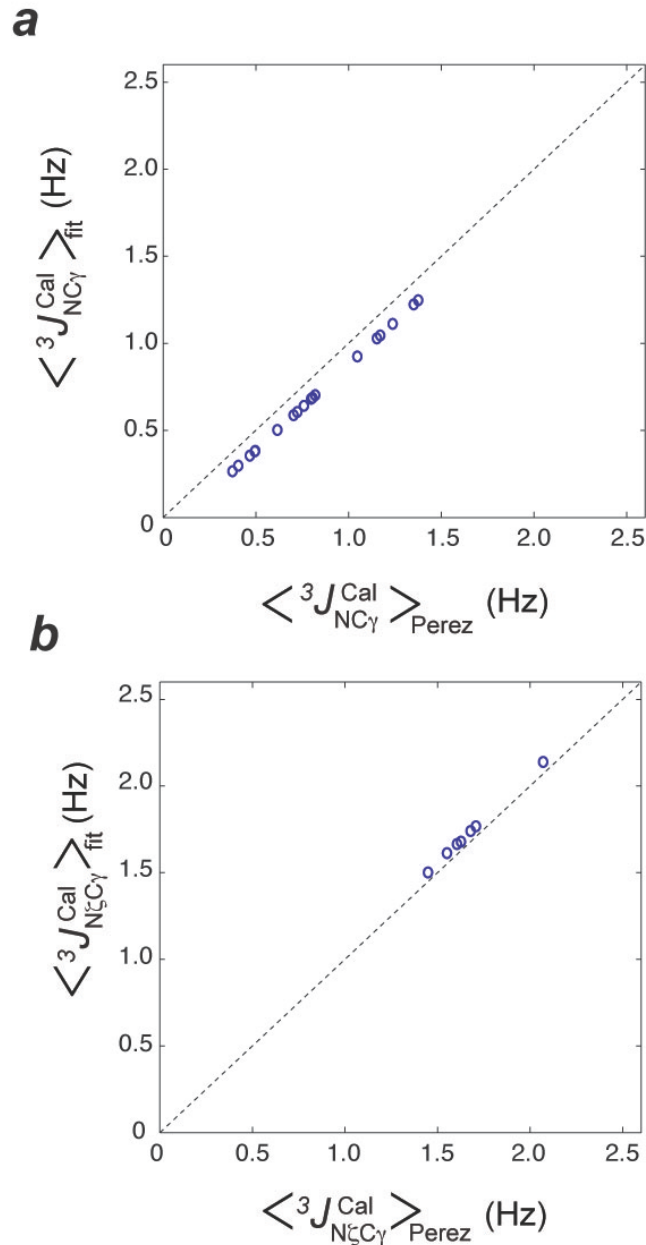
Figure 2.7 Correlation between ensemble averages $<^3J>$ calculated with the Karplus parameters optimized via direct fitting (see the caption for Figure 2.6) and those calculated with the parameters of Peréz et al (A = 1.29; B = -0.49; and C = 0.37). Panels a and b are for χ_1-related $<^3J_{NC\gamma}>$ and χ_4-related $<^3J_{N\zeta C\gamma}>$, respectively. Note that gauche conformers are major for the χ_1 angles of the Arg/Gln/Glu/Lys residues in ubiquitin (Figure 2.5), whereas trans conformers are major for the χ_4 angles of lysine (Figure 2.3). Due to this difference, the calculated $<^3J_{N\zeta C\gamma}>_{Perez}$ are generally larger than $<^3J_{NC\gamma}>_{Perez}$.

2.3.4 Measurement of $^{3h}J_{N\zeta C'}$ across hydrogen bonds involving lysine NH$_3^+$

In order to gain direct insight into hydrogen bonds involving the lysine side chains, we also studied ^{15}N-^{13}C ^3J-coupling constants $^{h3}J_{N\zeta C'}$ across hydrogen bonds involving lysine NH$_3^+$ groups of ubiquitin. Figure 2.8a shows the long-range correlation spectrum recorded for lysine NH$_3^+$ groups. Because of ^{13}C=O-selective IBURP-2 pulses that do not affect aliphatic ^{13}C nuclei, only NH$_3^+$ groups having sizable J-couplings with carbonyl or carboxyl ^{13}C nuclei across a hydrogen bond can give signals in this experiment. Signals for Lys29 and Lys33 NH$_3^+$ groups were clearly observed in the long-range correlation spectrum. The ^{13}C chemical shifts for the signals observed for the Lys29 and Lys33 NH$_3^+$ groups are in excellent agreement with those of backbone C=O groups of Glu16 and Thr14, respectively. In fact, these carbonyl groups are in close proximity to the lysine NH$_3^+$ groups in the crystal structure: the distances between the donor Nζ and acceptor O atoms are 2.8 Å for the Lys29-Glu16 pair and 3.5 Å for the Lys33-Thr14 pair. In the 1 μs MD simulation, the overall occupancies of the hydrogen bonds (Nζ-O distance < 3.5 Å) were 80% for the Lys29-Glu16 pair and 51% for the Lys33-Thr14 pair. The existence of the NMR signals arising from $^{h3}J_{N\zeta C'}$-coupling provides direct evidence for at least a part-time presence of these hydrogen bonds in solution.

To determine the values of $^{h3}J_{N\zeta C'}$, we carried out the spin-echo ^{15}N{^{13}C=O} difference constant-time ^{1}H-^{15}N correlation experiment for the lysine NH$_3^+$ groups, as shown in Figure 2.1b. The difference spectrum for the two subspectra a and b gives signals only if intensity modulation by $^{h3}J_{N\zeta C'}$ evolution is sizable. As expected from the result of the long-range correlation experiment, only the NH$_3^+$ groups of Lys29 and Lys33 exhibited signals in the difference spectrum (Figure 2.8b). The values of $^{h3}J_{N\zeta C'}$
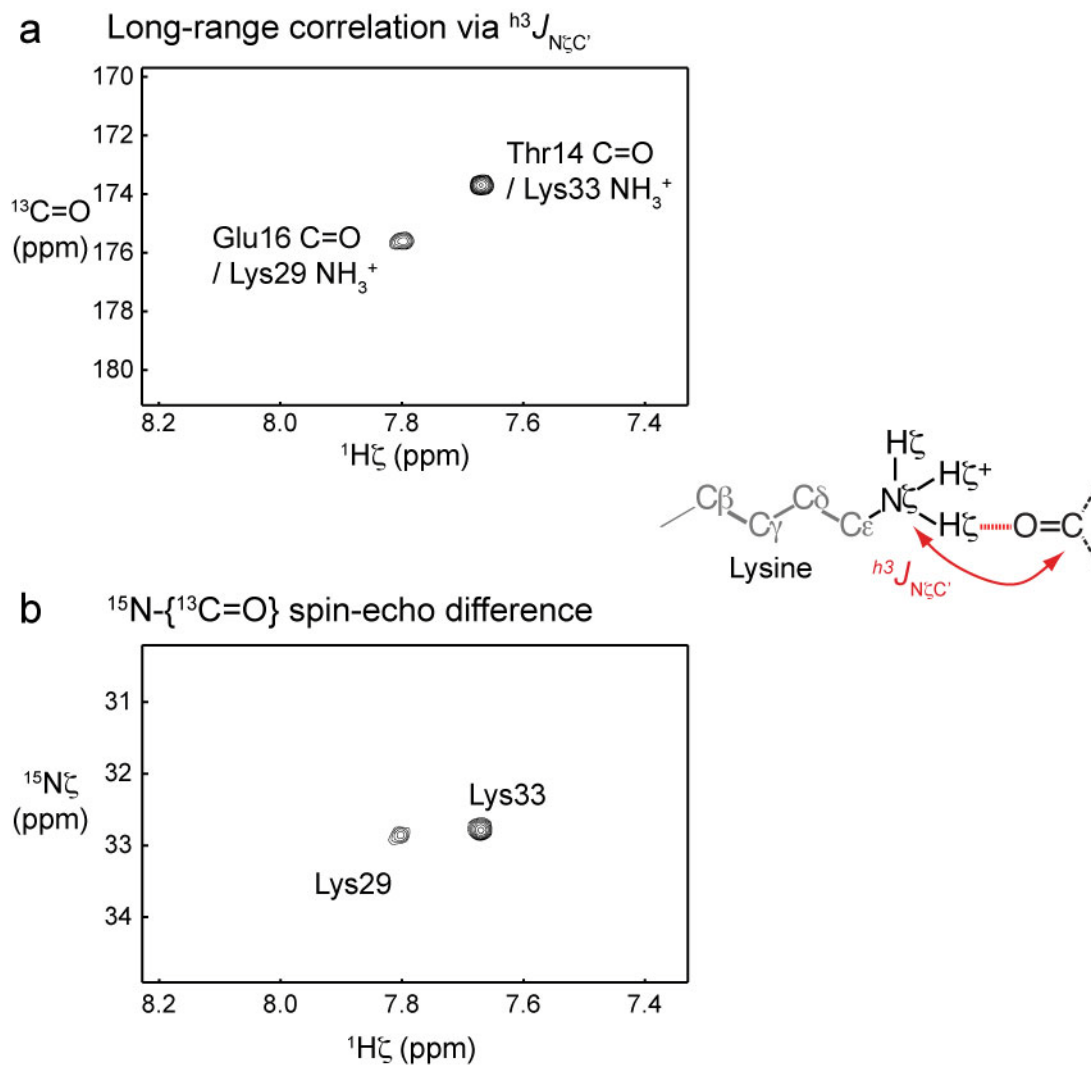
Figure 2.8. Direct Observation of hydrogen bonds via NMR. (a) Long-range correlations via $^{h3}J_{N\zeta C'}$-coupling across $NH_3^+ \cdots O=C$ hydrogen bonds. This spectrum was recorded using the pulse sequence shown in Figure 2.1a. (b) Spin-echo $^{15}N\{^{13}C=O\}$ difference spectrum for the two subspectra recorded with $^{13}C=O$-selective pulses at positions a and b in the pulse sequence shown in Figure 2.1B.

were determined to be 0.23 and 0.17 Hz for the Lys29-Glu16 and Lys33-Thr14 hydrogen bonds, respectively (Table 2.1). These values are relatively small in comparison with the $^{h3}J_{NC'}$ values for hydrogen bonds between backbone N-H and C=O groups, which can reach 0.9 Hz.[84-88] Since a scalar coupling across a hydrogen bond depends on the electronic configuration, the average geometry,[89-91] and internal motions,[92,93] quantitative interpretation of the observed $^{h3}J_{N\zeta C'}$ values for the NH_3^+ groups is not possible at this point. Nonetheless, the observed $^{h3}J_{N\zeta C'}$-couplings unequivocally indicate the part-time presence of hydrogen bonds between the NH_3^+ and C=O groups. Dynamic hydrogen bonding displayed by lysine side-chain NH_3^+ groups arises from the highly mobile χ_4 torsion angles combined with rapid rotation about the C3 symmetry axis along the Nζ-Cε bond, as depicted in Figure 2.9 for Lys29 and Lys33. In previous ^{15}N relaxation study by Iwahara group,[40] the order parameters for the NH_3^+ groups of the same amino acids indicated high mobility on the subnanosecond time scale (S^2_{axis} = 0.38 for Lys29 and 0.25 for Lys33). As shown in Figure 2.9, the MD simulation also showed a highly dynamic hydrogen-bonding process involving the side-chain NH_3^+ groups of Lys29 and Lys33. Together with the detected $^{h3}J_{N\zeta C'}$-couplings for Lys29 and Lys33, these data collectively suggest that the side-chain hydrogen bonds evidenced by detectable $^{h3}J_{N\zeta C'}$-couplings are of a highly mobile nature.

2.4 DISCUSSION

In this chapter we demonstrated the use of long-range ^{15}N-^{13}C three-bond J-coupling data for investigating the detailed dynamic behavior of lysine side chains in a protein. In conjunction with an extended MD simulation, our experimental data showed that lysine side-chains are extremely dynamic and should not be described with a single conformation observed in crystal structure. Instead, ensemble average should be
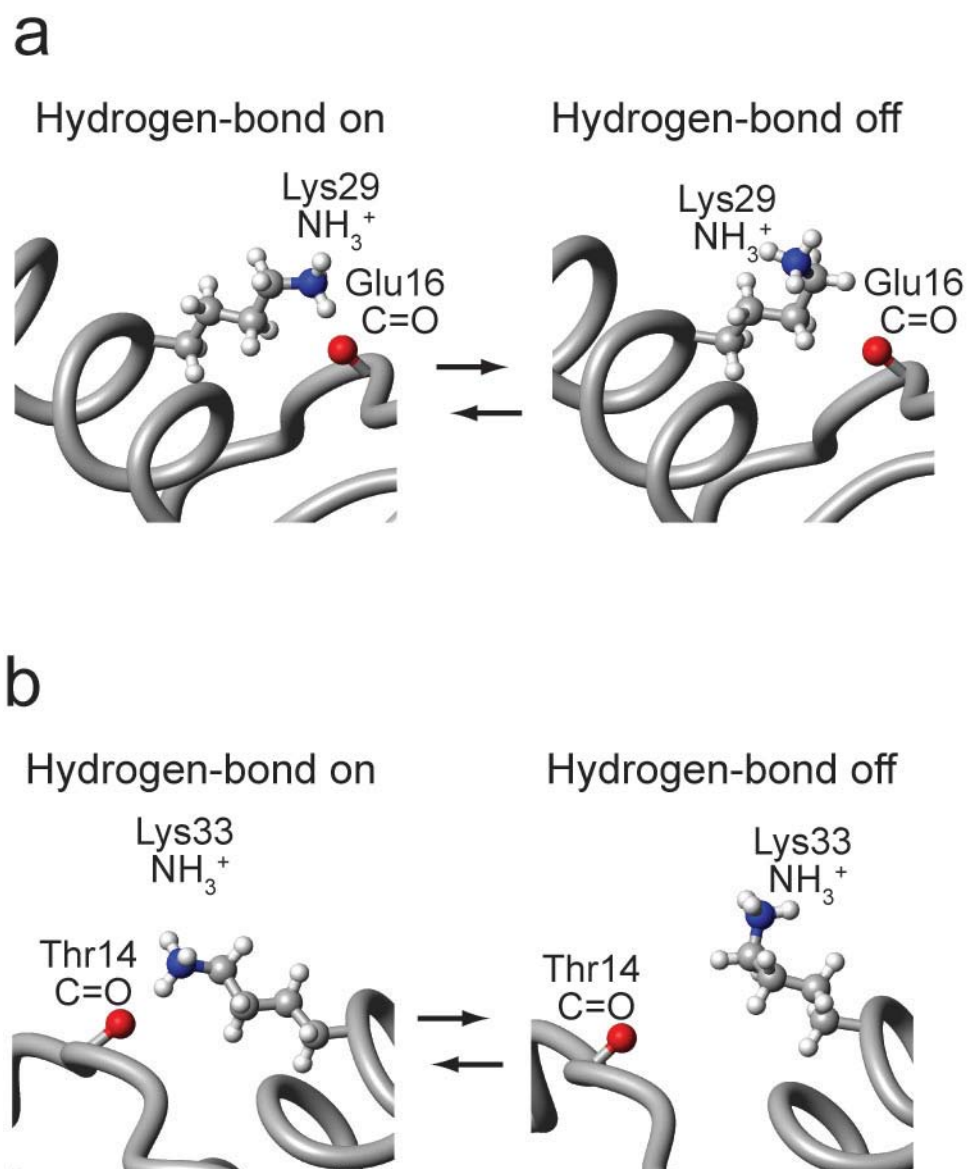
Figure 2.9. Dynamics of (a) Lys29-Glu16 and (b) Lys33-Thr14 hydrogen bonds seen in a 1 μs MD simulation at 275 K. The acceptors are backbone carbonyl groups. These hydrogen bonds were experimentally confirmed by means of measured $^{h3}J_{N\zeta C'}$ values (Figure 2.8). The average lifetimes of the hydrogen bonds in the MD simulation were 45 and 30 ps for Lys29 and Lys33, respectively.

considered for better interpretation of experimental data. Using our approach we also were able to measure *J*-coupling between Nζ of lysine side chain and C=O carbons for Lys29-Glu16 and Lys33-Thr14 pairs. The fact that we were able to detect magnetization transfer between these nuclei is direct evidence that these residues form hydrogen bonds, though relatively low values of these *J*-couplings (~0.2 Hz) suggest dynamic and transient nature of these bonds. This project provides a valuable information for understanding the relationship between protein function and dynamics, as amino acid side chain dynamics play critical role in intra- and intermolecular interaction. In future, this approach can be applied to elucidate dynamics of arginine guanidino groups.

CHAPTER 3:

Elucidation of dynamic mechanisms for target DNA search

3.1 INTRODUCTION

Over the decades, different groups have suggested a strong connection between protein dynamics and its functional kinetics.[5-9,14-17,21-24,26,29,33,34,94-96] Even though several papers have beautifully demonstrated such connection experimentally[8,15,21-23,35], this relationship remains poorly understood. Most of researchers focus on studying either protein dynamics, or functional kinetics alone, as investigating the relationship between two is experimentally challenging mostly due to experimental limitations. In this chapter we elucidate the role of domain dynamics of zinc-finger protein Egr-1 in its DNA target search kinetics and use our findings to improve target efficiency of artificial zinc-finger proteins. Our results suggest that domain dynamics play an important role in DNA scanning by zinc-finger proteins. Our findings could also provide valuable data for optimization of artificial zinc-finger nucleases, used in human gene therapy.

3.1.1 Egr-1: an inducible transcription factor

In cellular responses to various stimuli such as signals and stresses, gene regulation by transcription factors is of fundamental importance. Egr-1 (also known as Zif268) is an inducible transcription factor with crucial roles particularly in the brain and

cardiovascular systems in mammals. In the brain, Egr-1 is induced by synaptic signals in an activity-dependent manner and activates genes for long-term memory formation and consolidation.[97,98] In the cardiovascular system, Egr-1 is a stress-inducible transcription factor that activates the genes for initiating defense responses against vascular stress and injury.[99,100] Given the short lifetime of induced Egr-1 (typically ~2 h),[99] rapid gene activation by Egr-1 is important in these biological processes that require an immediate response to the stimuli.

The induced Egr-1 protein has to initiate its role by searching for its target DNA sites among billions of DNA base pairs in the nucleus. In the DNA scanning process, transcription factors need to discriminate their target sites from nonspecific sites based on relatively minor differences in DNA structure and sequence. Crystallographic studies demonstrated that Egr-1 recognizes its 9-bp target sequence, GCGTGGGCG, as a monomer via zinc finger domains 1, 2, and 3 (hereafter referred to as ZF1, ZF2, and ZF3) that contact 3 bp each (Figure 3.1a).[101] Although the three zinc finger domains are comprised of only 87 residues combined, Egr-1's interface with target DNA is as wide as 2,870 Å^2. Although such extensive contact with DNA is favorable for high specificity in target recognition as well as for high stability of the specific complex, it is unfavorable for rapid scanning of DNA because the protein has to break a larger number of interactions, such as hydrogen bonds and ion pairs, whenever it moves from one nonspecific DNA site to another.[102] This is problematic, particularly because the vast majority of Egr-1 molecules undergoing a target search should be bound to nonspecific DNA sites due to the extremely high density of DNA in the nucleus and Egr-1's micromolar affinity to nonspecific DNA. This situation represents the "speed-stability" paradox.[102,103]

How does Egr-1 achieve high specificity in recognition and rapidity in scanning? Even though there have been many structural studies on zinc finger proteins bound to their target DNA sites or free in solution,[101,104-109] this question remains unanswered

because very little is known about their binding to nonspecific DNA. In fact, there is a lack of knowledge on DNA scanning by eukaryotic transcription factors. The vast majority of the previous literature on the experimental studies of target DNA search is for prokaryotic gene-regulatory proteins or restriction enzymes;[103,110,111] however, because of the much higher DNA density and the presence of nucleosomes in eukaryotic nuclei, it is natural to consider that eukaryotic transcription factors may use a different strategy for their rapid target search. My present work shows that it is indeed the case. Here, we apply spectroscopic and computational methods to the highly dynamic complex in which Egr-1 is nonspecifically bound to DNA and perpetually changes its location on DNA. In this work WE provide structural, dynamic, and kinetic information on how Egr-1 efficiently scans DNA to find its target sites rapidly. Our results suggest that Egr-1's asymmetrical domain dynamics in the DNA-scanning process can play an important role in resolving the speed-stability paradox.

3.1.2 Zinc-finger nucleases (ZFNs)

Traditional gene therapy techniques have been plagued with problems of inaccurate targeting of undesirable genetic information and damaging otherwise functioning, non target sites.[112-114] These problems have been major hurdles that have prevented gene therapy techniques from successful integration into viable human treatment strategies. As an alternative, zinc-finger nuclease (ZFN) technology has been recently shown to be able to target specific genes.[115-122] ZFN technology is based on the work by Greisman and Pabo (Greisman, H.A. & Pabo, C.O. (1997) Science), where they described a new method to engineer artificial C_2H_2 type zinc-finger proteins with a desired target specificity.[123] To demonstrate their approach, they engineered protein with three zinc fingers which specifically binds to 3'-AAAATATCG-5' DNA sequence. They created library of zinc fingers where they randomized six amino acids located in the DNA

recognition helix. Their protocol consisted of three steps, each for selection of one of the zinc fingers from their zinc finger library. This method laid foundation to creation of artificial zinc-finger transcription factors used in biotechnology,[124] as well as ZFN technology.

ZFN proteins consist of artificial zinc fingers at the N-terminus and a FokI nuclease domain (ND) at the C-terminus.[115-122,125-128] A FokI ND alone can neither bind to, nor cleave DNA;[129,130] ZFN can cleave DNA only when two ZFN molecules bind to their recognition sites (separated by several base pairs) and two FokI nuclease domains are brought in close proximity to form a dimer. ZFN technology has been extensively used in gene therapy and gene manipulation in recent several years. The most notable of them is targeting CCR5 gene, which encodes for a T-cell surface receptor making the T-cells prone to HIV infection.[131] The use of ZFN for HIV treatment is currently in Phase 1/2 clinical trials.

Despite their high affinities for the target DNA sites, recent studies report kinetic defects of artificial zinc-finger proteins.[132-134] In fact, affinity-based selection of zinc fingers could be the contributing factor to ZFNs' low efficiency as it has strong bias towards stability.[135] While high specificity is good for DNA recognition, strong DNA binding can decrease ZFNs' translocation between nonspecific sites during target search that, as a result, decreases overall efficiency.[102,136] It has also been shown that excessive DNA-binding energy increases chances of nonspecific DNA cleavage, thus increasing toxicity of ZFN constructs.[137] In this project we applied knowledge which we gained while studying Egr-1's behavior in complex with DNA and its target search mechanism (Chapter 3.3.1-4) to create proteins that exhibit both rapid target search kinetics and high affinity for a target. We created ZFNs using wild-type and mutant Egr-1 proteins and FokI ND and investigated their DNA cleavage activities to find the most optimal conditions for the best ZFN efficiency.

3.2 MATERIALS AND METHODS

3.2.1 Preparation of Nonspecific and Specific Egr-1/DNA Complexes.

$^2H/^{15}N$- or $^2H/^{13}C/^{15}N$-labeled Egr-1 DNA-binding domain and its mutants were expressed in *E. coli* cultured in D_2O-based minimal media supplemented by 0.5 g/L ISOGRO (Sigma-Aldrich). Cell lysate with glutathione S-transferase (GST) fusion Egr-1 (wild-type or mutant) was loaded on GSTPrep FF column. Protein was washed with Tris-HCl (pH 7.5) buffer, containing 400 mM NaCl, 1% Triton X-100. After, protein was eluted with same buffer, containing 10 mM GSH instead of Triton. GST tag was cleaved off by overnight incubation at 4 ºC with 100 units of HRV-3C protease. Protein solution was concentrated down using 3,000 MWCO amicon ultra centrifugation tube and loaded on to a Sephacryl S100 26/60 column. Column was pre-equilibrated with 50 mM Tris-HCl (pH7.5), 1 M NaCl, 2 mM glutathione, and 0.2 mM $ZnCl_2$ buffer. Finally, fractions containing Egr-1 protein were buffer exchanged to 50 mM Tris-HCl (pH 7.0), 200 mM NaCl, 1 mM glutathione, 0.1 mM $ZnCl_2$, and 5% glycerol, and purified on Mono-S cation-exchange column. Protein was eluted using the same buffer using NaCl gradient from 200 mM to 500 mM. A plasmid for gene expression of the mutant proteins were made with the Quick Change Lightening site-directed mutagenesis kit (Agilent). The mutant protein was prepared in the same way as that for the wild type protein. Individual DNA strands for NMR studies were chemically synthesized and purified by anion-exchange chromatography. After annealing, 28-bp DNA duplexes were purified via anion exchange chromatography to remove excess single-stranded DNA. To prepare complexes, the Egr-1 protein was mixed with 28-bp DNA at a high ionic strength (~0.5 M NaCl), and the buffer was exchanged to reduce the ionic strength. NMR samples contained 0.8 mM complex, 10 mM d_{11}-Tris•HCl (pH 7.5), 20 mM KCl, and 7% D2O.

3.2.2 NMR Spectroscopy.

All NMR experiments for Egr-1/DNA complexes were performed at 35 °C. NMR experiments for Egr-1 in the free state were performed at 25 °C. The lower temperature for the free protein was necessary due to its lower stability. Backbone $^1H/^{13}C/^{15}N$ resonances of the specific and nonspecific Egr-1/DNA complexes and the free protein were assigned by HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, and HN(CO)CACB spectra.[138] ^{15}N relaxation rates R_1 data was measured at ^1H-frequencies of 600 and 800 MHz as previously described.[139] Kinetic rate constants for translocation of Egr-1 between two 28-bp nonspecific DNA duplexes NS28 and ZS28 were determined from HSQC spectra recorded independently for two different nonspecific complexes and their 1:1 mixture by using the NMR line-shape analysis as described previously.[140]

3.2.3 ZFN purification.

Pet-49b plasmids encoding wild-type, T23K/Q32E and T23K/Q32E/E60Q/K79T Egr-1/FokI genes were purchased from Integrated DNA Technologies (IDT). E60Q/K79T plasmid was created using QCL kit (Agilent). *E. coli* strain BL21(DE3) cells containing glutathione S-transferase (GST)-ZFN plasmids were cultured in minimal media at 37 ºC. At OD_{600}=0.4, temperature was decreased to 22 ºC. Protein expression was induced at OD_{600}=0.5-0.7 by addition of 0.7mM isopropyl β-D-1-thiogalactopyranosid and 200 µM $ZnCl_2$. After induction, cells were cultivated for 4-5 hours at 22 ºC. Harvested cells were re-suspended in ~100 ml buffer containing 50 mM Tris·HCl (pH 7.9), 0.5 M NaCl, 10% glycerol, 100 µM $ZnCl_2$, 1% Triton X-100 and two tablets of the EDTA-free protease inhibitor cocktail (Roche). Cells were disrupted by sonication at 4 ºC for 2.5 min. Cell lysate was centrifuged at 30000×g for 30 min at 4 ºC, followed by centrifugation of supernatant for additional 15 min. Supernatant was loaded on a GSTPrep FF column (GE Healthcare) at 4 ºC with the flow rate of <0.8 ml/min.

Column was washed using 150-200 ml of washing buffer, containing 50 mM Tris·HCl (pH 7.9), 0.5 M NaCl, 10% glycerol and 1% Triton X-100. Protein was eluted with 50-100 ml of elution buffer containing 50 mM Tris·HCl (pH 7.9), 0.5 M NaCl, 10% glycerol and 10 mM glutathione. GST tag was cleaved over two days at 4 ℃ using 100 units of HRV-3C protease (GE Healthcare). The cleavage was confirmed using SDS-PAGE. Reaction mixture was loaded onto HiTrap Heparin HP column (GE Healthcare) at 4 ℃ and washed with buffer containing 50 mM Tris·HCl (pH 7.9), 0.35 M NaCl, 2 mM β-ME, 100 µM $ZnCl_2$. Protein was eluted using the same buffer with concentration of NaCl at 0.5 M. Finally protein was buffer-exchanged to 20 mM Tris·HCl (pH 7.9), 0.5 M KCl, 2 mM β-ME, 200 µM $ZnCl_2$ and 10% glycerol and quantified using UV absorbance at 280 nm. For long term storage, protein solution was diluted using glycerol up to 50% glycerol concentration and stored at -70 ℃.

3.2.5 ZFN assay

Short palindromic primer containing two oppositely directed Egr-1 target sequences separated with 6 bp-s was inserted into pCR topo 2.1 plasmid using TA cloning kit. Product DNA was digested using BglII restriction enzyme. Digestion reaction was confirmed on 0.9% agarose TBE gel. Target DNA digestion by ZFNs was done in 20 mM Tris·HCl (pH 7.5), 150mM KCl, 5% glycerol, 2 mM β-ME, 1 mM $MgCl_2$ and 100 µM $ZnCl_2$ buffer with 1 nM target DNA, 37 ng/µl Calf Thymus (CT) DNA and 5 nM protein concentration at room temperature, unless noted otherwise. DNA digestion level was determined by analyzing bands on 0.9% agarose TBE gel. DNA was visualized using EZ-vision dye (Amresco).

3.3 RESULTS

3.3.1 NMR of Specific and Nonspecific DNA Complexes.

In our studies, we used the Egr-1 DNA-binding domain comprised of three zinc fingers (Egr-1 residues 349–421; hereafter, referred to as "Egr-1" for simplicity's sake) and two 28-bp DNA duplexes: SP28 and NS28 (Figure 3.1b). These two duplexes are identical except that SP28 contains a 9-bp target site (GCGTGGGCG) whereas the DNA duplex NS28 does not. The nonspecific 28-bp DNA NS28 does not contain any bp triplets recognized by individual domains. These 28-bp DNA duplexes were chosen for our experiments earlier studies in our lab showed that Egr-1 behaves well in complex with these DNA for NMR and fluorescence studies. It was also shown that affinity for nonspecific DNA is ~1,000-fold weaker than the affinity for target DNA under the same conditions (40 mM KCl). If an Egr-1 protein covers 11 bp, as observed in the crystal structure, the 28-bp nonspecific DNA can accommodate up to two protein molecules. To avoid complications from two proteins binding to the same DNA molecule, we studied the nonspecific complex using at least a two-fold molar excess of nonspecific DNA. Under these conditions, the population of the singly bound complexes should exceed 90%. Using TROSY-based triple resonance experiments with deuterium decoupling,[138] we assigned the $^1H/^{13}C/^{15}N$ resonances of the protein backbone for the nonspecific and specific complexes as well as for the free protein. Figure 3.1c shows chemical shift perturbation (CSP) upon DNA-binding for each Egr-1 amide group (Numbering of residues is according to previous literature).[101] CSP magnitudes were smaller for the nonspecific complex, though K_d data together with molecular concentrations suggest that >99% protein should be bound to DNA for both complex samples. Interestingly, ZF1's CSP upon binding to the nonspecific DNA duplex was particularly small (only two residues exhibit a weighted-average CSP larger than 0.1 ppm); whereas, all ZF domains,
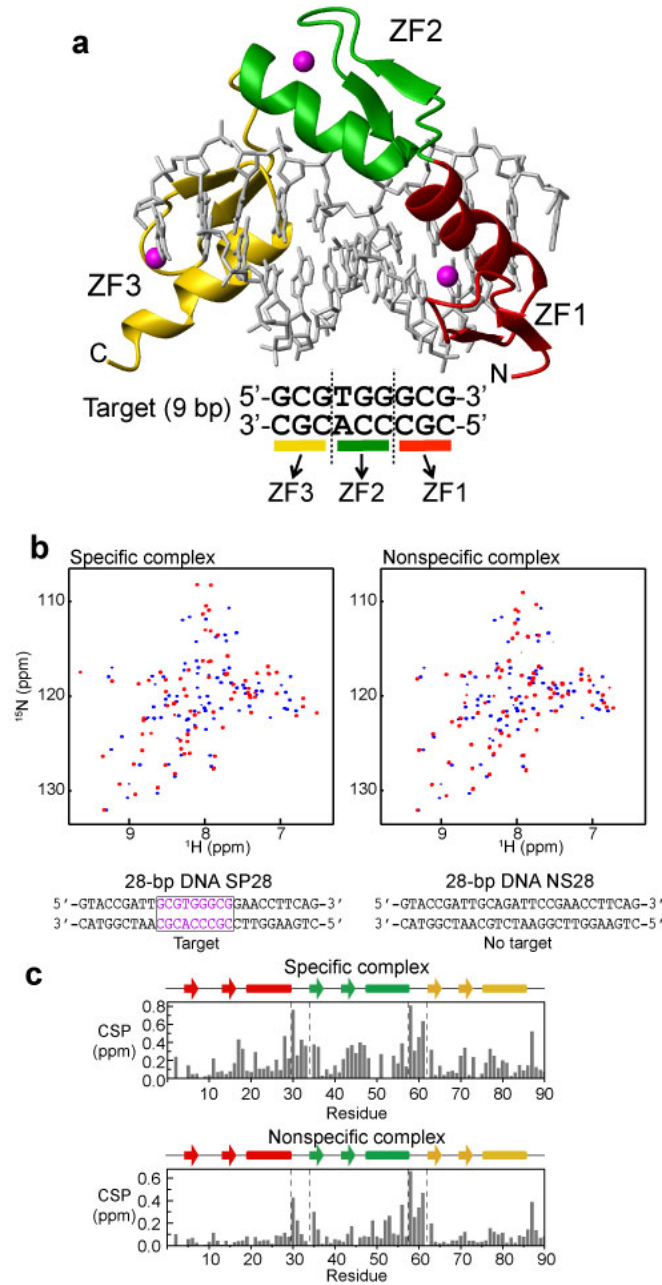
Figure 3.1   Specific and nonspecific DNA complexes of Egr-1. (a) Crystal structure (PDB 1AAY) of a complex of Egr-1 with a target DNA.[101,104] (b) ^{1}H-^{15}N TROSY spectra recorded for ^{2}H/^{15}N-labeled Egr-1 proteins in the free state (blue) and the complexes (red) with the specific (SP28) or nonspecific (NS28) DNA duplexes. (c) Weighted-average chemical shift perturbation (CSP) as defined $\{(|\Delta\delta_H|^2 + |0.2\Delta\delta_N|^2)/2\}^{1/2}$, where $\Delta\delta_H$ and $\Delta\delta_N$ represent chemical shift differences of backbone amide ^{1}H and ^{15}N, respectively, between the free and DNA-bound states.

including ZF1, exhibit substantial chemical shift perturbations upon binding to the specific DNA duplex (Figure 3.1c). This implies that Egr-1's binding mode for nonspecific DNA is quite different from that for specific DNA.

3.3.2 ZF1's Local Dissociation in the Nonspecific Complex.

To obtain structural and dynamic information on Egr-1 bound to nonspecific DNA, we conducted a comparative study of the nonspecific and specific complexes by analyzing ^{15}N longitudinal relaxation rates (R_1), ^{15}N transverse relaxation rates (R_2), and heteronuclear ^{1}H–^{15}N nuclear Overhauser effect (NOE) (Figure 3.2) for protein backbone amide groups in the specific and nonspecific Egr-1/DNA complexes. Comparison of heteronuclear NOE data for the specific and nonspecific complexes to those for the free protein (Figure 3.2) indicated that Linker 1 (between ZF1 and ZF2) in the nonspecific complex is as mobile as that in the free state, whereas Linker 2 in the nonspecific complex and both Linkers in the specific complex are immobilized upon complex formation. Furthermore, ZF1 in the nonspecific complex exhibited ^{15}N R_1 rates higher than those of ZF2 and ZF3 of the same complex (Figure 3.2b), whereas ^{15}N R_1 rates are similar for all the three zinc finger domains in the specific complex (Figure 3.2a). ^{15}N R_2 rates for many residues in the nonspecific complex were found to be substantially higher than those for the corresponding residues in the specific complex. CPMG relaxation dispersion data suggest that this is due to slow dynamics on a μs–ms timescale. The bottom graphs in Figures 3.3a and 3.3b show differences between apparent R_2 rates measured at CPMG field strengths V_{CPMG} of 33 Hz and 667 Hz [$\Delta R_2^{CPMG}=R_2(33$ Hz$)-R_2(667$ Hz$)$] for the complexes. Corresponding data for the free Egr-1 protein are shown in Figure 3.3c. From comparison of these CPMG data, it is clear that far more residues exhibit $\Delta R_2^{CPMG} > 5$ s^{-1} for the nonspecific complex. ΔR_2^{CPMG} can be large in the presence of μs–ms dynamics involving multiple states with different ^{15}N chemical shifts.
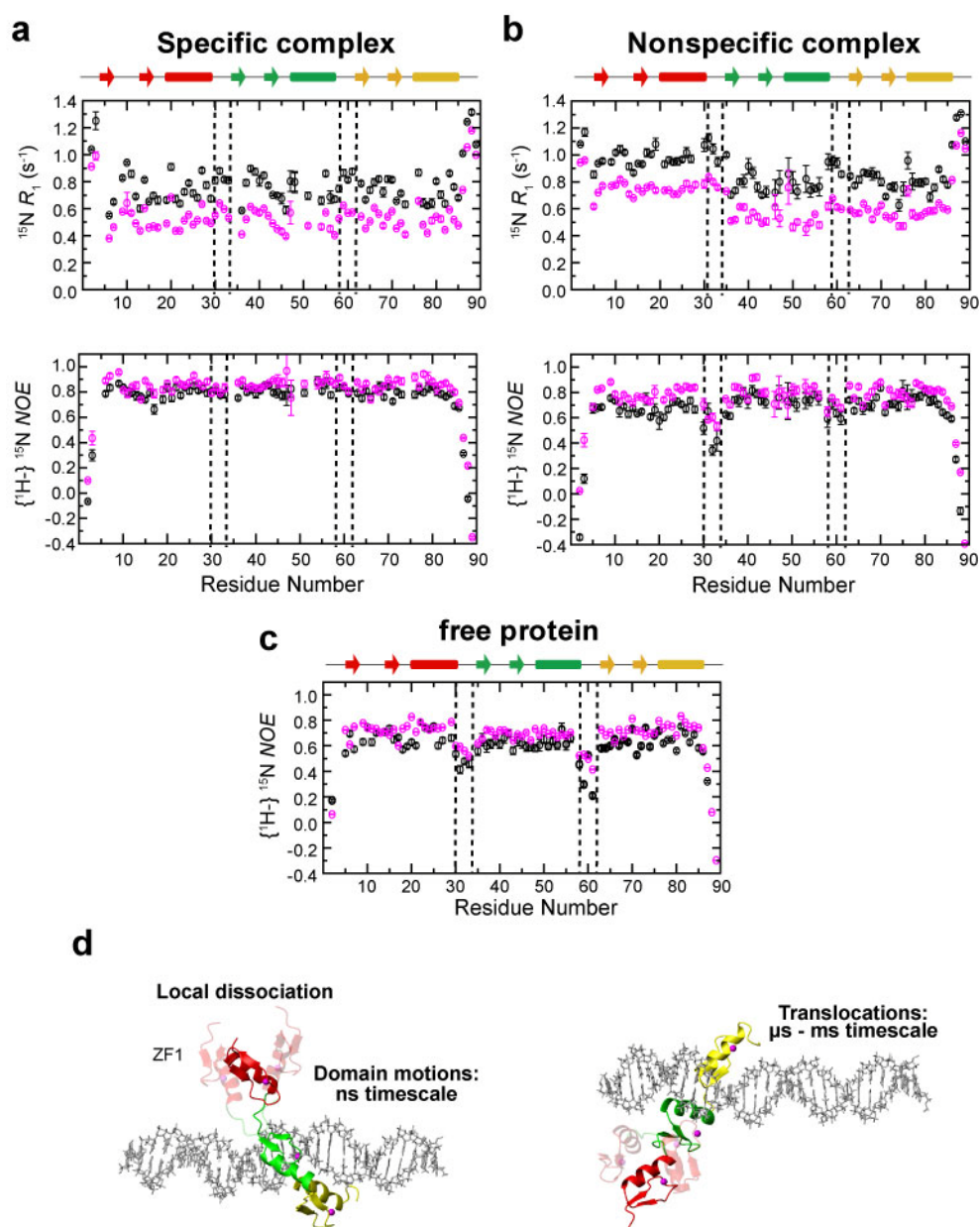
Figure 3.2. NMR relaxation experiments for Egr-1/DNA complexes. (a,b) Backbone ^{15}N R_1 and heteronuclear $\{^1\text{H-}\}^{15}\text{N}$ NOE for the specific (a) and nonspecific (b) complexes. (c) $\{^1\text{H-}\}^{15}\text{N}$ NOE for free protein. (d) Domain motions and translocation of Egr-1 in the nonspecific complex as suggested by the current study.
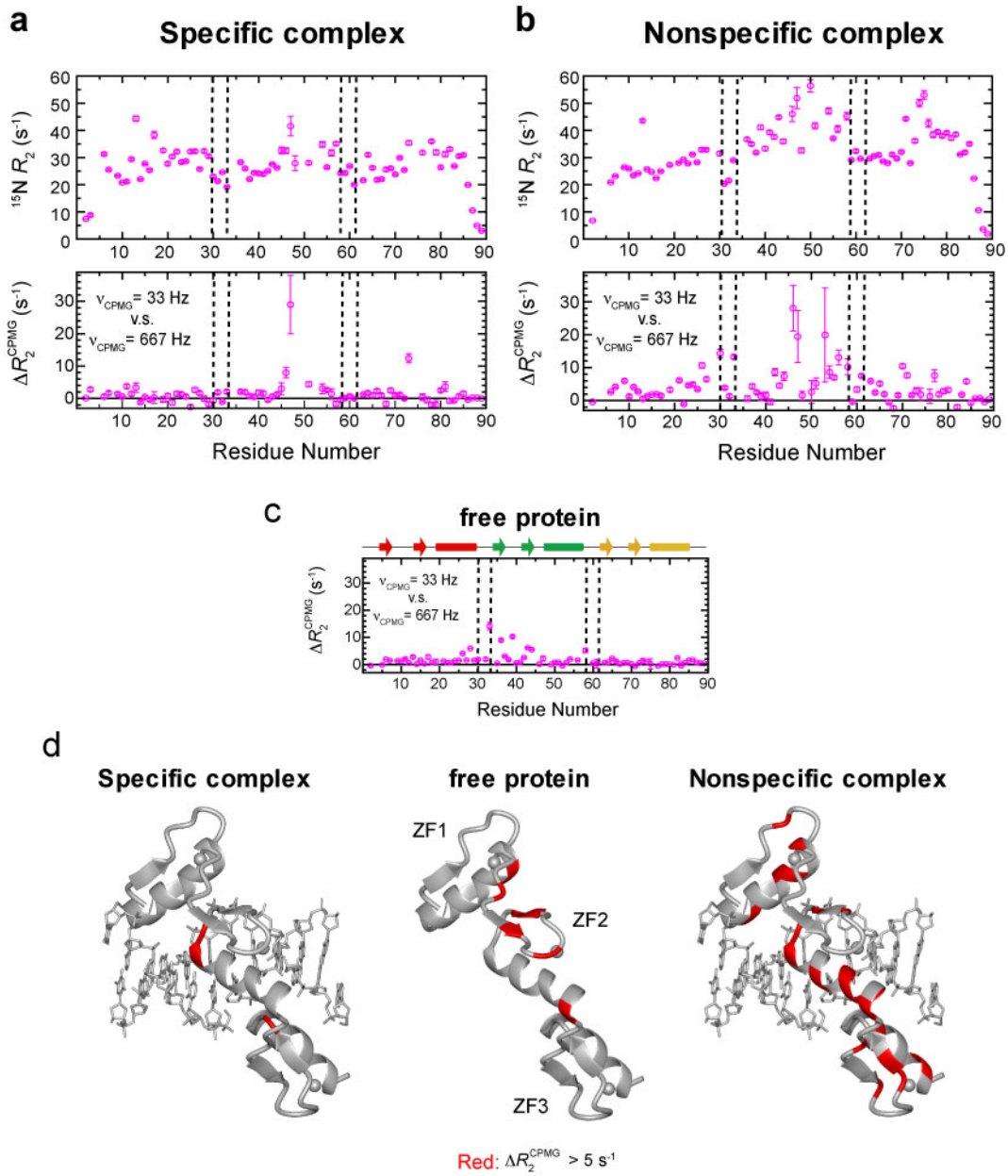
Figure 3.3 ^{15}N R_2 and CPMG relaxation experiments for the free and DNA bound Egr-1. Differences between apparent R_2 rates at CPMG frequencies ν_{CPMG} of 33 Hz and 667 Hz was defined as [ΔR_2^{CPMG} = R2(33 Hz) - R2(667 Hz)]. (a, b) ^{15}N R_2 and ΔR_2^{CPMG} for the (a) specific and (b) nonspecific DNA/Egr-1 complexes. (c) ΔR_2^{CPMG} for free Egr-1. (d) Mapping of residues that exhibited $\Delta R_2^{CPMG} > 5$ s^{-1} (red). Data are mapped on the crystal structure of the specific complex (PDB, 1AAY).

Importantly, most residues with $\Delta R_2^{\text{CPMG}} > 5 \text{ s}^{-1}$ were found to be close to DNA for the nonspecific complex (Figure 3.3d). It should be noted that ^{15}N chemical shifts of such residues near DNA could significantly change when the Egr-1 protein transfers from one nonspecific site to another on DNA via inter- or intra-molecular translocation. As shown below, intermolecular translocations of the Egr-1 protein between two nonspecific DNA duplexes were found to occur with an exchange rate constant of $\sim$1,000 s^{-1}. The intra-molecular translocation should be faster than the intermolecular translocation, as single molecule biophysical studies demonstrated that sliding of proteins on DNA occurs with a 1D diffusion coefficient in a range of 10^5–10^7 bp^2 s^{-1} in general.[103,141] Considering these together, it is likely that the μs–ms dynamics detected for the nonspecific complex correspond to translocation of protein on DNA rather than to intrinsic conformational dynamics; however, it is difficult to obtain further information from CPMG relaxation dispersion data because popular analytical expressions for a two-state exchange[142,143] are obviously inappropriate for the nonspecific complex that involves potentially 36 different states (18 binding sites and two opposite orientations for each).

In addition to R$_1$, R$_2$ and {^{1}H-^{15}N} NOE data, specific and nonspecific Egr-1/DNA complexes were analyzed using Residual Dipolar Coupling (RDC ^{1}D$_{\text{NH}}$) analysis by other members of Iwahara group. RDC data was in excellent agreement with other NMR experiments, indicating ZF1's unique behavior in the nonspecific complex. Taken together, these data clearly show that in the nonspecific complex with DNA, ZF1 domain is rather flexible and performs dynamic motions, while ZF2 and ZF3 domains are attached to DNA and act as anchors.

3.3.3 Modulation of Egr-1's Domain Motions in Nonspecific Complex.

High degree of domain motions for ZF1 in the nonspecific complex can be
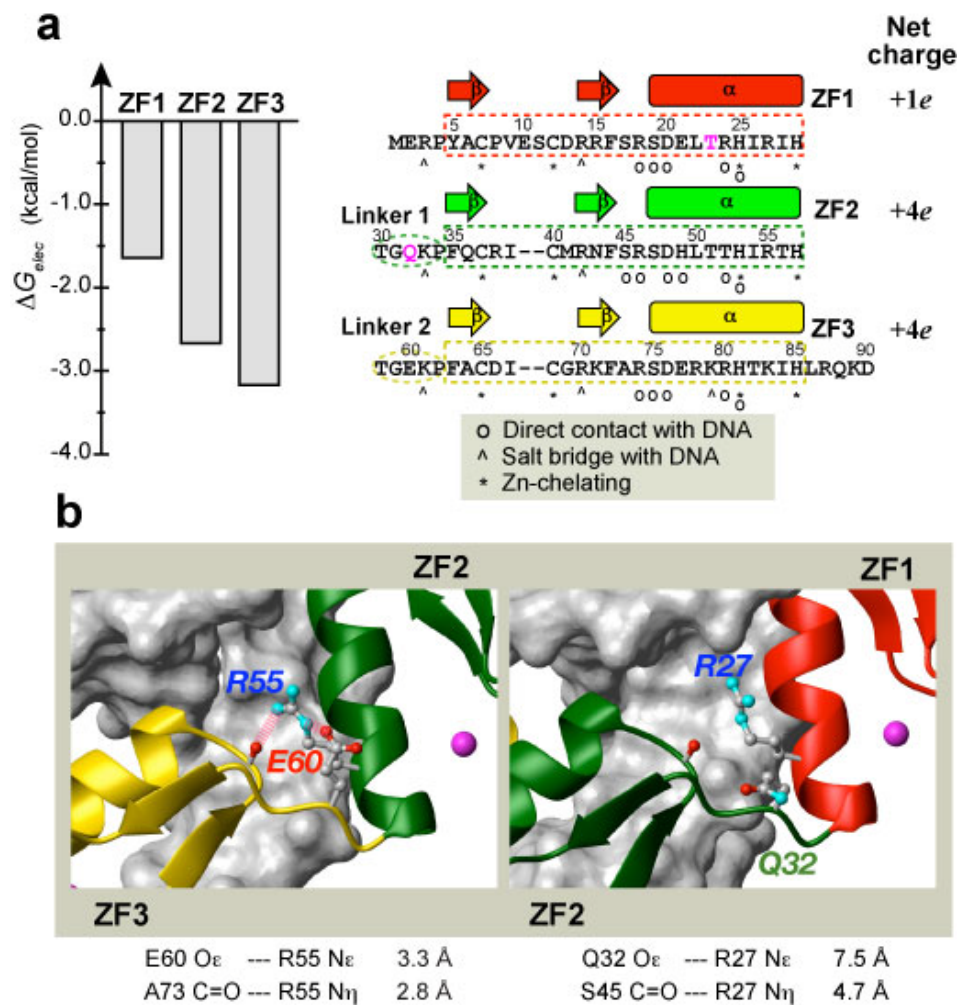
Figure 3.4   Determinants of ZF1's domain motions in Egr-1 bound to nonspecific DNA. (a) Electrostatic binding free energies for Egr-1's three zinc finger domains calculated using the Adaptive Poisson-Boltzmann Solver (APBS) software package for each domain in the crystal structure of the specific complex (PDB 1AAY). The energy calculation was carried out by Y. Levi's group as described previously[144]. The amino acid sequence and net charges of the three zinc-finger domains are also shown. Residues T23 and Q32 are highlighted in magenta. (b) Difference between interdomain interfaces in the crystal structure of the Egr-1/target DNA complex. Two hydrogen bonds and an ion pair stabilize interdomain interactions between ZF2 and ZF3; whereas, there is no corresponding stabilization between ZF1 and ZF2.

ascribed to ZF1's weaker interactions with DNA and with ZF2. Electrostatic binding free energy calculations (Figure 3.4a) suggested that among the three zinc finger domains, ZF1 is the weakest DNA-binder with the lowest absolute value of electrostatic binding free energy (−1.6 kcal/mol as opposed to −3.2 kcal/mol for ZF3). Qualitatively, this is due to the smallest net charge of ZF1 (Figure 3.4a). ZF1's weak interdomain interaction with ZF2 can be another determinant (Figure 3.4b). In all of 12 crystal structures available for wild type Egr-1 bound to its target DNA and for engineered Egr-1 proteins bound to their target DNAs,[101,104,145-148] the interface between ZF2 and ZF3 involves two hydrogen bonds supported by a salt bridge between R55 (ZF2) and E60 (Linker 2); however, the corresponding interdomain interactions are absent between ZF1 and ZF2 presumably because there is no corresponding electrostatic stabilization between R27 (ZF1) and Q32 (Linker 1). To examine whether or not these features make ZF1 highly dynamic in Egr-1 bound to nonspecific DNA, we investigated the domain motions of the T23K/Q32E mutant protein. T23K mutation was intended to enhance ZF1's electrostatic interaction with DNA by introducing an additional lysine/phosphate salt bridge found at the corresponding position in ZF3 (see Figure 3.4a) whereas Q32E mutation was intended to enhance interdomain interaction between ZF1 and ZF2 by introducing a salt bridge corresponding to that found at the ZF2-ZF3 interface (see Figure 3.4b). We measured backbone ^{15}N R_1 relaxation rates for the T23K/Q32E mutant protein bound to the 28-bp nonspecific DNA NS28 (Figure 3.5b). Although ZF1 in the wild type nonspecific complex exhibited clearly elevated R_1 rates (Figure 3.2b), all zinc finger domains in the T23K/Q32E mutant nonspecific complex exhibited R_1 rates in the same range as that for ZF2 and ZF3 in the wild type nonspecific complex. These results suggest that the T23K/Q32E double mutations indeed suppress ZF1's domain motions in the nonspecific complex to a level comparable to those of ZF2 and ZF3. To further modulate Egr-1's domain dynamics, we designed and created T23K/Q32E/E60Q/K69T

58
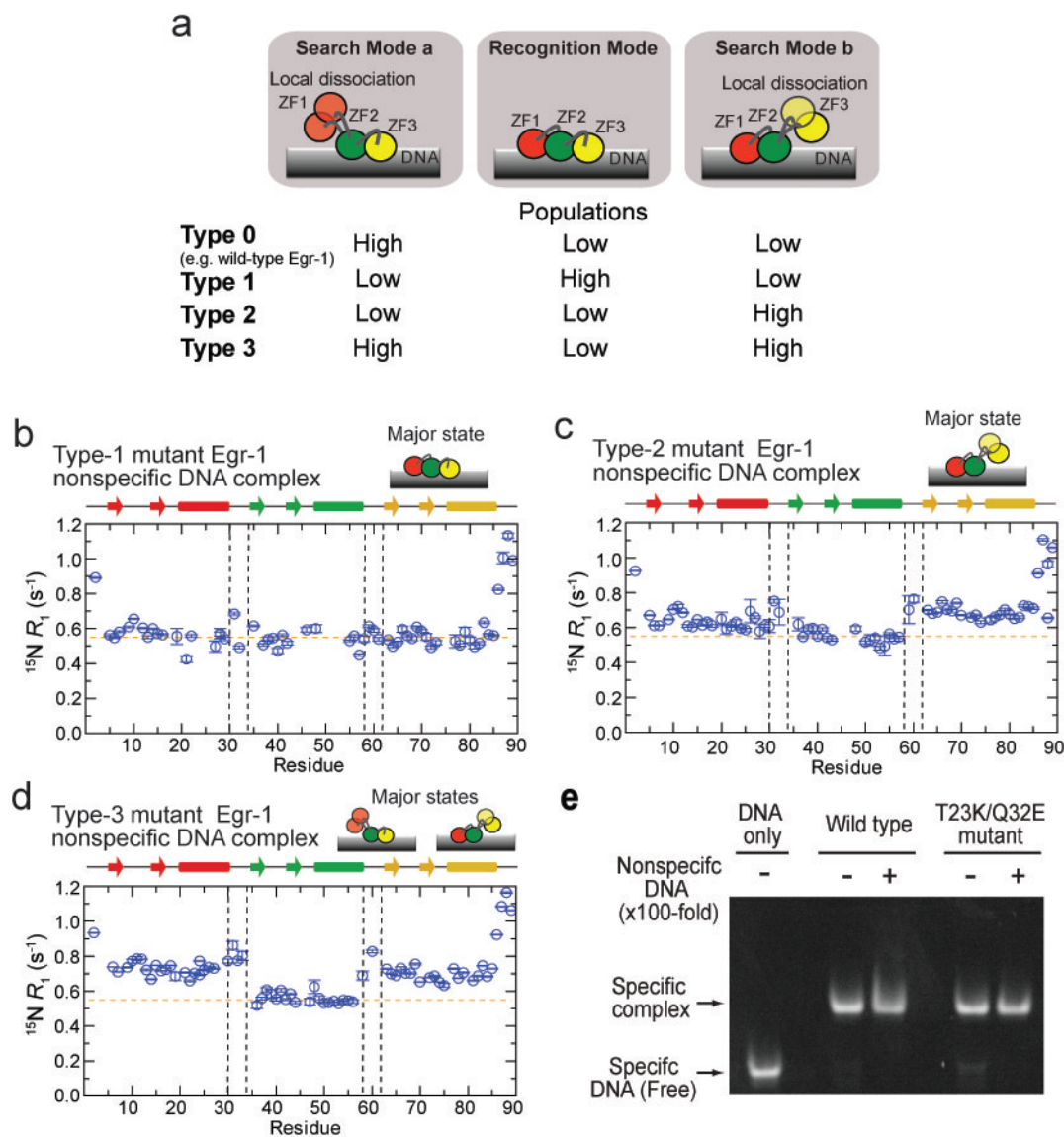
Figure 3.5   Impact of mutations on Egr-1's domain flexibility. (a) Four types of proteins used in this project. (b, c, d) ^{15}N R_1 profiles for (b) Type-1, (c) Type-2 and (d) Type-3 Egr-1 proteins. Different level of ^{15}N R_1 indicates altered domain flexibility in these mutants. (e) Gel based competition assay for wild-type and Type-1 mutant of Egr-1. Shown results indicate that the Type-1 mutant retains affinity and specificity to the target DNA sequence.

and E60Q/K79T mutants of Egr-1. Additional mutations at E60 and K79 positions were introduced in order to make ZF3 domain more mobile. Specifically, E60Q mutation should disrupt hydrogen bonding network between domains 2 and 3, and K79T mutation was introduced to remove the salt bridge between ZF3 and DNA at corresponding position. Therefore, we expected in T23K/Q32E/E60Q/K69T mutant ZF1 domain to show no flexibility at all and ZF3 domain to be more flexible, while in E60Q/K79T mutant both ZF1 and ZF3 domains would be flexible. And indeed, T23K/Q32E/E60Q/K69T mutant exhibited elevated ^{15}N R_1 profiles for only for ZF3 (Figure 3.5c), while for E60Q/K79T mutant both ZF1 and ZF3 showed elevated ^{15}N R_1 values (Figure 3.5d), what strongly suggested that ZF3 domain in T23K/Q32E/E60Q/K69T, and both ZF1 and ZF3 domains in E60Q/K79T are flexible in nonspecific complex. For simplicity sake, we will refer to wild-type, T23K/Q32E, T23K/Q32E/E60Q/K69T and E60Q/K69T proteins as type-0, type-1, type-2 and type-3 proteins (Figure 3.5a). To verify that our mutations don't change protein's specificity we used gel-based competition assay. Results shown in Figure 3.5e indicated that Type-1 mutant of Egr-1 retains affinity and specificity towards target DNA. Similar experiments were done for other two mutants as well, which showed that they also retain target specificity (data not shown).

3.3.4 Significance of Domain Motions for Intersegment Transfer.

By using an NMR approach described previously,[140] we have also analyzed the kinetics of Egr-1's translocation between two nonspecific DNA duplexes. In this approach, translocation kinetics was analyzed with NMR line-shapes from the three samples: two nonspecific complexes with different 28-bp DNA duplexes and a 1:1 mixture of these complexes. All of these samples gave single sets of NMR signals,
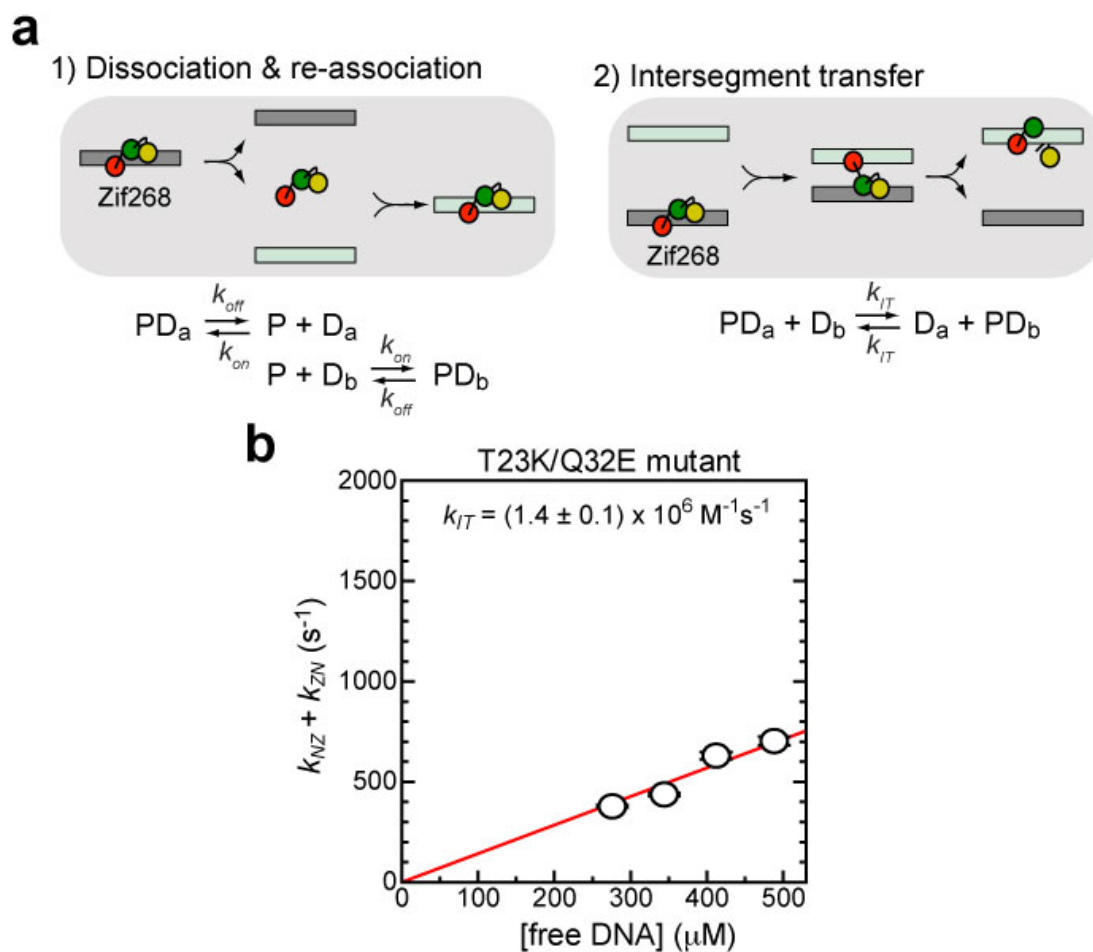
Figure 3.6 Intersegment transfer by Type-1 Egr-1. (a) Mechanisms of proteins' translocation from one nonspecific DNA to another. (b) Kinetics of intermolecular translocation of the T23K/Q32E (Type-1) mutant between nonspecific DNA duplexes.

indicating that Egr-1 is in fast exchange mode on NMR chemical shift timescale for all of those samples. Due to additional exchange contributions arising from Egr-1's translocation between the two DNA duplexes, the NMR line-shapes for the 1:1 mixture are broader when the difference between the chemical shifts of the two complexes is large. By analyzing the NMR lineshapes for the individual complexes and the mixture, we can determine the pseudo first-order rate constants for translocation between two nonspecific DNA duplexes at varying concentrations of DNA. Intermolecular translocation between two nonspecific DNA duplexes can occur via two possible mechanisms: (i) dissociation and reassociation and (ii) intersegment transfer (also known as direct transfer) (Figure 3.6a).[140,149-151] Because the rate-limiting step in the former is dissociation when [DNA] $\gg$ K$_d$ (note that this inequality leads to k_{on}[DNA] $\gg k_{off}$), the overall rate of protein translocation via this mechanism should be virtually independent of the concentration of free DNA. Intersegment transfer, however, is a second-order reaction whose rate is proportional to the concentration of free DNA. Measurements of the apparent kinetic rate constants for intermolecular translocation at different DNA concentrations permit determination of the second-order rate constant k_{IT} for intersegment transfer. By stopped-flow fluorescence experiment other members of Iwahara group have determined the rate constant k_{off} for Egr-1's dissociation from nonspecific DNA NS28 to be 0.34 ± 0.01 s^{-1}. From these kinetic data, they determined the second-order rate constant k_{IT} for intersegment transfer between nonspecific DNA molecules to be $(3.6 \pm 0.2) \times 10^6$ M$^{-1} \cdot$s^{-1}. This second order rate constant indicated that Egr-1's intersegment transfer between nonspecific DNA molecules was extremely efficient. In fact, it was $>10^6$-fold faster than intersegment transfer between target DNA sites ($k_{IT,specific} = 0.8$ M$^{-1} \cdot$s^{-1}), which was also determined in earlier study from our lab.

Because T23K/Q32E mutations cause the decrease of ZF1's domain motions, this mutant protein is suitable for examining the role of ZF1's domain motions in intersegment transfer. We confirmed that the Type-1 mutant protein retains DNA-binding

specificity and affinity to the target DNA (Figure 3.6b). Interestingly, the apparent nonspecific DNA-binding affinity of this mutant protein was found to be the same as that of the wild type protein, though one may expect a higher affinity for the mutant with more stable interactions between ZF1 and nonspecific DNA. This could be due to an entropy–enthalpy compensation arising from loss of ZF1's freedom. Using the same NMR approach as that used for the wild type protein, we measured the kinetic rate constants for intermolecular translocation of the Type-1 mutant protein between the two 28-bp nonspecific DNA duplexes at varying concentrations (Figure 3.6c). From these data, we determined the second-order rate constant k_{IT} for intersegment transfer to be (1.4 $\pm$ 0.1) $\times$ 10^6 $M^{-1} \cdot s^{-1}$. Thus, T23K/Q32E mutations that reduce ZF1's domain motions caused a 2.6-fold decrease of intersegment transfer efficiency. These results suggest that the domain dynamics is important for intersegment transfer of Egr-1. In addition, another member of Iwahara group, A. Esadze analyzed target search kinetics of all four Egr-1 proteins by using stopped-flow kinetic experiments (data not shown). He demonstrated that Type-1 mutant of Egr-1 showed the slowest target search kinetics, while Type-3 mutant was the fastest in locating the target DNA site. Type-0 and Type-2 proteins, which both have only one flexible domain had comparable intermediate speeds compared to two other constructs. These data also supported my observation and indicated that Egr-1's domain motions facilitate its translocation from one nonspecific DNA site to another.

3.3.5 Domain Motions could Enhance ZFNs' Activity.

To examine whether domain motions can also facilitate ZFNs' rapid translocation and enhance their activity, we created ZFNs by combining all four Egr-1 proteins with FokI ND (Figure 3.7b). FokI ND cannot bind to or cleave DNA in nonnumeric form and is known not to take part in translocation process, therefore only zinc-fingers should be
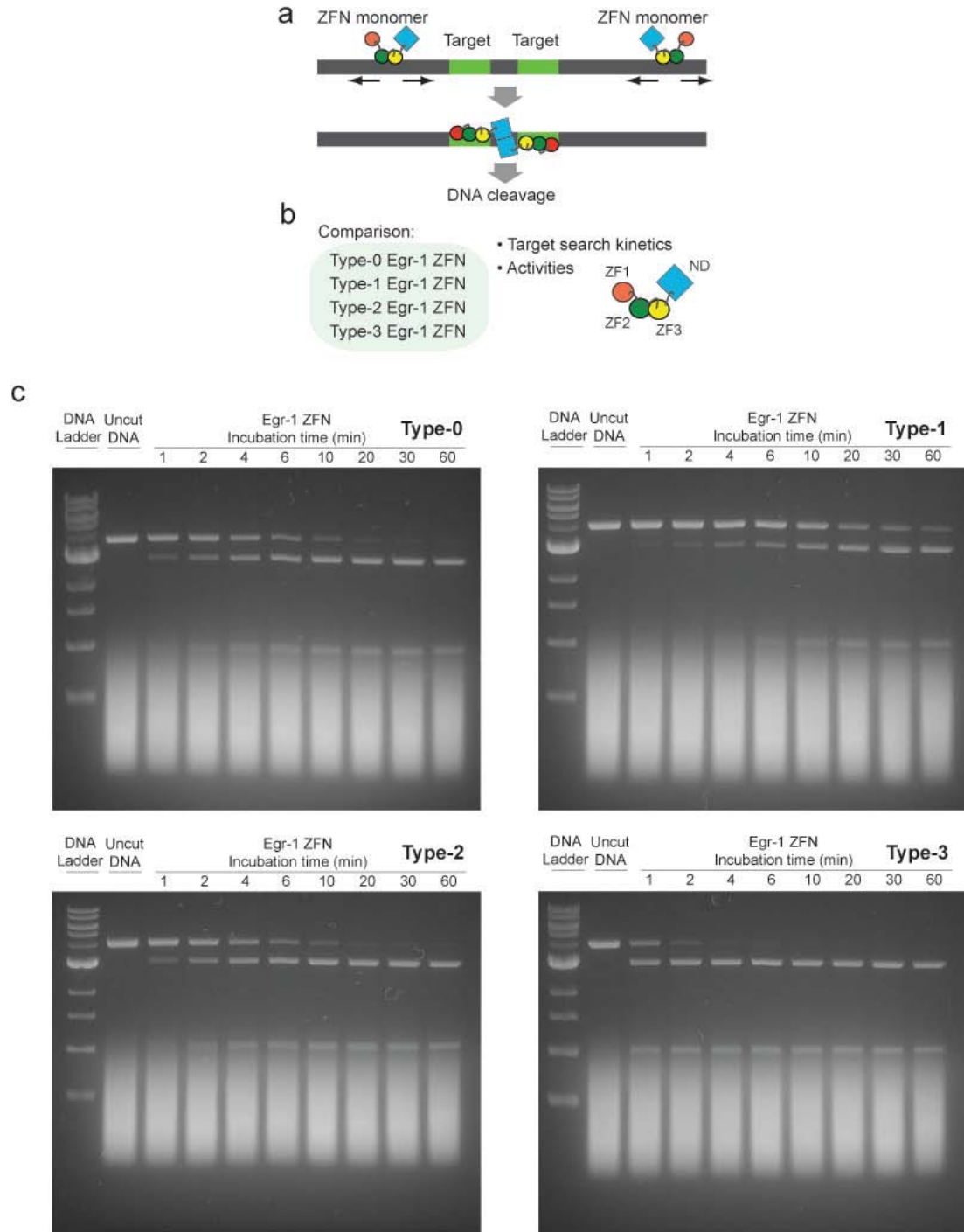
Figure 3.7   DNA digestion by four ZFN constructs. (a) Schematic presentation of DNA cleavage by ZFN, (b) Proof-of-concept study for improvement of ZFN engineering, (c) Time course of DNA digestion by type-0, -1, -2 and -3 Egr-1 ZFNs. Gels clearly show the difference in ZFN activity between these proteins. Exactly same conditions were used for all experiments.
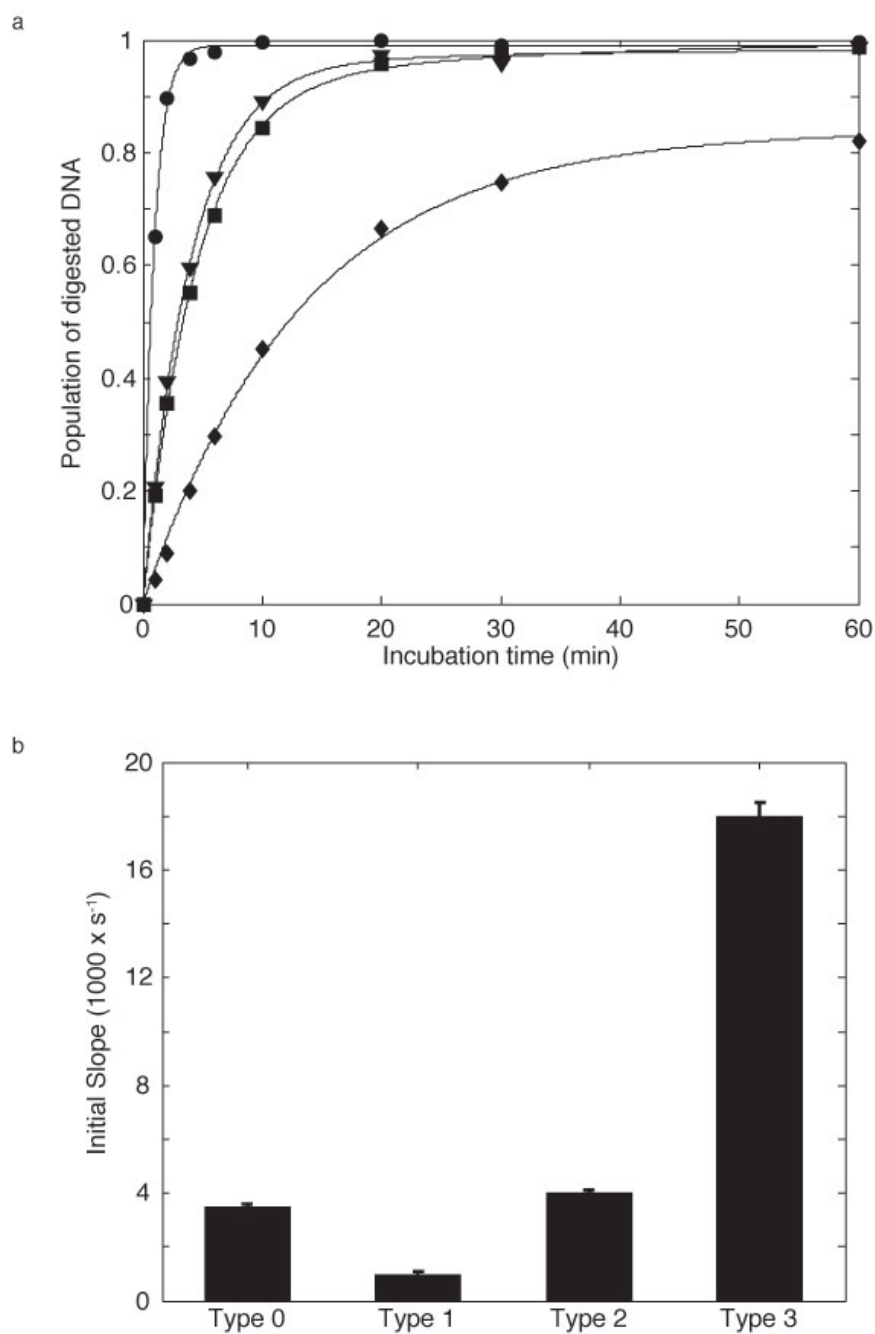
Figure 3.8   Kinetics of DNA cleavage by ZFN proteins. (a) Changes in populations of cleaved DNA for (■) Type-0, (♦) Type-1, (▼) Type-2 and (●) Type-3 Egr-1 ZFNs over one hour of incubation at room temperature. Non-linear least-square fittings were done using double-exponential model. (b) Initial slopes from fittings of changes in population of cleaved DNA.

responsible for ZFNs' translocation.[129,130,152] To check DNA cleavage by these constructs, we designed ~4kb DNA plasmid with two Egr-1 target sequences with opposite orientations, separated with 6bp-s (Figure 3.7a). Next, target DNA plasmid was linearised by BglII digestion. We examined digestion of linear target DNA by ZFNs in presence of Calf Thymus DNA as a competitor. Calf Thymus DNA is a good analogues to nuclear environment, as it contains different length DNA molecules, with an average length of ~500 bp. To evaluate ZFNs efficiency, target DNA was digested by ZFNs over one hour and population of digested vs. full-length species at different time points was determined using agarose gel based electrophoresis. Our results showed that Type-1 Egr-1 ZFN with no dynamic domains was the least efficient in specific DNA cleavage compared to other proteins, while Type-3 ZFN construct with two dynamic domains was the most efficient in DNA digestion (Figure 3.7c). Due to absence of a good numerical model, we performed double exponential fitting to population of digested species and used the initial slope to compare relative activities of Egr-1/FokI constructs (Figure 3.8a). As we determined from experimental results, the ZFN construct without any dynamic domains was four times slowed than any of proteins with a single dynamic domain, and 16 times slower than the protein with two dynamic domains (Figure 3.8b). This data was in good agreement with studies of Egr-1 and its mutants, suggesting that domain motions in nonspecific complex can facilitate ZFNs' translocation between nonspecific sites, thus enhance ZFNs' overall activity.

3.4 DISCUSSION

3.4.1 Target search by Egr-1

Mechanisms that allow transcription factors to find their target sites rapidly in a sea of nonspecific DNA have been the subject of considerable interest in molecular biophysics.[103,110,111] Despite the long history of the field and a wealth of theoretical

studies, very little has been experimentally known about structural details of target DNA search by transcription factors.[51,140,153] Our present work together with some other experimental data from other lab members demonstrates the asymmetrical roles of Egr-1's three zinc fingers in the dynamic DNA-scanning process and provides important insights into how this mammalian transcription factor can achieve high specificity in binding and rapidityin DNA-scanning.

An important finding is that Egr-1's binding modes for specific and nonspecific DNA duplexes are substantially different in structural and dynamic terms. Based on similar experimental observation for Escherichia coli lac repressor in a pioneering work by Kalodimos *et al.*,[51] some groups theoretically pursued the conformational switch model as a mechanism that resolves the speed-stability paradox.[102,103,154-156] Our finding for Egr-1 is remarkably consistent with this model, although the structures are completely different for lac repressor and Egr-1. In the conformational switch model, proteins on DNA undergo rapid transitions between two modes: the search mode and recognition mode. The search mode is suitable for proteins' rapid translocation but unsuitable for high stability and specificity; whereas, the recognition mode is unsuitable for rapid translocation but suitable for high stability and specificity.[102,156] The state observed in the crystal structure of the specific complex with all three zinc-finger domains bound to DNA corresponds to the recognition mode. The dynamic state of Egr-1 on nonspecific DNA with only ZF2 and ZF3 being bound corresponds to the search mode (Figure 3.9a). Domain motions seem to permit the transitions between the search and recognition modes of Egr-1. Because of ZF1's domain motions on a nanosecond timescale, transitions between the two modes can occur more rapidly than translocation. This rapid conformational switch would allow Egr-1 to efficiently scan DNA and specifically locate its target sites. The domain motions of locally dissociated ZF1 in Egr-1 bound to nonspecific DNA can also promote intersegment transfer via a transient bridging of two
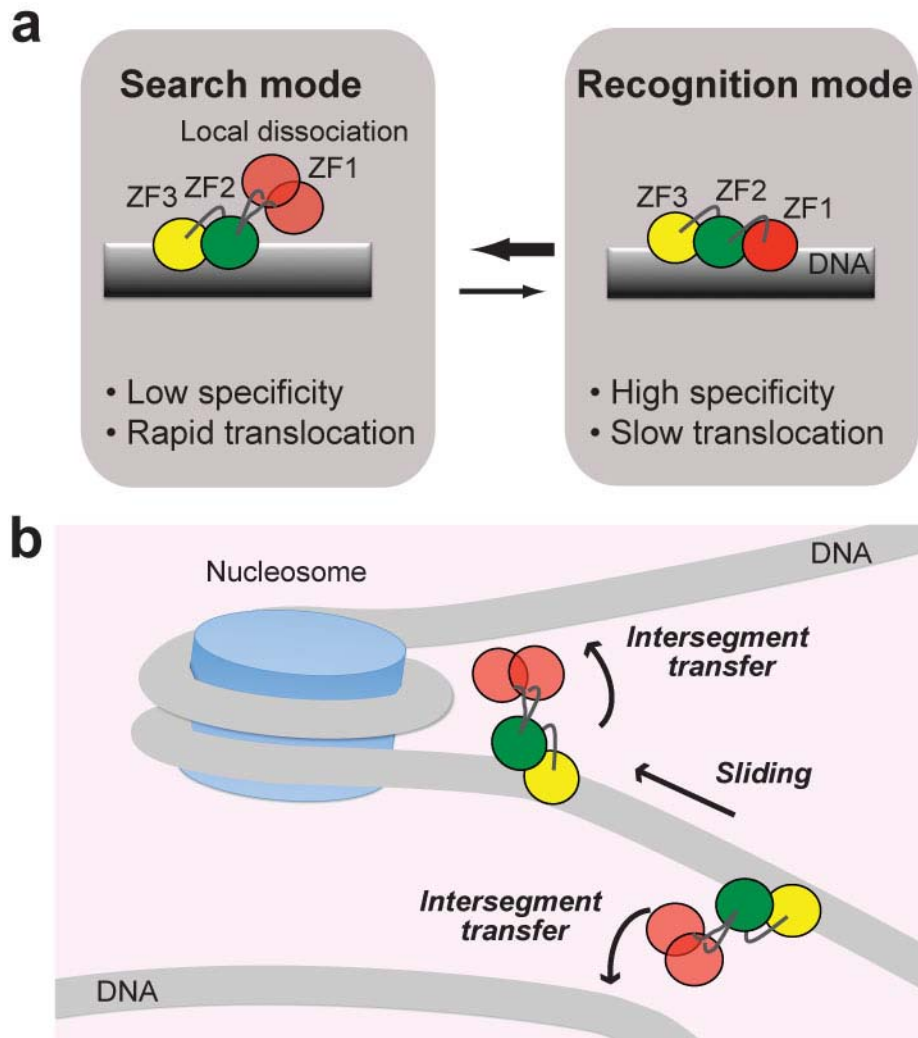
Figure 3.9 Dynamic mechanism for rapid target search by Egr-1. (a) Dynamic transitions between search and recognition modes via domain dynamics that can allow Egr-1 to resolve the speed-stability paradox. Local dissociation of ZF1 in the search mode can accelerate translocation of Egr-1 on DNA by decreasing energy barriers for sliding and enhancing intersegment transfer. (b) Bypassing a nucleosome via intersegment transfer. Because the two DNA ends of a nucleosome are separated by only ∼60 Å, Egr-1 may bypass nucleosomes via intersegment transfer and carry out continuous scanning.

DNA molecules in close proximity as observed in the CGMD simulations. While ZF2 and ZF3 are bound to the original DNA site and act as anchors, ZF1 acts as an explorer that transiently searches for potential DNA binding sites. In fact, our experimental data indicate highly efficient intersegment transfer between nonspecific DNA duplexes. Due to the extremely high DNA density in the nucleus ($\sim$100 mg/mL),[157] intersegment transfer can be kinetically predominant in translocation between distant DNA sites. Even if 90% of the DNA is wound up by nucleosomes, the concentration of naked 30-bp segments in the nucleus is as high as $\sim$0.5 mM. At this concentration, a pseudo first-order rate constant k_{IT}[DNA] for intersegment transfer is calculated to be $\sim$1,800 s^{-1}. Thus, Egr-1's intersegment transfer between nonspecific DNA at physiological concentrations is far faster than translocation via the "dissociation and re-association" mechanism for which the rate-limiting step is dissociation with a rate constant of 0.3 s^{-1}. Furthermore, studies of Type-2 and Type-3 mutants of Egr-1 demonstrated that by additional mutations it is possible to modulate Egr-1's translocation speed through regulation of its domain dynamics. To better understand how domain dynamics facilitates protein's translocation, our collaborators (Dr. Y. Levy's group) performed coarse-grained molecular dynamics (CGMD) simulations. Their results on Egr-1's dynamics also suggested that ZF1 exhibits higher flexibility in the nonspecific complex compared to ZF2 and ZF3. The most interesting in their simulations was that wild-type Egr-1 was able to form a bridging structure between two parallel DNA molecules using flexible ZF1 domain, what facilitated intersegment transfer. And such bridging structure was formed mostly through ZF1 domain. Their data was also in perfect agreement with my results, showing that forming bridging structure was harder for Type-1 mutant, what also decreased k_{IT} of this mutant in CGMD simulations. Our experimental observation, together with CGMD simulations strongly suggested that flexibility of Egr-1's domains facilitates its intersegment transfer via forming a transient bridging structure, though additional studies

69

are necessary to fully support this hypothesis. In other words, this work strongly suggests relationship between domain dynamics of Egr-1 and its target search kinetics.

3.4.2 Modulation of ZFN activity through zinc finger domain dynamics

After getting deeper insight on how natural zinc-finger proteins are able to efficiently scan DNA, we tested whether the same approach can be used to solve problems associated with artificial zinc-finger nucleases and improve their activities. For this purpose, we designed four ZFNs: each of these ZFNs was a fusion of one of four Egr-1 proteins used in this project and of FokI ND. We expected that due to absence of domain flexibility, Type-1 ZFN would be least efficient in locating target DNA sequence, while constructs with flexible domains would be faster to bind to target DNA sites. Earlier studies on FokI have shown that k_{cat} for this enzyme is on the order of 0.5 s^{-1}.[158] At our experimental conditions the rate-limiting step in our ZFNs' activity should be target search, rather than k_{cat}, as it should take at least several, or even dozens of minutes to locate the target site. Therefore, proteins exhibiting fast DNA translocation should also be more efficient in DNA cleavage. Indeed, our experiments clearly demonstrated that Type-1 Egr-1, which demonstrated the slowest translocation on DNA, also was the least efficient to introduce double strand breaks in target DNA. On the other hand, Type-3 ZFN with two flexible domains was the most efficient to break target DNA molecule. As FokI domain doesn't take part in DNA search process, and k_{cat} for all constructs should be independent from mutations in zinc-fingers, it is safe to conclude that the difference in the activities of four ZFN proteins is due to the difference in their target search kinetics. These experimental results underline the flaw in designing the ZFNs for use in medical applications, as zinc-fingers for these proteins are selected exclusively based on their DNA binding affinity, completely ignoring their target search kinetics. This kind of selection results in proteins which are stuck in recognition mode, what is great for DNA

specificity, but due low population of search mode is disadvantageous for target search kinetics.[103] As a conclusion, my data underlines a significant flaw which exists in current protocols for designing ZFNs, and suggest a strategy for improving ZFNs activities via enhancing their target search kinetics by modulating their zinc-finger dynamics in nonspecific complex.

CHAPTER 4:

Exploring NMR studies of protein dynamics and kinetics in situ.

4.1 INTRODUCTION

NMR is a major tool for studying protein dynamics and its relationship with biological function. Though, using NMR it is possible to acquire information on molecular behavior on a very wide range of timescales at atomic level resolution, it has its own limitations.[9,11,14,19,20] One of the main problems is that the vast majority of NMR experiments are done *in vitro*, under non-physiological conditions, where proteins are "shielded" from nonspecific interactions. While such conditions are making it easier to analyze proteins' properties, they leave many open questions and leave us wondering how interaction with sea of other macromolecules in cells or extracellular fluids affect behavior of these proteins. Over decades different researchers have suggested or demonstrated the effect of molecular crowding and nonspecific interactions on proteins' dynamic behavior and its functional characteristics.[44-46] To date, there are only few experimental methods, which can provide pseudo-atomic or atomic resolution data on proteins in their physiological environment, with most of them being either underdeveloped or having some serious limitations.[48] Such lack of necessary methodology creates the huge knowledge gap, which often results in unexplained discrepancy between *in vitro* vs. *in vivo* or *in situ* studies. Therefore, there is an urgent need to develop new approaches, which will allow elucidation of protein characteristics
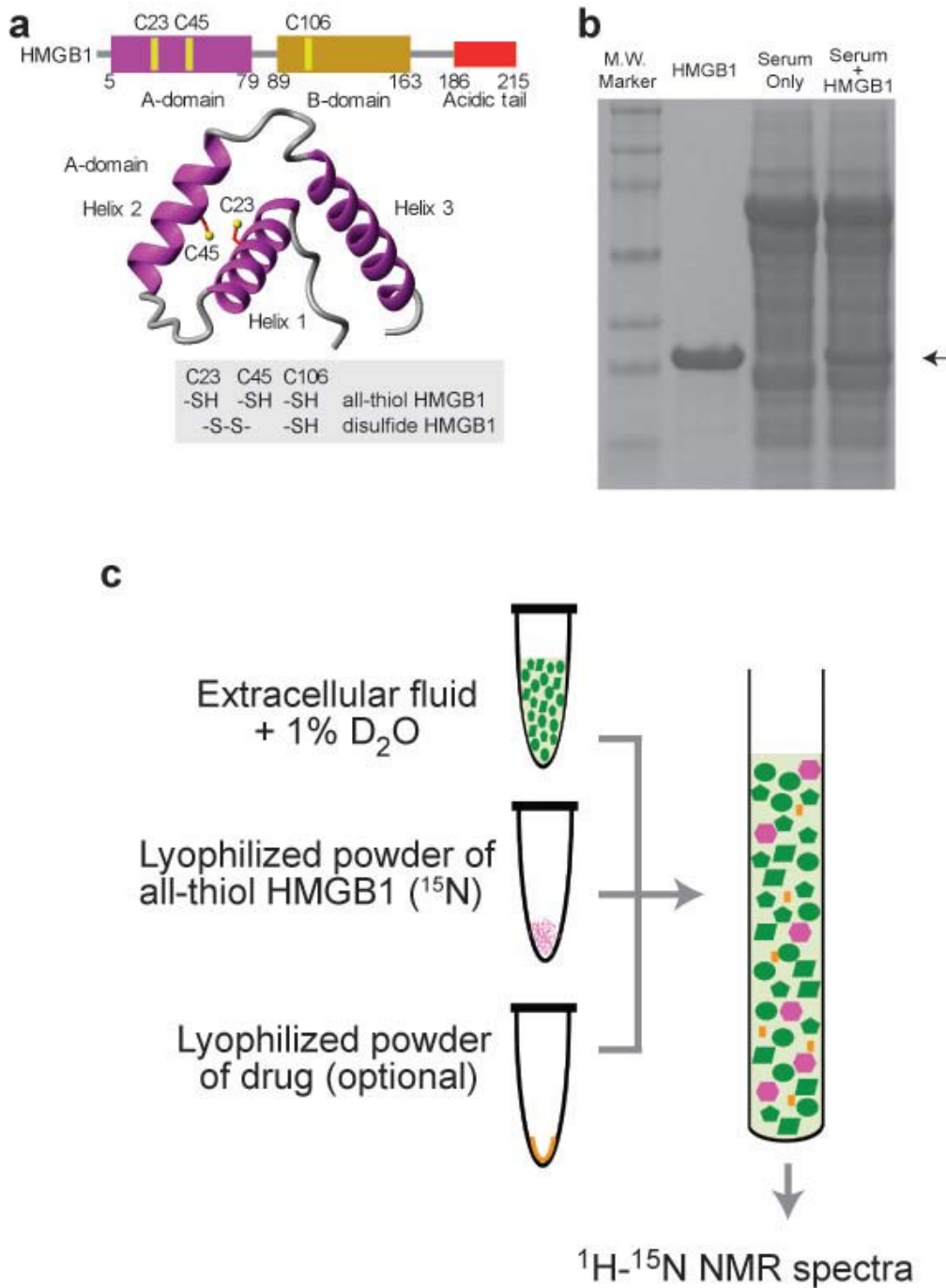
Figure 4.1. *In situ* NMR experimental design (a) HMGB1 and its three cysteine residues. Cys-23 and Cys-45 form a disulfide bond under oxidative conditions. (b) SDS-PAGE of the serum samples used for the NMR experiments. 3 µl of serum before and after dissolving 135 µM ^{15}N-labeled HMGB1 was loaded. The arrow indicates the position of the HMGB1 band. The strongest bands are albumin. (c) experimental scheme for the *in situ* protein NMR approach to investigate the behavior of HMGB1 in actual extracellular fluids.

under their natural habitats. In this study, we used HMGB1 (high-mobility group box 1) protein as a model system. This 25-kDa protein, which is normally located in nuclei, is passively released during cell injury and necrosis and actively secreted by some cell types via a non-canonical pathway that bypasses the endoplasmic reticulum.[159,160] Extracellular HMGB1 acts as a danger signal and inflammatory mediator via binding to cell surface receptors such as RAGE, TLR4 (Toll-like receptor 4), and CXCR4.[159-162] Extracellular HMGB1/receptor interactions promote inflammation and angiogenesis via activation of NF-B[163,164] and tumor proliferation via activation of p44/p42, p38, and SAPK/JNK MAPKs.[165,166] When released into the extracellular space, HMGB1 is initially in the reduced state ("all-thiol HMGB1") (see Figure. 4.1a) but becomes oxidized due to the oxidative environment.[161,167,168] HMGB1 has three cysteine residues. Cys-23 and Cys-45 in the A-domain are located in close proximity (see Figure 4.1a). With a standard redox potential of -237 mV,[168] these two cysteine residues can easily form a disulfide bond under relatively mild oxidative conditions ("disulfide HMGB1"). Cys-106 in the B-domain remains in the reduced state but can be sulfonated in the extracellular space when exposed to a large amount of reactive oxygen species from activated leukocytes.[169] The three different redox states (i.e. all-thiol, disulfide, and sulfonated) of extracellular HMGB1 play distinct roles in inflammation.[162,170] All-thiol HMGB1,but not disulfide HMGB1,is able to form a complex with CXCL12 for signaling via the CXCR4 receptor and exhibits chemoattractant activity.[171] Only disulfide HMGB1 can interact with TLR4 and exhibits cytokine-inducing activity.[171-173] Sulfonated HMGB1 is involved in the resolution of inflammation.[169] Therefore, knowledge of the lifetimes of the different species is essential to better understand the relative roles of these different redox species.

In order to fill up the gap, which exists between *in vitro* and *in situ* studies, we developed a novel NMR-based approach, allowing investigation of proteins' functionally important properties in various biological fluids. We used this approach to investigate the kinetics of HMGB1 oxidation and the half-lives of all-thiol and disulfide HMGB1 species

in serum, saliva, and cell culture medium. We also studied how oxidation kinetics is affected by exogenous ligands such as heparin, which mimics cell surface heparan sulfate. By these studies we demonstrate that our approach can be used elucidate proteins' properties in different biological fluids, and has a potential to provide information on proteins' behavior under such conditions on the pseudo-atomic and atomic resolutions. This will bring us one step closer to understanding the effect of molecular crowding and nonspecific interactions on macromolecules' dynamics and function.

4.2 MATERIALS AND METHODS

4.2.1 Human Serum

The serum sample for this study was prepared from a blood sample obtained from a consenting healthy male (71 years old) by our approved Institutional Review Board protocol. The blood was centrifuged at $1000 \times g$ for 5 min at 20 °C. The supernatant was transferred into BD Vacutainer red-top tubes (non-heparinized) and allowed to clot by leaving it undisturbed for 30 min at room temperature. The clot was removed by centrifugation at $200 \times g$ for 5 min at room temperature. Without disturbing the coagulated pellet, the resultant supernatant (i.e. serum) was carefully transferred into 2-ml polypropylene microtubes and stored at -80 °C until used.

4.2.2 Human Saliva

Human saliva was sampled from a healthy male (42 years old). Before sampling, the mouth was washed with 18-megohm water, and the saliva was collected 10 min later. The saliva was centrifuged at $2000 \times g$ for 10 min at room temperature, and the supernatant was immediately used for the NMR experiment.

4.2.3 Extracellular Fluids of PC-3M Cell Culture

Human prostate cancer PC-3M cells[174] were cultured at 37 °C in RPMI 1640 medium (Invitrogen) with 10% FBS (Atlanta Biologicals). Cells were trypsinized for seeding. To remove trypsin, the cells were centrifuged at 1000 rpm for 5 min at room temperature, and the cell pellet was washed once with phosphate-buffered saline (Invitrogen) and then suspended with RPMI 1640 medium plus 10% FBS. Three different titrating amounts of the cell suspension were used for seeding to obtain cultures with three different cell confluences at the same time. Culture supernatants were collected after the cells were cultured for 48 h at 37 °C under 5% CO_2. As a control, a blank (without cells) RPMI 1640 medium plus 10% FBS was also incubated and collected at the same time. Each fluid was filtered with a 0.22-µm filter and kept at -80 °C until used.

4.2.4 HMGB1 Ligands

Glycyrrhizin was purchased from Nacalai USA, Inc. Heparin octasaccharide was purchased from Iduron. These materials were dissolved in 18-megohm water and lyophilized before use for in situ protein NMR experiments (see Figure 4.1b).

4.2.5 Preparation of HMGB1

Full-length HMGB1 protein was expressed in *Escherichia coli* strain Rosetta 2(DE3) cells harboring a pET-11d-derived plasmid with the human HMGB1 gene inserted between the NcoI and BamHI sites. For preparation of ^{15}N- or ^{13}C/^{15}N-labeled proteins, ^{15}N ammonium chloride and ^{13}C glucose were used as the sole sources of nitrogen and carbon, respectively, in the *E. coli* culture medium. Protein expression was induced with 0.4 mM isopropyl β-D-thiogalactopyranoside, and the *E. coli* culture was continued at 18 °C for 16 h. Harvested cells were suspended in buffer containing 20 mM

sodium phosphate (pH 6.0), 1 mM EDTA, 100 mM NaCl, 2 mM DTT, and 5% glycerol and disrupted at 4 °C by sonication. The supernatant of the cell lysate was loaded onto an SP cation exchange column equilibrated with 20 mM sodium phosphate (pH 6.0) and 100 mM NaCl and eluted with a gradient of 100–2000 mM NaCl. Fractions containing HMGB1 were concentrated and passed through an S-100 column equilibrated with buffer containing 20 mM Tris-HCl (pH 8.0), 1 mM EDTA, and 200 mM NaCl. After size exclusion chromatography, HMGB1 was further purified by RESOURCE Q anion exchange chromatography with a gradient of 0–1500 mM NaCl in 50 mM Tris-HCl (pH 7.5). No reducing regent was used in the column chromatography procedures except in the sonication buffer. Complete disulfide bond formation in HMGB1 was confirmed by NMR.[168]

4.2.6 Preparation of Salt-free Powders of ^{15}N-Labeled All-thiol HMGB1

To break the disulfide bond, 5 mM DTT was added to a solution of ^{15}N-labeled disulfide HMGB1, and the mixture was kept overnight at 4 °C. Using an Amicon Ultra-15 unit, the solvent was exchanged with water containing 1 mM β-mercaptoethanol. Because of the lower solubility of HMGB1 in the absence of salt, the process of the solvent exchange was carried out at protein concentrations lower than 200 µM. The final solution was divided into aliquots (each containing 68 nmol of HMGB1), frozen at -80 °C, and lyophilized. β-Mercaptoethanol was removed during the lyophilization process, and all-thiol HMGB1 remained as a powder. The obtained powder was stored at -20 °C until used.

4.2.7 Preparation of Salt-free Powders of ^{15}N-Labeled Disulfide HMGB1

Salt-free powders of ^{15}N-labeled disulfide HMGB1 were prepared in the same way as for all-thiol HMGB1 except that the samples were not treated with either DTT or β-mercaptoethanol. Disulfide HMGB1 in water was lyophilized as described above.

4.2.8 In Situ NMR Experiments to Measure the Half-lives of All-thiol and Disulfide HMGB1 Species in Extracellular Fluids

For the kinetic measurements, the salt-free powder of ^{15}N-labeled all-thiol HMGB1 was dissolved in 500 µl of extracellular fluid (i.e. serum, saliva, or cell culture medium) plus 1% D2O (for the NMR lock). The solution was immediately sealed in a 5-mm NMR tube with ambient oxygen pressure, and a series of [1]H-[15]Ncorrelation spectra were recorded at 37 °C using Bruker AVANCE III spectrometers equipped with cryogenic probes (1H frequencies, 800 and 600 MHz). In all experiments except those involving the cell culture medium, band-selective optimized flip-angle short-transient heteronuclear multiple quantum coherence (SOFAST-HMQC) spectra[175] were recorded with a recycle delay of 0.1 s. For the cell culture medium, in which HMGB1 oxidation is relatively slow, the kinetics were measured with a series of sensitivity-enhanced [1]H-^{15}N transverse relaxation optimized spectroscopy spectra[142,176] recorded with a recycle delay of 0.8 or 1.0 s. Values of pH for the NMR samples were confirmed to be ~7.4. Kinetic rate constants were determined via nonlinear least-squares fitting with MATLAB software (MathWorks). Other details of the calculations are given in the figure legends.

4.2.9 NMR Resonance Assignment

Because the spectra of HMGB1 samples dissolved in extracellular fluids and in buffers are very similar, resonances of HMGB1 in extracellular fluids were assigned

using the NMR spectra recorded for HMGB1 in a buffer. ^{1}H/^{13}C/^{15}N resonances for oxidized and reduced HMGB1 proteins were assigned with three-dimensional HNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCACB, CBCA(CO)NH, C(CO)NH, ^{15}N-edited NOESY-heteronuclear single quantum coherence; and F1-^{15}N/F2-^{15}N-edited HMQC-NOESY-heteronuclear single quantum coherence spectra[177] recorded at 25 °C on 0.7mM ^{13}C/^{15}N-labeled proteins in buffer containing 20 mM Tris-HCl (pH 7.5), 120 mM NaCl, and 5% D_2O. For all-thiol HMGB1 in the buffer, 5 mM DTT was also added, and the NMR tube was sealed in the presence of argon gas. Resonance assignment was also aided by previous assignment data for HMGB1 under different conditions.[168,178-180] NMR data were processed and analyzed using the NMRPipe[56] and NMRView[57] programs.

In this study, I use the residue numbering scheme from the initial methionine in the gene (i.e. Met-1–Gly-2–Lys-3 . . .) because this has become the accepted norm, although the numbering scheme from the actual N-terminal glycine (i.e. Gly-l–Lys-2 . . .) of HMGB1 has been used in many previous articles (including ours).

4.3 RESULTS

4.3.1 NMR of All-thiol and Disulfide HMGB1 Proteins Dissolved in Serum

We prepared salt-free lyophilized powders of ^{15}N-labeled all-thiol HMGB1 and HMGB1 proteins and dissolved them in human serum to characterize the actual redox kinetics of extracellular HMGB1 by NMR. Figure 4.2a shows NMR spectra recorded at 37 °C for all-thiol HMGB1 and disulfide HMGB1 in human serum. The spectra were similar to spectra recorded for all-thiol HMGB1 and disulfide HMGB1 in buffer containing 20 mM HEPES-NaOH (pH 7.4), 120 mM NaCl, and 5% D2O (Figure 4.2b). This suggests that the HMGB1 structures in serum and in buffer are identical, although serum is a molecular crowding environment with numerous proteins as shown by the SDSPAGE data (Figure 4.1b). Our NMR data also indicate that the proteins from the salt-
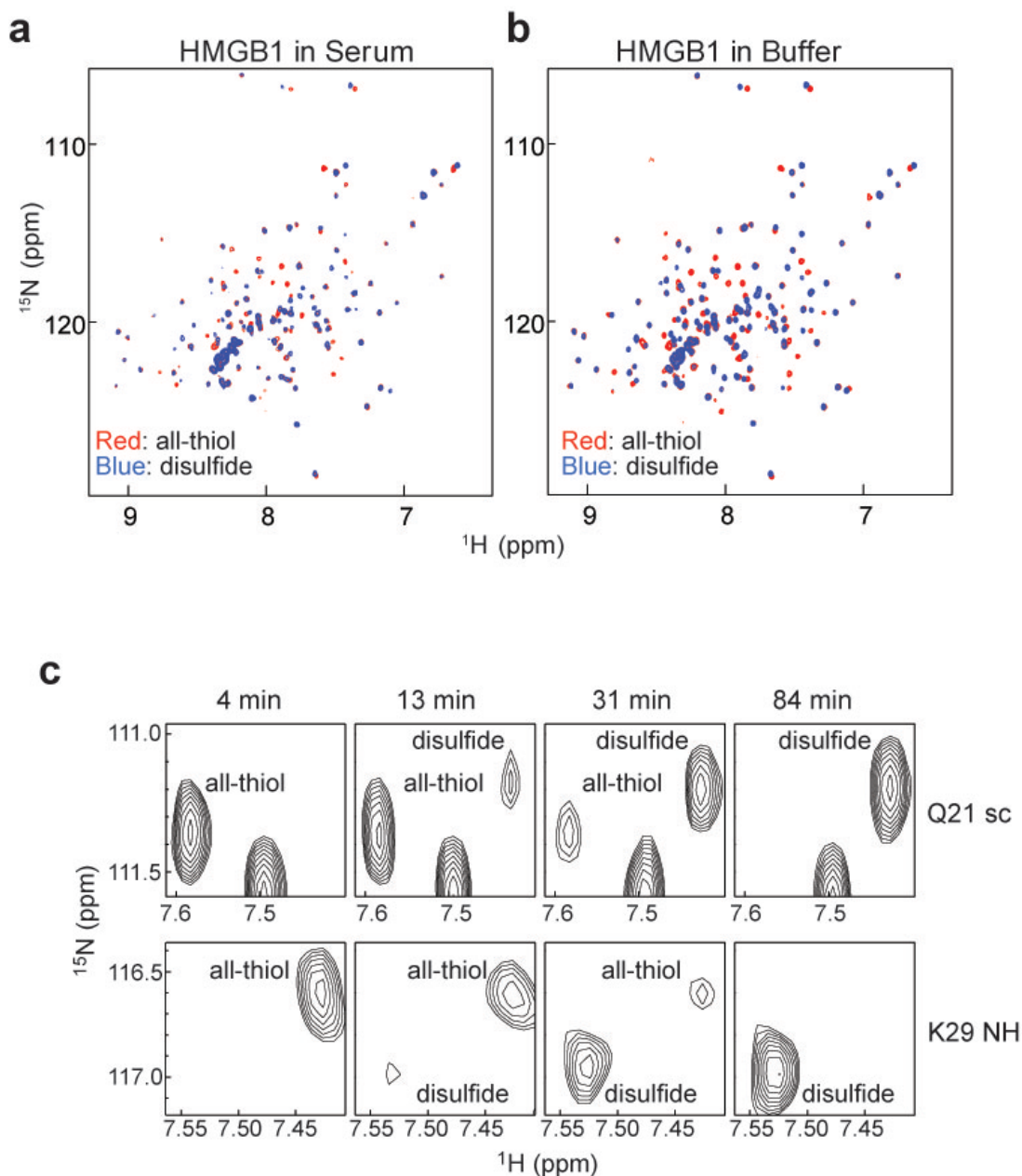
Figure 4.2. HMGB1's NMR spectra in different fluids. (a) ^{1}H-^{15}N SOFAST-HMQC spectra recorded with eight scans at 37 °C for all-thiol HMGB1 (red) and disulfide HMGB1 (blue) immediately after dissolving in human serum. (b) ^{1}H-^{15}N SOFAST-HMQC spectra of all-thiol HMGB1 (red) and disulfide HMGB1 (blue) immediately after dissolving in 20 mM HEPES-NaOH (pH 7.4), 120 mM NaCl, and 5% D_2O. Line shapes of NMR signals from HMGB1 in serum were substantially broader than those in buffer presumably due to the molecular crowding environment of the serum. (c) changes in the ^{1}H-^{15}N SOFAST-HMQC spectra recorded at 37 °C for all-thiol HMGB1 in human serum. Due to oxidation, signals from all-thiol HMGB1 became weaker, whereas those from disulfide HMGB1 became stronger.

free lyophilized powders were properly folded when dissolved in extracellular fluids. As demonstrated in previous study from Iwahara lab on the A-domain of HMGB1,[168] residues in close proximity to Cys-23 or Cys-45 exhibited large ^{1}H/^{15}N chemical shift differences between all-thiol HMGB1 and disulfide HMGB1. Due to rapid oxidation, however, the spectra recorded for the sample of all-thiol HMGB1 in serum gradually changed and finally became identical to the spectra of disulfide HMGB1 (Figure 4.2c). We used this change in NMR spectra to investigate the oxidation kinetics of HMGB1 in serum.

4.3.2 Kinetics of HMGB1 Oxidation in Serum

We recorded a series of ^{1}H-^{15}N SOFAST-HMQC spectra immediately after dissolving ^{15}N-labeled all-thiol HMGB1 in serum. SOFAST-HMQC allows acquisition of high-resolution ^{1}H-^{15}N correlation spectra in a few minutes and is therefore ideally suited for measuring real-time kinetics.[175] NMR signals from disulfide HMGB1 gradually became stronger and predominant, whereas signals from all-thiol HMGB1 became weaker and eventually disappeared in 50 min (Figure 4.2c). This clearly reflects the oxidation process of all-thiol HMGB1 in serum. Using the SOFAST-HMQC spectra, we obtained the time course data of the fraction of oxidation (Figure 4.3a). By nonlinear least-squares fitting with the experimental data, we determined the apparent pseudo-first-order rate constant (k_{ox}) for HMGB1 oxidation. The half-life ($t_{1/2}$) of all-thiol HMGB1, which was calculated as $k_{ox}^{-1} \, ln2$, was determined to be 17 ± 1 min (Table 4.1). We also measured k_{ox} at different concentrations of HMGB1 (Figure 4.3b). k_{ox} was slightly smaller at concentrations higher than 200 µM. This is probably because the overall concentration of the oxidants that oxidized HMGB1 was not high enough to make the process completely pseudo-first-order. Because the in vivo concentration of extracellular
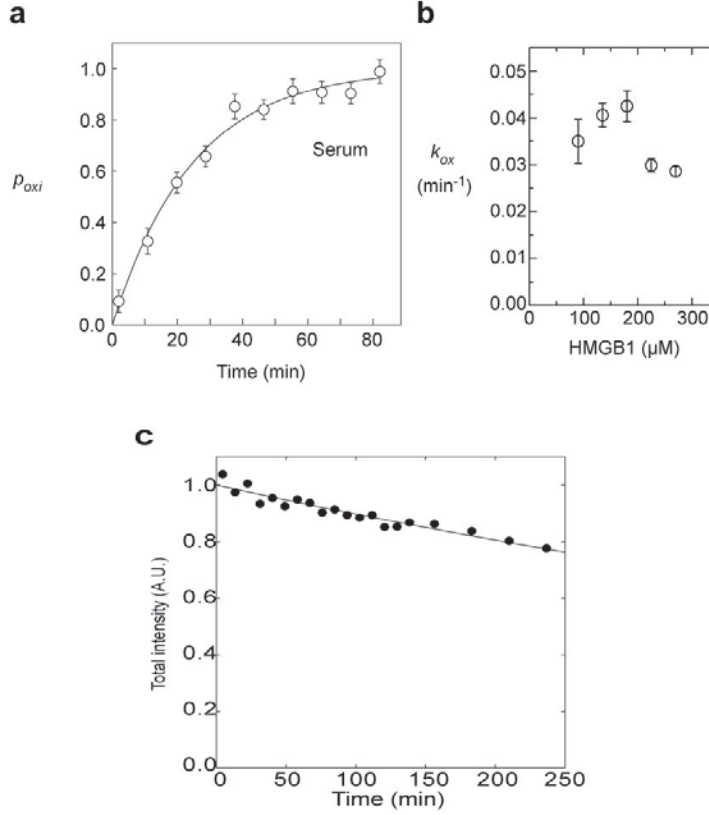
Figure 4.3.   Kinetics of oxidation and clearance of HMGB1 in human serum. (a) oxidation kinetics of HMGB1 in human serum measured by in situ protein NMR. The vertical axis represents the fraction of oxidation (p_{oxi}) measured as $(1/N)\sum_{i}^{N} I_{D,i}/(I_{D,i} + I_{A,i})$, where I_A and I_D are signals intensities for all-thiol and disulfide HMGB1 proteins, respectively, and i is indices for analyzed residues. Intensities were measured for NMR signals from residues that exhibit relatively strong, well isolated signals for both all-thiol and disulfide states (i.e. those from Thr-22, Lys-29, His-31, Thr-51, and Gly-58 backbone amide groups and the Gln-21 side chain NH_2 group). Experimental p_{ox} data were fitted to $1 - \exp(-k_{ox}(t - t_0))$, where k_{ox} represents a pseudo-first-order rate constant for the oxidation of all-thiol HMGB1, and t_0 represents a time shift due to the difference between the effective times and the mid-times for recording NMR spectra. This shift was applied to the horizontal axis of the time course data shown.(b) kinetic rate constants (k_{ox}) for HMGB1 at different concentrations (90, 135, 180, 225, and 270 µM) in human serum. (c) decrease in the total intensity of NMR signals from amide groups whose $^1H/^{15}N$ chemical shifts were unaffected by oxidation. SOFAST-HMQC spectra recorded for 135 µM ^{15}N-labeled HMGB1 in human serum were analyzed. The curve represents the best fit to $a \exp(-k_{cl}t)$, where k_{cl} is a pseudo-first-order rate constant for HMGB1 clearance. The vertical axis represents the total intensity divided by a. *A.U.*, arbitrary unit.

HMGB1 is lower than micromolar, HMGB1 oxidation in vivo should occur in a completely pseudo-first-order manner, and the half-life *in vivo* should be only slightly shorter than what we observed *in situ*.

4.3.3 Clearance of HMGB1 in Serum

Our NMR data also provide kinetic information on the clearance of HMGB1 due to protease activities in extracellular fluids. Figure 4.3c shows the time course of the loss in total intensity of NMR signals whose positions (i.e. ^{1}H/^{15}N chemical shifts) were unaffected by the oxidation. The HMGB1 band in SDS-PAGE became weaker over hours as well, which indicates that the decrease in NMR signal intensities is caused by a decrease in the total amount of HMGB1. Apparent kinetic rate constants for HMGB1 clearance (k_{cl}) were determined via monoexponential fitting with the time course data shown in Figure 4.3c. The half-life of disulfide HMGB1 in serum was calculated as $k_{ox}^{-1} \, ln2$ to be 642 ± 49 min (Table 4.1). Interestingly, neither ^{1}H-^{15}N spectra nor SDS-PAGE showed degradation products of HMGB1. These results imply that, once cleavage occurs, HMGB1 molecules become highly susceptible to complete digestion by extracellular proteases to an amino acid level.

4.3.4 Kinetics of HMGB1 Oxidation and Clearance in Saliva

Salivary gland cells are known to produce extracellular HMGB1.[181] To examine whether or not the molecular behavior of HMGB1 in saliva is similar to that in serum, I also investigated oxidation and clearance of HMGB1 in saliva using the *in situ* protein NMR approach. Interestingly, HMGB1 clearance in saliva was found to be 10-fold faster
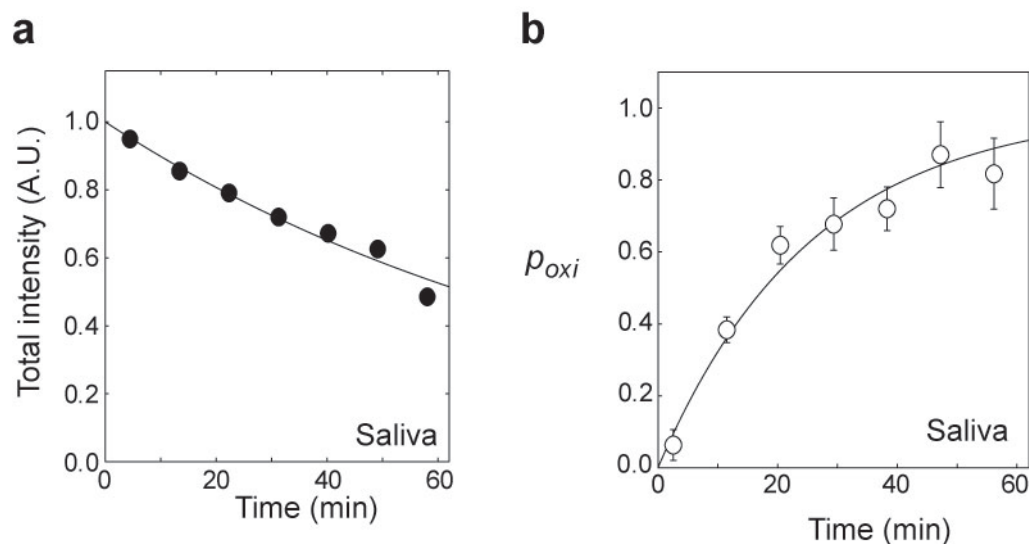
Figure 4.4 Kinetics of HMGB1 clearance (a) and oxidation (b) in human saliva measured by in situ protein NMR. The clearance of HMGB1 in saliva was 10-fold faster than that in serum (see Table 4.1). A.U., arbitrary unit.

Table 4.1 Half-lives of all-thiol HMGB1 and disulfide HMGB1 in extracellular fluids measured by in situ NMR

Extracellular fluids	Oxidation $t_{1/2}$[a]	Clearance $t_{1/2}$[b]
	min	*min*
Human serum	17 ± 1	642 ± 49
Human saliva	18 ± 2	65 ± 6
RPMI 1640 medium _ 10% FBS[c]	25 ± 3	464 ± 41
PC-3M cell culture medium		
30% cell confluency[d]	27 ± 3	480 ± 48
60% cell confluency[e]	84 ± 3	>1500[f]
95% cell confluency[g]	201 ± 6	>1500[f]
PC-3M[g] + heparin[h]	43 ± 2	>1500[f]
PC-3M[g] + glycyrrhizin[i]	289 ± 9	>1500[f]

[a] Determined as $k_{ox}^{-1} \, ln2$. This corresponds to the half-life of all-thiol HMGB1.

[b] Determined as $k_{cl}^{-1} \, ln2$. This corresponds to the half-life of disulfide HMGB1.

[c] Control for PC-3M data. No cells were cultured (see "Experimental Procedures").

[d] Extracellular fluid at a cell confluency of 30%.

[e] Extracellular fluid at a cell confluency of 60%.

[f] Too slow to determine.

[g] Extracellular fluid at a cell confluency of 95%.

[h] 2.5 mM heparin octasaccharide dissolved together with HMGB1.

[i] 2.0 mM glycyrrhizin dissolved together with HMGB1.

than that in serum (Figure 4.4a), whereas HMGB1 oxidation in saliva was as fast as that in serum (Figure 4.4b). Thus, the function of different redox forms of extracellular HMGB1 is very likely dependent on the nature of the extracellular fluid.

4.3.5 Kinetics of HMGB1 Oxidation and Clearance in Extracellular Fluids from Cell Culture

Using the *in situ* protein NMR approach, we also investigated the kinetics of HMGB1 oxidation and clearance in extracellular fluids from the culture of the prostate cancer cell line PC-3M.[174] Extracellular fluids at different cell confluences (0, 30, 60, and 95%) were analyzed. Interestingly, the oxidation kinetics were found to strongly depend on the cell confluency. HMGB1 oxidation was slower at a higher confluency (Figure 4.5). At the highest confluency, the halflife of all-thiol HMGB1 was longer than 3 h (Table 4.1), suggesting a more hypoxic environment in the extracellular fluid. In such a hypoxic environment (e.g. in cancer), slow oxidation can suppress the role of disulfide HMGB1.

4.3.6 Impact of Exogenous Ligands on the Oxidation of HMGB1

Since blockade of extracellular HMGB1 signaling was shown to be effective for suppressing tumor growth and metastases,[166] a large number of strategies to block HMGB1 signaling have been proposed for therapeutics of cancer and inflammatory diseases.[160,161,182-186] Compounds that directly bind to HMGB1 were discovered previously. Among them are glycyrrhizin and heparin.[187-190] Using extracellular fluids from PC-3M cell culture (95% confluency), I investigated the impact of these compounds on the oxidation and degradation of HMGB1. The dissociation constant (K_d) for the
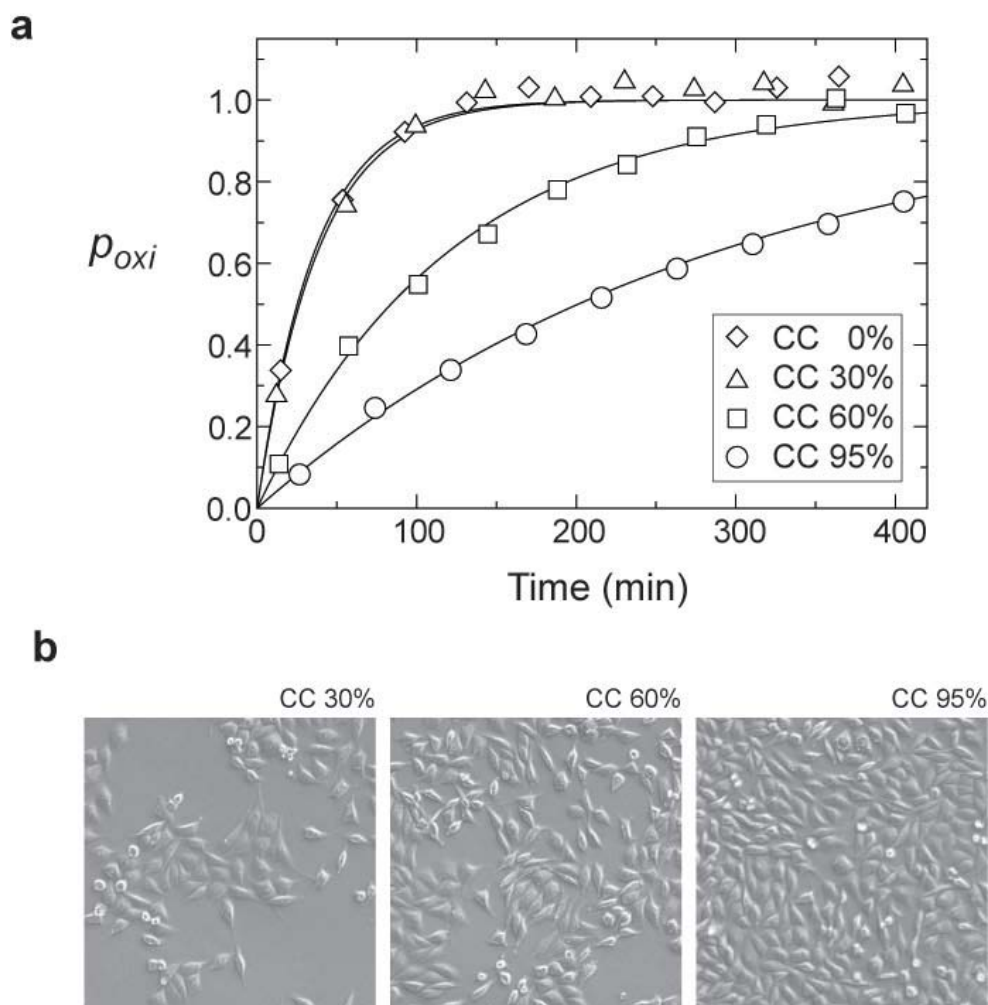
Figure 4.5. HMGB1 in extracellular fluids from PC-3M cells. (a) oxidation kinetics of HMGB1 in extracellular fluids from cultures of the prostate cancer cell line PC-3M studied by in situ protein NMR. Oxidation of HMGB1 was slower at higher cell confluency (CC). (b) images of the cultures when the extracellular fluids were sampled. The total time of incubation at 37 °C was identical for all of the cell culture samples (see "Experimental Procedures"). Different cell confluences were achieved by using different amounts of seeding.

HMGB1-glycyrrhizin complex was reported to be 156 ± 3 µM.[190] By NMR-based titration, we determined the K_d for the HMGB1-heparin octasaccharide complex to be 259 ± 72 µM (Figure 4.6). For these ligands, we compared the kinetics of HMGB1 oxidation for free and ligand-bound HMGB1 in PC-3M extracellular fluids. In these experiments, the concentrations of glycyrrhizin and heparin were 2.0 and 2.5 mM, respectively, at which >90% of HMGB1 should be bound. The oxidation time course data for these samples are shown in Figure 4.7, and the half-lives determined from the data are shown in Table 4.1. Although both glycyrrhizin and heparin are inactive in terms of thiol/disulfide redox chemistry, the binding of these compounds significantly affected the oxidation of HMGB1. The binding of heparin was found to promote the oxidation by 4-fold. It is well known that HMGB1 undergoes conformational equilibrium between the open and closed states via intramolecular interactions between the C-terminal acidic tail and A/B domains.[179,180] Because heparin is also highly negatively charged, HMGB1/heparin interactions may weaken intramolecular interactions between the A-domain and the acidic tail and increase the exposure of the pair of the Cys-23/Cys-45 thiol groups, thereby enhancing their oxidation. On the contrary, glycyrrhizin binding to HMGB1 was found to slow HMGB1 oxidation by ~1.4-fold (Figure 4.7). According to the NMR-based structure model of the HMGB1-glycyrrhizin complex of Mollica *et al.*,[190] glycyrrhizin bound to HMGB1 covers the regions of the Cys-23 and Cys-45 side chains, and so glycyrrhizin may block oxidant accessibility to the Cys-23/Cys-45 thiol groups. A computational study suggested that the HMGB1/glycyrrhizin interactions could be non-directional and dynamic,[189] which may weaken the protection of the thiol pair from oxidants. Our present data suggest that non-redox-active ligands can regulate HMGB1 oxidation.
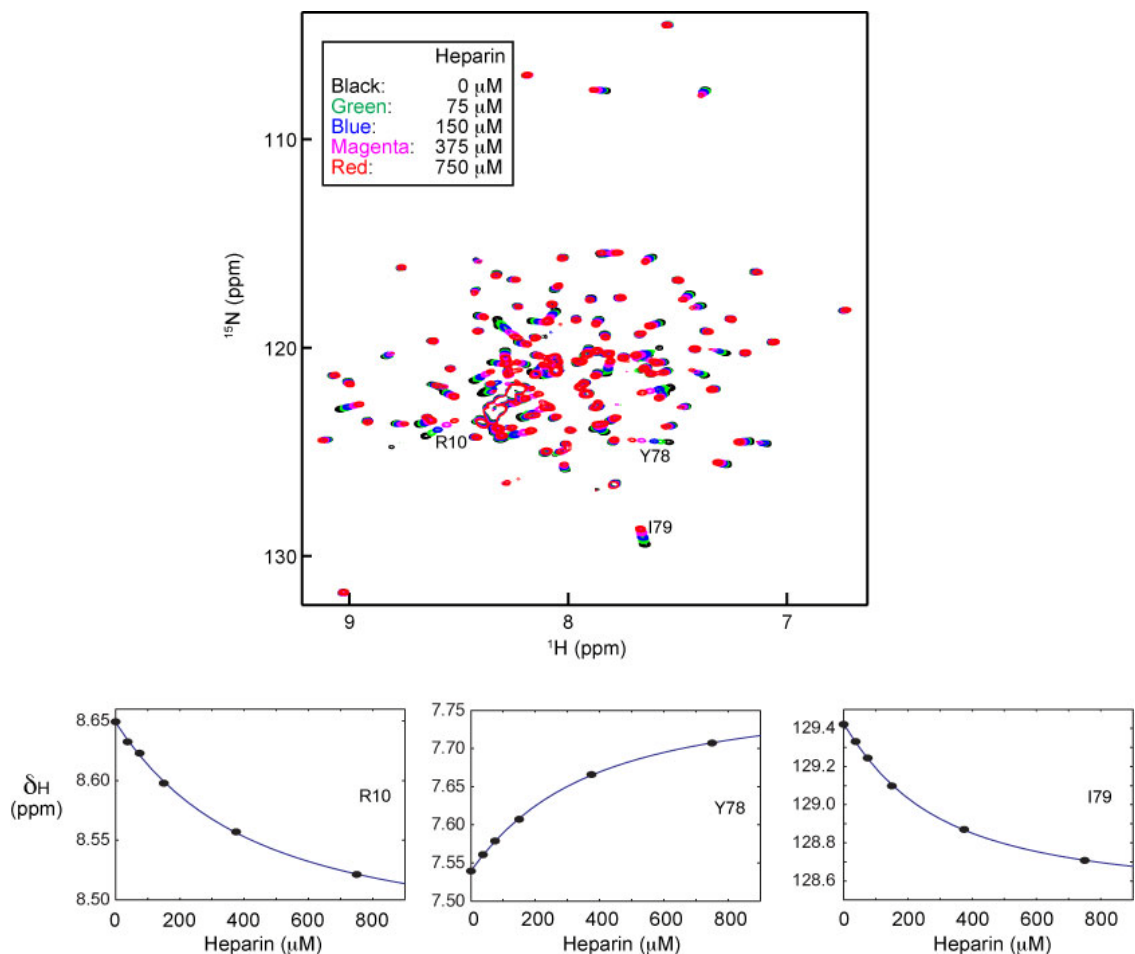
Figure 4.6. HMGB1/heparin interactions observed by NMR. (a) ^{1}H-^{15}N transverse relaxation optimized spectroscopy spectra recorded at 37 °C for 135 µM all-thiol HMGB1 in the presence of heparin octasaccharide at different concentrations. The samples were dissolved in 20 mM Tris-HCl (pH 7.5), 120 mM NaCl, and 5 mM DTT. (b) determination of the dissociation constant (K_d) from ^{1}H and ^{15}N chemical shifts as a function of heparin concentration. The titration data, together with the best fit curves, are shown for three residues whose signals are indicated in (a). Fourteen signals from residues showing large chemical shift perturbation were analyzed, from which the K_d was determined to be 259 $\pm$ 72 µM.
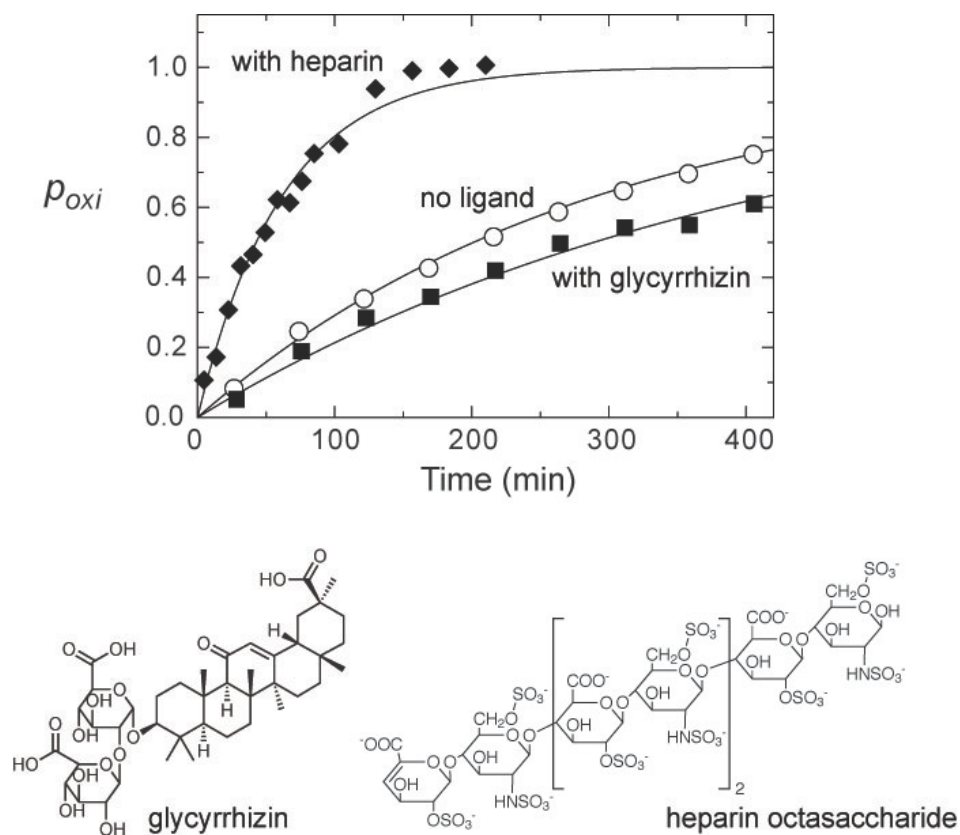
Figure 4.7. Impact of ligand binding on HMGB1 oxidation in extracellular fluids from PC-3M cell culture (95% cell confluency). Time courses of oxidation obtained from in situ protein NMR data are plotted. ○, data for the sample without ligands; ♦ and ■, data for the samples containing heparin octasaccharide (2.5 mM) and glycyrrhizin (2 mM), respectively.

4.4 DISCUSSION

4.4.1 Relevance to Inflammation

All-thiol HMGB1 is considered to recruit leukocytes in inflammation.[162,170,171] Venereau *et al.* found that this role is associated with the binding of the all-thiol HMGB1-CXCL12 heterodimer to CXCR4 on leukocytes. In this study, we have shown that the half-life of all-thiol HMGB1 is only ~17 min in serum. Our data suggest that the half-life can become even shorter via interactions with heparan sulfate, which is structurally analogous to heparin and highly abundant on the cell surface. To recruit leukocytes, the binding of the all-thiol HMGB1-CXCL12 heterodimer to CXCR4 should occur rapidly before the HMGB1 molecule is oxidized. Heterodimer formation with CXCL12 may protect HMGB1 from oxidants. This is possible, as we found that the binding of ligands significantly modulates the kinetics of HMGB1 oxidation. Because of the short half-life of all-thiol HMGB1, most of the extracellular HMGB1 molecules at the injury site should be disulfide HMGB1 when leukocytes arrive as a result of the chemoattractant activity of all-thiol HMGB1. Disulfide HMGB1, but not all-thiol HMGB1, is capable of activating the leukocytes and triggering the release of pro-inflammatory cytokines/chemokines.[162,170,171,173] In this sense, the short half-life of all-thiol HMGB1 in serum seems advantageous for the timely switching of HMGB1 function in inflammation.

4.4.2 Variations in the Half-lives of HMGB1 in Different Extracellular Fluids

We found large variations in the half-lives of HMGB1 in different extracellular fluids. For example, the half-life of all-thiol HMGB1 in PC-3M cell culture medium (95% confluency) was as long as ~3 h, whereas those in serum and saliva were ~17 min. Because all experiments were conducted under ambient oxygen pressure, the kinetic

variation in HMGB1oxidation can be largely due to different levels of redox factors intrinsically present in individual fluids. Our data also indicate that the half-lives of disulfide HMGB1 in serum and saliva differ by 10-fold, most likely due to different degrees of protease activity. These data suggest that the balance between the roles of all-thiol HMGB1 and disulfide HMGB1 depends significantly on the extracellular environment.

4.4.3 Potential Application of *in Situ* Protein NMR

This work has demonstrated that *in situ* protein NMR can provide important insights into the molecular behavior of a particular protein in actual extracellular fluids. The NMR data for ^{15}N-labeled HMGB1 in human serum provided information specific to the behavior of HMGB1, although serum is a molecular crowding fluid with an enormous number of different proteins (60–80 mg/ml) and metabolites. Requiring only 0.50 ml of fluid, *in situ* protein NMR can potentially be used to investigate the molecular behaviors of extracellular proteins in various clinical specimens (e.g. serum, saliva, mucus, etc.). By more detailed analysis, this method has a potential to provide accurate information on proteins' characteristics at pseudo-atomic and atomic resolution, and as a result enabling us to validate observations obtained through *in vitro* experiments.

CHAPTER 5:

Conclusion and Future Directions

Understanding the relationship between protein dynamics and functional kinetics is of fundamental importance in biological sciences. Due to various technical and methodological problems, this relationship remains poorly understood. In this project we pursued three specific aims: 1) to develop methods for protein dynamics analysis on atomic resolution; 2) to elucidate the role of protein dynamics in translocation kinetics; and 3) to find new ways for analysis of protein characteristics in biological fluids. Each of these aims gave us additional understanding of dynamics/function relationship.

For Aim 1, we have developed a novel NMR based method for elucidating the side-chain dynamics of lysine residues. Protein amino-acid side-chains play an essential role in proteins' function. Due to technical difficulties and lack of appropriate biophysical methods, hydrogen-bonding dynamics involving side chains and its role in proteins' function had not been well understood. Our new methods allow us to gain direct insight into functionally important dynamics of lysine side chains in proteins.

On the basis of ^{1}H/^{13}C/^{15}N heteronuclear correlation experiments selective for lysine NH_3^+ groups, we analyzed three different types of long-range ^{15}N-^{13}C J-coupling constants: one between ^{15}Nζ and ^{13}Cγ nuclei ($^3J_{N\zeta C\gamma}$) and the other between ^{15}Nζ and carbonyl ^{13}C' nuclei across a hydrogen bond ($^{h3}J_{N\zeta C'}$). Three bond J-couplings can be interpreted using Karplus equation (chapter 2.3.1), according to which J-coupling is expressed as a function of dihedral angle and three constants (Karplus parameters A, B and C). Initially, we used dihedral angles from ubiquitin's crystal structure and previously reported Karplus parameters to see the correlation between calculated and experimental J-couplings. As expected, correlation between these two was quite poor. But when we

used ensemble average $\langle\, {}^{3}J_{\text{N}\zeta\text{C}\gamma}\rangle$ calculated from the χ_4 angles sampled during a 1 μs MD simulation, we saw significant improvement in correlation with experimental data. This result indicated that lysine side chains have a very dynamic nature and simple use of conformations from crystal structure for analysis of experimental data can be un-appropriate and might result in misinterpretation of data.

After, we used the same approach to measure *J*-coupling between lysine ^{15}Nζ and ^{13}C' of another residue. Relatively low values of *J*-coupling indicated a dynamic and transient nature of these hydrogen bonds, but the fact that they were measurable, was a clear evidence of their existence. Our results were in excellent agreement with crystal structure, which also indicated that side chains of these lysine are in involved in hydrogen bonding. This work shows that our method can be used for other systems to get deeper insight into protein intra- and intermolecular interactions. Our method can also be applied to other residues, such as arginine's guadino groups.

For Aim 2, we have elucidated the role of domain motions in translocation of Egr-1 on nonspecific DNA during target search process. NMR based dynamic studies of Egr-1's behavior in complex with specific and nonspecific DNA revealed very interesting characteristics of this protein: our experiments demonstrated that while in specific complex all three zinc fingers of Egr-1 are firmly attached to DNA, in nonspecific complex ZF1 domain exhibits flexible motions and is mainly dissociated from DNA and ZF2 and ZF3 remain attached to DNA in a way similar to specific complex. In order to understand the role of ZF1's flexibility in Egr-1's translocation, we designed mutant Egr-1 proteins, for which we were able to modulate domain dynamics. More specifically, the first mutant exhibited no domain flexibility in nonspecific complex (type 1), in the second mutant we diminished ZF1's flexibility, but made ZF2 flexible (type 2), and finally in ZF3 we made both ZF1 and ZF3 flexible (type 3). Kinetics studies of Egr-1's translocation on DNA strongly suggested that domain dynamics plays important role in Egr-1's translocation on DNA.

To demonstrate more medical application of our findings, we created Egr-1/FokI fusion protein constructs using wild-type and type-1, -2 and -3 mutants. Zinc-finger protein fusions with FokI nuclease domain (ZFNs) are extensively tested for use in human gene therapy for treatment of various viral diseases. Though ZFNs give very promising results in some cases (e.g. treatment of HIV virus through CCR5 gene mutation), they still have some major problems. Many groups have problems with designing effective ZFNs, as they often exhibit low efficiency in target DNA cleavage, or have high toxicity, in other words high nonspecific cleavage activity. We hypothesized that one of the problems in ZFN technology could be the way zinc finger domains are selected. Since zinc-finger domains are selected based only on their high affinity towards target site, they are biased towards specificity, but lack necessary means for rapid translocation on DNA. We studied DNA cleavage activity of our Egr-1/FokI proteins and found that constructs with flexible zinc-finger domains exhibited considerably high activity compared to construct without flexible zinc fingers. As mutations in zinc fingers should not affect ZFN's cleavage activity, these data suggest that different cleavage activity is caused by difference in DNA translocation kinetics, though we need to do some additional kinetics experiments for confirmation. In other words, our results suggest that enzymatic activity can be modulated and improved through modulation of its zinc-finger domains' dynamics.

In the future, it should be examined whether Egr-1 truly forms a bridging structure between two DNA molecules, or is the difference in DNA translocation between wild-type and mutant Egr-1's caused by other factors. Also, using stopped-flow fluorescence experiments we want to confirm different target search kinetics for different Egr-1/FokI constructs. We demonstrated that our approach was effective in improvement of Egr-1/FokI's DNA cleavage activity, but Egr-1 is a natural protein, which is already optimized for rapid target search. Our next step should be testing effectiveness of this approach for ZFNs used in clinical studies. As artificial ZFNs are designed to have very

strong DNA binding, it might be necessary to take more drastic measures to improve their activity. And finally, it is also important to test our constructs in cell. There are several standard ZFN assays, which can be used to check whether our ZFN constructs exhibit same kinetic differences in cells, that they have *in vitro*.

For Aim 3, we have presented a novel *in situ* methods for analysis of protein characteristics in biological fluids. Using high-mobility group box 1 (HMGB1) protein, we demonstrated how our method can be used to elucidate proteins' functional properties in various biological fluids. We investigated HMGB1's oxidation and clearance rates in human saliva, human blood serum and in extracellular fluids collected at different cell confluences of PC-3M prostate cancer cells. Our results showed that using this *in situ* method one can easily and accurately determine these rates (average error in oxidation rates was ~7%). Results also indicated that oxidation and clearance rates strongly depend on the environment. To our knowledge, our paper published in the *Journal of Biological Chemistry*, where we described this method, is the first work where the oxidation speed of HMGB1 was presented. This information is valuable for drug development since HMGB1 has been a potential target for therapeutic treatment. Our *in situ* method could also be used for medical diagnostics. Different diseases disturb balance of anti-oxidant and pro-oxidant species in blood. HMGB1's oxidation kinetics in blood serum from different patients could be used to detect such abnormal anti-oxidant/pro-oxidant levels. This method can potentially be used to study any other protein which will give appropriate quality NMR spectra in biological fluids. Also, it is possible to do more detailed, pseudo-atomic resolution dynamics analysis of proteins in biological fluids.

In conclusion, we have successfully achieved our three specific aims. Through the three projects described in this dissertation, we have gained better knowledge and methodology to answer the central question: what is relationship between protein dynamics and its functional kinetics? In order to find an answer for this question, it is
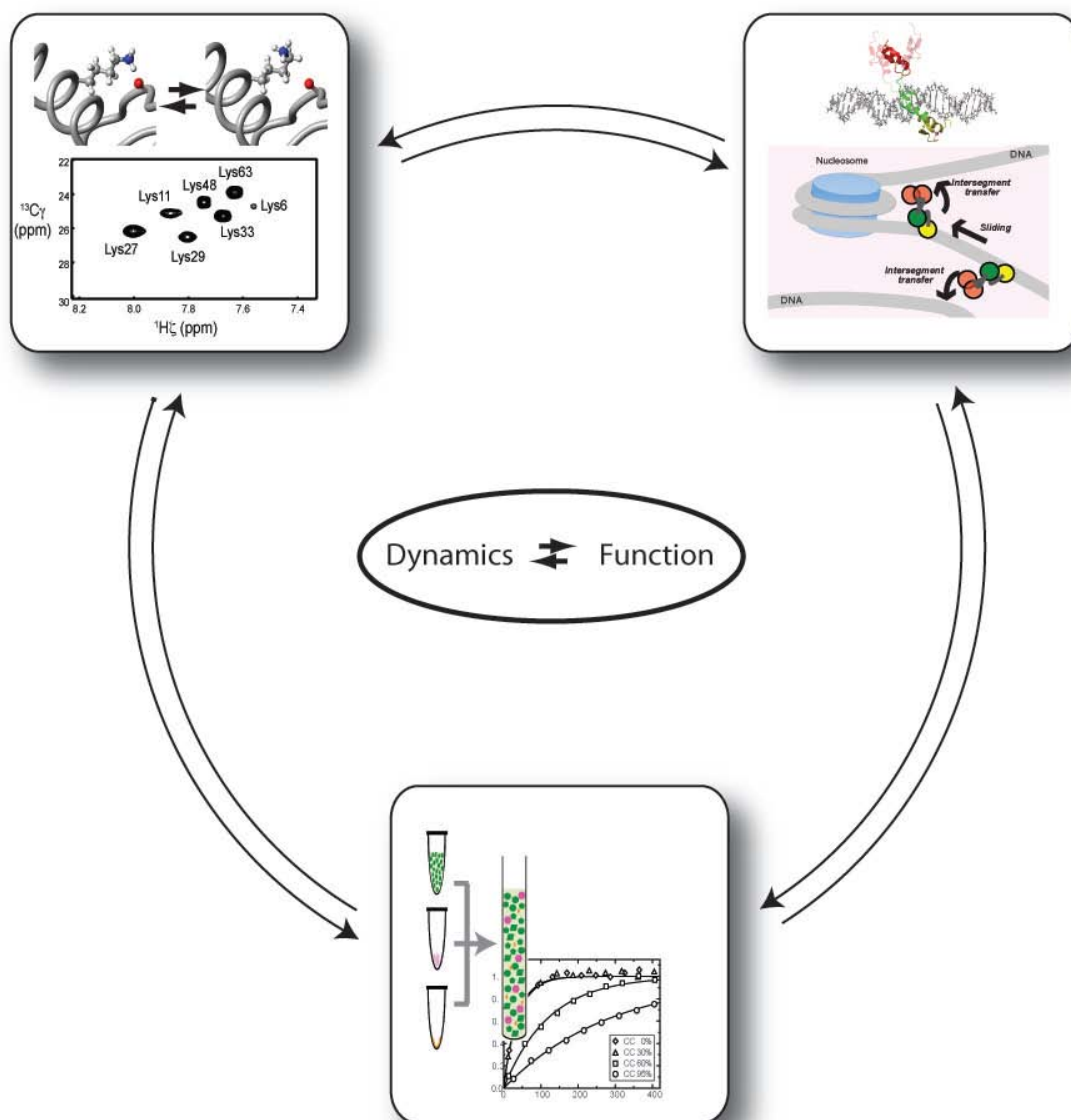
Figure 5.1 Only through simultaneous use of several orthogonal approaches is it possible to gain deeper insight into protein dynamics/function relationship.

essential to realize the importance of pursuing this question from different angles (Figure 5.1). For example, in chapter 2 we elucidated the dynamics of hydrogen bonds involving lysine side chains. Lysine side chains play an important role in proteins' binding to DNA. For protein to translocate from one nonspecific site to another, it needs to break all of its bond. By detailed analysis of hydrogen bonds between Egr-1 and DNA, we will gain deeper insight into determinants of Egr-1's domain dynamics and its translocation kinetics. And the timescale and population of hydrogen bonds provide valuable information for prediction of kinetics of facilitated DNA scanning mechanisms, such as sliding and intersegment transfer. Also, *in situ* NMR method described in chapter 4 can be used to analyze Egr-1's dynamics in more natural environment, to verify *in vitro* observations. Nonspecific interactions and crowding could significantly alter our results, therefore verification of *in vitro* experimental results in biological fluids is essential. With our current method, sensitivity of *J*-coupling measurements of hydrogen bonds in biological fluids would be very poor, but further modification of our methods could make it possible to gain atomic resolution data on protein side chain dynamics in natural environment. There remain still many issues to be addressed in order to gain full understanding of protein dynamics/function relationship, but by pursuing the three aims described in this project, we have made a significant progress to improve our knowledge on this relationship.

Appendix

Copyright Clearance Center RightsLink®

Home Account Info Help

ACS Publications
High quality. High impact.

Title: Signature of Mobile Hydrogen Bonding of Lysine Side Chains from Long-Range 15N–13C Scalar J-Couplings and Computation

Author: Levani Zandarashvili, Da-Wei Li, Tianzhi Wang, Rafael Brüschweiler, and Junji Iwahara

Publication: Journal of the American Chemical Society

Publisher: American Chemical Society

Date: Jun 1, 2011

Copyright © 2011, American Chemical Society

Logged in as:
Levani Zandarashvili
Account #:
3000664097

LOGOUT

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

BACK CLOSE WINDOW

Copyright Clearance Center

RightsLink®

Home   Account Info   Help

jbc

Title: Real-time Kinetics of High-mobility Group Box 1 (HMGB1) Oxidation in Extracellular Fluids Studied by in Situ Protein NMR Spectroscopy

Author: Levani Zandarashvili, Debashish Sahu, Kwanbok Lee, Yong Sun Lee, Pomila Singh, Krishna Rajarathnam, Junji Iwahara

Publication: Journal of Biological Chemistry

Publisher: The American Society for Biochemistry and Molecular Biology

Date: Apr 26, 2013

Copyright © 2013, by the American Society for Biochemistry and Molecular Biology

Logged in as:
Levani Zandarashvili
Account #:
3000664097

LOGOUT

Download Letter of Permission for Reuse in Thesis/Dissertation

The publisher allows article reuse in thesis or dissertations at no charge. Click here to receive a letter of permission for this purpose.

BACK   CLOSE WINDOW

Bibliography/References

1   Berman, H. M., Coimbatore Narayanan, B., Di Costanzo, L., Dutta, S., Ghosh, S., Hudson, B. P., Lawson, C. L., Peisach, E., Prlic, A. *et al.* Trendspotting in the Protein Data Bank. *FEBS Letters* **587**, 1036-1045,(2013).

2   Zandarashvili, L., Vuzman, D., Esadze, A., Takayama, Y., Sahu, D., Levy, Y. & Iwahara, J. Asymmetrical roles of zinc fingers in dynamic DNA-scanning process by the inducible transcription factor Egr-1. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E1724-1732,(2012).

3   Zandarashvili, L., Sahu, D., Lee, K., Lee, Y. S., Singh, P., Rajarathnam, K. & Iwahara, J. Real-time Kinetics of High-mobility Group Box 1 (HMGB1) Oxidation in Extracellular Fluids Studied by in Situ Protein NMR Spectroscopy. *Journal of Biological Chemistry* **288**, 11621-11627,(2013).

4   Zandarashvili, L., Li, D. W., Wang, T., Bruschweiler, R. & Iwahara, J. Signature of mobile hydrogen bonding of lysine side chains from long-range 15N-13C scalar J-couplings and computation. *Journal of the American Chemical Society* **133**, 9192-9195,(2011).

5   Yonetani, T. & Laberge, M. Protein dynamics explain the allosteric behaviors of hemoglobin. *Biochimica et Biophysica Acta* **1784**, 1146-1158,(2008).

6   Wrabl, J. O., Gu, J., Liu, T., Schrank, T. P., Whitten, S. T. & Hilser, V. J. The role of protein conformational fluctuations in allostery, function, and evolution. *Biophysical Chemistry* **159**, 129-141,(2011).

7   Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z. & Kern, D. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nature Structural Molecular Biology* **11**, 945-949,(2004).

8   Villali, J. & Kern, D. Choreographing an enzyme's dance. *Current Opinion in Chemical Biology* **14**, 636-643,(2010).

9   Tzeng, S. R. & Kalodimos, C. G. Protein dynamics and allostery: an NMR view. *Current Opinion in Structural Biology* **21**, 62-67,(2011).

10   Teilum, K., Olsen, J. G. & Kragelund, B. B. Protein stability, flexibility and function. *Biochimica et Biophysica Acta* **1814**, 969-976,(2011).

11   Sapienza, P. J. & Lee, A. L. Using NMR to study fast dynamics in proteins: methods and applications. *Current Opinion in Pharmacology* **10**, 723-730,(2010).

12   Salsbury, F. R., Jr. Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. *Current Opinion in Pharmacology* **10**, 738-744,(2010).

13   Otten, R., Villali, J., Kern, D. & Mulder, F. A. Probing microsecond time scale dynamics in proteins by methyl (1)H Carr-Purcell-Meiboom-Gill relaxation dispersion NMR measurements. Application to activation of the signaling protein NtrC(r). *Journal of the American Chemical Society* **132**, 17004-17014,(2010).

14   Osawa, M., Takeuchi, K., Ueda, T., Nishida, N. & Shimada, I. Functional dynamics of proteins revealed by solution NMR. *Current Opinion in Structural Biology* **22**, 660-669,(2012).

15   Ogino, S., Nishida, N., Umemoto, R., Suzuki, M., Takeda, M., Terasawa, H., Kitayama, J., Matsumoto, M., Hayasaka, H. *et al.* Two-state conformations in the hyaluronan-binding domain regulate CD44 adhesiveness under flow condition. *Structure* **18**, 649-656,(2010).

16   McGeagh, J. D., Ranaghan, K. E. & Mulholland, A. J. Protein dynamics and enzyme catalysis: insights from simulations. *Biochimica et Biophysica Acta* **1814**, 1077-1092,(2011).

17   Liu, T., Whitten, S. T. & Hilser, V. J. Functional residues serve a dominant role in mediating the cooperativity of the protein ensemble. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 4347-4352,(2007).

18   Liberles, D. A., Teichmann, S. A., Bahar, I., Bastolla, U., Bloom, J., Bornberg-Bauer, E., Colwell, L. J., de Koning, A. P., Dokholyan, N. V. *et al.* The interface of protein structure, protein biophysics, and molecular evolution. *Protein Science* **21**, 769-785,(2012).

19   Jarymowycz, V. A. & Stone, M. J. Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. *Chemical Reviews* **106**, 1624-1671,(2006).

20   Igumenova, T. I., Frederick, K. K. & Wand, A. J. Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. *Chemical Reviews* **106**, 1672-1699,(2006).

21   Henzler-Wildman, K. A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., Pozharski, E., Wilson, M. A., Petsko, G. A. *et al.* Intrinsic motions along an enzymatic reaction trajectory. *Nature* **450**, 838-844,(2007).

22   Henzler-Wildman, K. A., Lei, M., Thai, V., Kerns, S. J., Karplus, M. & Kern, D. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913-916,(2007).

23   Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964-972,(2007).

24   Dorn, J. F. & Maddox, P. S. Kinetochore dynamics: how protein dynamics affect chromosome segregation. *Current Opinion in Cell Biology* **24**, 57-63,(2012).

25   Cargnello, M., Tcherkezian, J., Dorn, J. F., Huttlin, E. L., Maddox, P. S., Gygi, S. P. & Roux, P. P. Phosphorylation of the eukaryotic translation initiation factor 4E-transporter (4E-T) by c-Jun N-terminal kinase promotes stress-dependent P-body assembly. *Molecular and Cellular Biology* **32**, 4572-4584,(2012).

26   Bu, Z. & Callaway, D. J. Proteins move! Protein dynamics and long-range allostery in cell signaling. *Advances in protein chemistry and structural biology* **83**, 163-221,(2011).

27   Boehr, D. D., McElheny, D., Dyson, H. J. & Wright, P. E. Millisecond timescale fluctuations in dihydrofolate reductase are exquisitely sensitive to the bound ligands. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1373-1378,(2010).

28   Benkovic, S. J. & Hammes-Schiffer, S. A perspective on enzyme catalysis. *Science* **301**, 1196-1202,(2003).

29   Jardetzky, O. Protein dynamics and conformational transitions in allosteric proteins. *Progress in Biophysics and Molecular Biology* **65**, 171-219,(1996).

30   Taylor, S. S., Kim, C., Vigil, D., Haste, N. M., Yang, J., Wu, J. & Anand, G. S. Dynamics of signaling by PKA. *Biochimica et Biophysica Acta* **1754**, 25-37,(2005).

31   Teriete, P., Banerji, S., Noble, M., Blundell, C. D., Wright, A. J., Pickford, A. R., Lowe, E., Mahoney, D. J., Tammi, M. I. *et al.* Structure of the regulatory hyaluronan binding domain in the inflammatory leukocyte homing receptor CD44. *Molecular Cell* **13**, 483-496,(2004).

32   Takeda, M., Terasawa, H., Sakakura, M., Yamaguchi, Y., Kajiwara, M., Kawashima, H., Miyasaka, M. & Shimada, I. 1H, 13C and 15N backbone resonance assignments of the hyaluronan-binding domain of CD44. *Journal of Biomolecular NMR* **29**, 97-98,(2004).

33   Takeda, M., Ogino, S., Umemoto, R., Sakakura, M., Kajiwara, M., Sugahara, K. N., Hayasaka, H., Miyasaka, M., Terasawa, H. *et al.* Ligand-induced structural changes of the CD44 hyaluronan-binding domain revealed by NMR. *Journal of Biological Chemistry* **281**, 40089-40095,(2006).

34   Ponta, H., Sherman, L. & Herrlich, P. A. CD44: from adhesion molecules to signalling regulators. *Nature Reviews Molecular Cell Biology* **4**, 33-45,(2003).

35   Banerji, S., Wright, A. J., Noble, M., Mahoney, D. J., Campbell, I. D., Day, A. J. & Jackson, D. G. Structures of the Cd44-hyaluronan complex provide insight into a fundamental carbohydrate-protein interaction. *Nature Structural Molecular Biology* **14**, 234-239,(2007).

36   Iwahara, J. & Clore, G. M. Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* **440**, 1227-1230,(2006).

37   Clore, G. M., Tang, C. & Iwahara, J. Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Current Opinion in Structural Biology* **17**, 603-616,(2007).

38   Palmer, A. G., 3rd, Kroenke, C. D. & Loria, J. P. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods in Enzymology* **339**, 204-238,(2001).

39   Anderson, K. M., Esadze, A., Manoharan, M., Bruschweiler, R., Gorenstein, D. G. & Iwahara, J. Direct observation of the ion-pair dynamics at a protein-DNA interface by NMR spectroscopy. *Journal of the American Chemical Society* **135**, 3613-3619,(2013).

40   Esadze, A., Li, D. W., Wang, T., Bruschweiler, R. & Iwahara, J. Dynamics of lysine side-chain amino groups in a protein studied by heteronuclear 1H-15N NMR spectroscopy. *Journal of the American Chemical Society* **133**, 909-919,(2011).

41   Takayama, Y., Sahu, D. & Iwahara, J. Observing in-phase single-quantum 15N multiplets for NH2/NH3+ groups with two-dimensional heteronuclear correlation spectroscopy. *Journal of Magnetic Resonance* **194**, 313-316,(2008).

42   Takayama, Y., Castaneda, C. A., Chimenti, M., Garcia-Moreno, B. & Iwahara, J. Direct evidence for deprotonation of a lysine side chain buried in the hydrophobic core of a protein. *Journal of the American Chemical Society* **130**, 6714-6715,(2008).

43   Iwahara, J., Jung, Y. S. & Clore, G. M. Heteronuclear NMR spectroscopy for lysine NH(3) groups in proteins: unique effect of water exchange on (15)N transverse relaxation. *Journal of the American Chemical Society* **129**, 2971-2980,(2007).

44   Elcock, A. H. Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Current Opinion in Structural Biology* **20**, 196-206,(2010).

45   Woronowicz, K., Sha, D., Frese, R. N., Sturgis, J. N., Nanda, V. & Niederman, R. A. The effects of protein crowding in bacterial photosynthetic membranes on the flow of quinone redox species between the photochemical reaction center and the ubiquinol-cytochrome c2 oxidoreductase. *Metallomics* **3**, 765-774,(2011).

46   Zhou, H. X. Influence of crowded cellular environments on protein folding, binding, and oligomerization: biological consequences and potentials of atomistic modeling. *FEBS Letters* **587**, 1053-1061,(2013).

47   Binolfi, A., Theillet, F. X. & Selenko, P. Bacterial in-cell NMR of human alpha-synuclein: a disordered monomer by nature? *Biochemical Society Transactions* **40**, 950-954,(2012).

48   Inomata, K., Ohno, A., Tochio, H., Isogai, S., Tenno, T., Nakase, I., Takeuchi, T., Futaki, S., Ito, Y. *et al.* High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **458**, 106-109,(2009).

49   Aden, J., Verma, A., Schug, A. & Wolf-Watz, M. Modulation of a pre-existing conformational equilibrium tunes adenylate kinase activity. *Journal of the American Chemical Society* **134**, 16562-16570,(2012).

50   Kalodimos, C. G., Boelens, R. & Kaptein, R. Toward an integrated model of protein-DNA recognition as inferred from NMR studies on the Lac repressor system. *Chemical Reviews* **104**, 3567-3586,(2004).

51   Kalodimos, C. G., Biris, N., Bonvin, A. M., Levandoski, M. M., Guennuegues, M., Boelens, R. & Kaptein, R. Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science* **305**, 386-389,(2004).

52   Kalodimos, C. G., Bonvin, A. M., Salinas, R. K., Wechselberger, R., Boelens, R. & Kaptein, R. Plasticity in protein-DNA recognition: lac repressor interacts with its natural operator 01 through alternative conformations of its DNA-binding domain. *EMBO Journal* **21**, 2866-2876,(2002).

53   Gong, S., Worth, C. L., Bickerton, G. R., Lee, S., Tanramluk, D. & Blundell, T. L. Structural and functional restraints in the evolution of protein families and superfamilies. *Biochemical Society Transactions* **37**, 727-733,(2009).

54   Zandarashvili, L., Li, D. W., Wang, T., Bruschweiler, R. & Iwahara, J. Signature of Mobile Hydrogen Bonding of Lysine Side-Chains from Long-Range 15N-13C Scalar J-Couplings and Computation. *Journal of the American Chemical Society*,(2011).

55   Kupče, E., Boyd, J. & Campbell, I. D. Short selective pulses for biochemical applications. *Journal of Magnetic Resonance B* **106**, 300-303,(1995).

56   Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes. *Journal of Biomolecular NMR* **6**, 277-293,(1995).

57   Johnson, B. A. & Blevins, R. A. Nmr View - a Computer-Program for the Visualization and Analysis of Nmr Data. *Journal of Biomolecular NMR* **4**, 603-614,(1994).

58   Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R. & Zhu, G. Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods in Enzymology* **239**, 79-105,(1994).

59   Hu, J. S. & Bax, A. χ_1 angle information from a simple two-dimensional NMR experiment that identifies trans $^3J_{NC\gamma}$ couplings in isotopically enriched proteins. *Journal of Biomolecular NMR* **9**, 323-328,(1997).

60   Hu, J. S., Grzesiek, S. & Bax, A. Two-dimensional NMR methods for determining (chi 1) angles of aromatic residues in proteins from three-bond J(C'C gamma) and J(NC gamma) couplings. *Journal of the American Chemical Society* **119**, 1803-1804,(1997).

61   Karplus, M. Contact electron-spin coupling of nuclear magnetic moments. *Journal of Chemical Physics* **30**, 11-15,(1959).

62   Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, Rance, M. & Skelton, N. J. *Protein NMR Sprectroscopy: Principles and Practice*. 2 edn, (Elsevier Academic Press, 2007).

63   Perez, C., Löhr, F., Rüterjans, H. & Schmidt, J. M. Self-consistent Karplus parametrization of 3J couplings depending on the polypeptide side-chain torsion chi1. *Journal of the American Chemical Society* **123**, 7081-7093,(2001).

64   Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. Structure of ubiquitin refined at 1.8 A resolution. *Journal of Molecular Biology* **194**, 531-544,(1987).

65   Pachler, K. G. R. Nuclear magnetic resonance study of some α amino acids. I. Coupling constants in alkaline and acidic medium. *Spectrochimica Acta* **19**, 2085-2092,(1963).

66   Pachler, K. G. R. Nuclear magnetic resonance study of some α amino acids. II. Rotational isomerism. *Spectrochimica Acta* **20**, 581-587,(1964).

67   Gibbons, W. A., Nemethy, G., Stern, A. & Craig, L. C. An approach to conformational analysis of peptides and proteins in solution based on a combination of nuclear magnetic resonance spectroscopy and conformational energy calculations. *Proceedings of the National Academy of Sciences of the United States of America* **67**, 239-246,(1970).

68   Rinkel, L. J. & Altona, C. Conformational analysis of the deoxyribofuranose ring in DNA by means of sums of proton-proton coupling constants: a graphical method. *Journal of Biomolecular Structure and Dynamics* **4**, 621-649,(1987).

69   Madi, Z. L., Griesinger, C. & Ernst, R. R. Conformational Dynamics of Proline Residues in Antamanide - J-Coupling Analysis of Strongly Coupled Spin Systems Based on E-Cosy Spectra. *Journal of the American Chemical Society* **112**, 2908-2914,(1990).

70   Dzakula, Z., Edison, A. S., Westler, W. M. & Markley, J. L. Analysis of Chi-1 Rotamer Populations from Nmr Data by the Cupid Method. *Journal of the American Chemical Society* **114**, 6200-6207,(1992).

71   Dzakula, Z., Westler, W. M., Edison, A. S. & Markley, J. L. The Cupid Method for Calculating the Continuous Probability-Distribution of Rotamers from Nmr Data. *Journal of the American Chemical Society* **114**, 6195-6199,(1992).

72   van Wijk, J., Huckriede, B. D., Ippel, J. H. & Altona, C. Furanose sugar conformations in DNA from NMR coupling constants. *Methods in Enzymology* **211**, 286-306,(1992).

73   Brunne, R. M., Vangunsteren, W. F., Brüschweiler, R. & Ernst, R. R. Molecular-Dynamics Simulation of the Proline Conformational Equilibrium and Dynamics in Antamanide Using the Gromos Force-Field. *Journal of the American Chemical Society* **115**, 4764-4768,(1993).

74   Schmidt, J. M., Brüschweiler, R., Ernst, R. R., Dunbrack, R. L., Joseph, D. & Karplus, M. Molecular-Dynamics Simulation of the Proline Conformational Equilibrium and Dynamics in Antamanide Using the Charmm Force-Field. *Journal of the American Chemical Society* **115**, 8747-8756,(1993).

75   Brüschweiler, R. & Case, D. A. Adding Harmonic Motion to the Karplus Relation for Spin-Spin Coupling. *Journal of the American Chemical Society* **116**, 11199-11200,(1994).

76   Brüschweiler, R. & Wright, P. E. NMR Order Parameters of Biomolecules - a New Analytical Representation and Application to the Gaussian Axial Fluctuation Model. *Journal of the American Chemical Society* **116**, 8426-8427,(1994).

77   Case, D. A., Scheurer, C. & Bruschweiler, R. Static and dynamic effects on vicinal scalar J couplings in proteins and peptides: A MD/DFT analysis. *Journal of the American Chemical Society* **122**, 10390-10397,(2000).

78   Chou, J. J., Case, D. A. & Bax, A. Insights into the mobility of methyl-bearing side chains in proteins from (3)J(CC) and (3)J(CN) couplings. *Journal of the American Chemical Society* **125**, 8959-8966,(2003).

79   Best, R. B. & Hummer, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J Phys Chem B* **113**, 9004-9015,(2009).

80   Wickstrom, L., Okur, A. & Simmerling, C. Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. *Biophysical Journal* **97**, 853-856,(2009).

81   Li, D. W. & Brüschweiler, R. NMR-based protein potentials. *Angewandte Chemie. International Ed. In English* **49**, 6778-6780,(2010).

82   Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950-1958,(2010).

83   Markwick, P. R. L., Showalter, S. A., Bouvignies, G., Brüschweiler, R. & Blackledge, M. Structural dynamics of protein backbone phi angles: extended molecular dynamics simulations versus experimental [3]J scalar couplings. *Journal of Biomolecular NMR* **45**, 17-21,(2009).

84   Cordier, F. & Grzesiek, S. Direct observation of hydrogen bonds in proteins by interresidue (3h)J(NC ') scalar couplings. *Journal of the American Chemical Society* **121**, 1601-1602,(1999).

85   Cornilescu, G., Hu, J. S. & Bax, A. Identification of the hydrogen bonding network in a protein by scalar couplings. *Journal of the American Chemical Society* **121**, 2949-2950,(1999).

86   Cornilescu, G., Ramirez, B. E., Frank, M. K., Clore, G. M., Gronenborn, A. M. & Bax, A. Correlation between (3h)J(NC ') and hydrogen bond length in proteins. *Journal of the American Chemical Society* **121**, 6275-6279,(1999).

87   Wang, Y. X., Jacob, J., Cordier, F., Wingfield, P., Stahl, S. J., Lee-Huang, S., Torchia, D., Grzesiek, S. & Bax, A. Measurement of (3h)J(NC ') connectivities across hydrogen bonds in a 30 kDa protein. *Journal of Biomolecular NMR* **14**, 181-184,(1999).

88   Grzesiek, S., Cordier, F., Jaravine, V. & Barfield, M. Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings. *Progress in Nuclear Magnetic Resonance Spectroscopy* **45**, 275-300,(2004).

89   Scheurer, C. & Brüschweiler, R. Quantum-chemical characterization of nuclear spin-spin couplings across hydrogen bonds. *Journal of the American Chemical Society* **121**, 8661-8662,(1999).

90   Czernek, J. & Brüschweiler, R. Geometric dependence of (3h)J(P-31-N-15) and (2h)J(P-31-H-1) scalar couplings in protein-nucleotide complexes. *Journal of the American Chemical Society* **123**, 11079-11080,(2001).

91   Barfield, M. Structural dependencies of interresidue scalar coupling (h3)J(NC), and donor H-1 chemical shifts in the hydrogen bonding regions of proteins. *Journal of the American Chemical Society* **124**, 4158-4168,(2002).

92   Bouvignies, G., Bernado, P., Meier, S., Cho, K., Grzesiek, S., Brüschweiler, R. & Blackledge, M. Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 13885-13890,(2005).

93   Juranic, N., Atanasova, E., Streiff, J. H., Macura, S. & Prendergast, F. G. Solvent-induced differentiation of protein backbone hydrogen bonds in calmodulin. *Protein Science* **16**, 1329-1337,(2007).

94   Krukenberg, K. A., Street, T. O., Lavery, L. A. & Agard, D. A. Conformational dynamics of the molecular chaperone Hsp90. *Quarterly Reviews of Biophysics* **44**, 229-255,(2011).

95   Lamboy, J. A., Kim, H., Dembinski, H., Ha, T. & Komives, E. A. Single-Molecule FRET Reveals the Native-State Dynamics of the IkappaBalpha Ankyrin Repeat Domain. *Journal of Molecular Biology*,(2013).

96   Wang, H., Chalovich, J. M. & Marriott, G. Structural dynamics of troponin I during Ca2+-activation of cardiac thin filaments: a multi-site Forster resonance energy transfer study. *PLoS One* **7**, e50420,(2012).

97   Bozon, B., Davis, S. & Laroche, S. A requirement for the immediate early gene zif268 in reconsolidation of recognition memory after retrieval. *Neuron* **40**, 695-701,(2003).

98   Lee, J. L., Everitt, B. J. & Thomas, K. L. Independent cellular processes for hippocampal memory consolidation and reconsolidation. *Science* **304**, 839-843,(2004).

99   Khachigian, L. M., Lindner, V., Williams, A. J. & Collins, T. Egr-1-induced endothelial gene expression: a common theme in vascular injury. *Science* **271**, 1427-1431,(1996).

100   Yan, S. F., Fujita, T., Lu, J., Okada, K., Shan Zou, Y., Mackman, N., Pinsky, D. J. & Stern, D. M. Egr-1, a master switch coordinating upregulation of divergent gene families underlying ischemic stress. *Nature Medicine* **6**, 1355-1361,(2000).

101   Pavletich, N. P. & Pabo, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**, 809-817,(1991).

102   Slutsky, M. & Mirny, L. A. Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophysical Journal* **87**, 4021-4035,(2004).

103   Mirny, L., Slutsky, M., Wunderlich, Z., Tafvizi, A., Leith, J. & Kosmrlj, A. How a protein searches for its site on DNA: the mechanism of facilitated diffusion. *Journal of Physics A: Mathematical and Theoretical* **42**, 401335,(2009).

104   Elrod-Erickson, M., Rould, M. A., Nekludova, L. & Pabo, C. O. Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure* **4**, 1171-1180,(1996).

105   Foster, M. P., Wuttke, D. S., Radhakrishnan, I., Case, D. A., Gottesfeld, J. M. & Wright, P. E. Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA. *Nature Structural Biology* **4**, 605-608,(1997).

106   Laity, J. H., Dyson, H. J. & Wright, P. E. DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers. *Journal of Molecular Biology* **295**, 719-727,(2000).

107   Laity, J. H., Dyson, H. J. & Wright, P. E. Molecular basis for modulation of biological function by alternate splicing of the Wilms' tumor suppressor protein. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 11932-11935,(2000).

108   Wolfe, S. A., Nekludova, L. & Pabo, C. O. DNA recognition by Cys_2His_2 zinc finger proteins. *Annual Review of Biophysics and Biomolecular Structure* **29**, 183-212,(2000).

109   Stoll, R., Lee, B. M., Debler, E. W., Laity, J. H., Wilson, I. A., Dyson, H. J. & Wright, P. E. Structure of the Wilms tumor suppressor protein zinc finger domain bound to DNA. *Journal of Molecular Biology* **372**, 1227-1245,(2007).

110   von Hippel, P. H. & Berg, O. G. Facilitated target location in biological systems. *Journal of Biological Chemistry* **264**, 675-678,(1989).

111   Halford, S. E. & Marko, J. F. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Research* **32**, 3040-3052,(2004).

112   Mazari, P. M. & Roth, M. J. Library screening and receptor-directed targeting of gammaretroviral vectors. *Future Microbiology* **8**, 107-121,(2013).

113   Lim, S. P., Neilsen, P., Kumar, R., Abell, A. & Callen, D. F. The application of delivery systems for DNA methyltransferase inhibitors. *BioDrugs* **25**, 227-242,(2011).

114   Tong, A. W. & Nemunaitis, J. Modulation of miRNA activity in human cancer: a new paradigm for cancer gene therapy? *Cancer Gene Therapy* **15**, 341-355,(2008).

115   Camenisch, T. D., Brilliant, M. H. & Segal, D. J. Critical parameters for genome editing using zinc finger nucleases. *Mini-Reviews in Medicinal Chemistry* **8**, 669-676,(2008).

116   Carroll, D. Progress and prospects: zinc-finger nucleases as gene therapy agents. *Gene Therapy* **15**, 1463-1468,(2008).

117   Carroll, D. Zinc-finger nucleases: a panoramic view. *Current Gene Therapy* **11**, 2-10,(2011).

118   Carroll, D. Genome engineering with zinc-finger nucleases. *Genetics* **188**, 773-782,(2011).

119   Cathomen, T. & Joung, J. K. Zinc-finger nucleases: the next generation emerges. *Molecular Therapy* **16**, 1200-1207,(2008).

120   Davis, D. & Stokoe, D. Zinc finger nucleases as tools to understand and treat human diseases. *BMC Medicine* **8**, 42,(2010).

121   Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M. H. & Chandrasegaran, S. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Research* **33**, 5978-5990,(2005).

122   Le Provost, F., Lillico, S., Passet, B., Young, R., Whitelaw, B. & Vilotte, J. L. Zinc finger nuclease technology heralds a new era in mammalian transgenesis. *Trends in Biotechnology* **28**, 134-141,(2010).

123   Greisman, H. A. & Pabo, C. O. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science* **275**, 657-661,(1997).

124   Guan, X., Stege, J., Kim, M., Dahmani, Z., Fan, N., Heifetz, P., Barbas, C. F., 3rd & Briggs, S. P. Heritable endogenous gene regulation in plants with designed polydactyl zinc finger transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 13296-13301,(2002).

125   Porteus, M. H. & Carroll, D. Gene targeting using zinc finger nucleases. *Nature Biotechnology* **23**, 967-973,(2005).

126   Rahman, S. H., Maeder, M. L., Joung, J. K. & Cathomen, T. Zinc-finger nucleases for somatic gene therapy: the next frontier. *Human Gene Therapy* **22**, 925-933,(2011).

127   Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics* **11**, 636-646,(2010).

128   Wu, J., Kandavelou, K. & Chandrasegaran, S. Custom-designed zinc finger nucleases: what is next? *Cellular and Molecular Life Sciences* **64**, 2933-2944,(2007).

129   Kaczorowski, T., Skowron, P. & Podhajska, A. J. Purification and characterization of the FokI restriction endonuclease. *Gene* **80**, 209-216,(1989).

130   Looney, M. C., Moran, L. S., Jack, W. E., Feehery, G. R., Benner, J. S., Slatko, B. E. & Wilson, G. G. Nucleotide sequence of the FokI restriction-modification system: separate strand-specificity domains in the methyltransferase. *Gene* **80**, 193-208,(1989).

131   Galvani, A. P. & Novembre, J. The evolutionary history of the CCR5-Delta32 HIV-resistance mutation. *Microbes and Infection* **7**, 302-309,(2005).

132   Le Douce, V., Janossy, A., Hallay, H., Ali, S., Riclet, R., Rohr, O. & Schwartz, C. Achieving a cure for HIV infection: do we have reasons to be optimistic? *Journal of Antimicrobial Chemotherapy* **67**, 1063-1074,(2012).

133   Kim, H. J., Lee, H. J., Kim, H., Cho, S. W. & Kim, J. S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Research* **19**, 1279-1288,(2009).

134   Cradick, T. J., Keck, K., Bradshaw, S., Jamieson, A. C. & McCaffrey, A. P. Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis B virus DNAs. *Molecular Therapy* **18**, 947-954,(2010).

135   Shimizu, Y., Sollu, C., Meckler, J. F., Adriaenssens, A., Zykovich, A., Cathomen, T. & Segal, D. J. Adding fingers to an engineered zinc finger nuclease can reduce activity. *Biochemistry* **50**, 5033-5041,(2011).

136   Zhou, H. X. & Szabo, A. Enhancement of association rates by nonspecific binding to DNA and cell membranes. *Physical Review Letters* **93**, 178101,(2004).

137   Pattanayak, V., Ramirez, C. L., Joung, J. K. & Liu, D. R. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nature Methods* **8**, 765-770,(2011).

138   Pervushin, K. Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology. *Quarterly Reviews of Biophysics* **33**, 161-197,(2000).

139   Palmer, A. G., 3rd. NMR probes of molecular dynamics: overview and comparison with other techniques. *Annual Review of Biophysics and Biomolecular Structure* **30**, 129-155,(2001).

140   Iwahara, J., Zweckstetter, M. & Clore, G. M. NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15062-15067,(2006).

141   Baber, J. L., Szabo, A. & Tjandra, N. Analysis of slow interdomain motion of macromolecules using NMR relaxation data. *Journal of the American Chemical Society* **123**, 3953-3959,(2001).

142   Loria, J. P., Rance, M. & Palmer, A. G. A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *Journal of the American Chemical Society* **121**, 2331-2332,(1999).

143   Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D. & Kay, L. E. Slow dynamics in folded and unfolded states of an SH3 domain. *Journal of the American Chemical Society* **123**, 11341-11352,(2001).

144   Vuzman, D., Polonsky, M. & Levy, Y. Facilitated DNA search by multidomain transcription factors: cross talk via a flexible linker. *Biophysical Journal* **99**, 1202-1211,(2010).

145   Elrod-Erickson, M., Benson, T. E. & Pabo, C. O. High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition. *Structure* **6**, 451-464,(1998).

146   Miller, J. C. & Pabo, C. O. Rearrangement of side-chains in a Zif268 mutant highlights the complexities of zinc finger-DNA recognition. *Journal of Molecular Biology* **313**, 309-315,(2001).

147   Peisach, E. & Pabo, C. O. Constraints for zinc finger linker design as inferred from X-ray crystal structure of tandem Zif268-DNA complexes. *Journal of Molecular Biology* **330**, 1-7,(2003).

148   Wolfe, S. A., Grant, R. A., Elrod-Erickson, M. & Pabo, C. O. Beyond the "recognition code": structures of two Cys_2His_2 zinc finger/TATA box complexes. *Structure* **9**, 717-723,(2001).

149   Fried, M. G. & Crothers, D. M. Kinetics and mechanism in the reaction of gene regulatory proteins with DNA. *Journal of Molecular Biology* **172**, 263-282,(1984).

150   Iwahara, J. & Clore, G. M. 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,(2006).

151   Lieberman, B. A. & Nordeen, S. K. DNA intersegment transfer, how steroid receptors search for a target site. *Journal of Biological Chemistry* **272**, 1061-1068,(1997).

152 Skowron, P., Kaczorowski, T., Tucholski, J. & Podhajska, A. J. Atypical DNA-binding properties of class-IIS restriction endonucleases: evidence for recognition of the cognate sequence by a FokI monomer. *Gene* **125**, 1-10,(1993).

153 Takayama, Y. & Clore, G. M. Intra- and intermolecular translocation of the bi-domain transcription factor Oct1 characterized by liquid crystal and paramagnetic NMR. *Proceedings of the National Academy of Sciences of the United States of America* **108**, E169-176,(2011).

154 Marcovitz, A. & Levy, Y. Frustration in protein-DNA binding influences conformational switching and target search kinetics. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 17957-17962,(2011).

155 Murugan, R. Theory of site-specific DNA-protein interactions in the presence of conformational fluctuations of DNA binding domains. *Biophysical Journal* **99**, 353-359,(2010).

156 Zhou, H. X. Rapid search for specific sites on DNA through conformational switch of nonspecifically bound proteins. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 8651-8656,(2011).

157 Lewin, B. *Genes VII*. (Oxford Univ Press, 2000).

158 Pernstich, C. & Halford, S. E. Illuminating the reaction pathway of the FokI restriction endonuclease by fluorescence resonance energy transfer. *Nucleic Acids Research* **40**, 1203-1213,(2012).

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

160 Lotze, M. T. & Tracey, K. J. High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. *Nature Reviews Immunology* **5**, 331-342,(2005).

161 Tang, D., Kang, R., Zeh, H. J., 3rd & Lotze, M. T. High-mobility group box 1, oxidative stress, and disease. *Antioxidants & Redox Signaling* **14**, 1315-1335,(2011).

162 Venereau, E., Schiraldi, M., Uguccioni, M. & Bianchi, M. E. HMGB1 and leukocyte migration during trauma and sterile inflammation. *Molecular Immunology* (2012).

163 Orlova, V. V., Choi, E. Y., Xie, C., Chavakis, E., Bierhaus, A., Ihanus, E., Ballantyne, C. M., Gahmberg, C. G., Bianchi, M. E. *et al.* A novel pathway of HMGB1-mediated inflammatory cell recruitment that requires Mac-1-integrin. *EMBO Journal* **26**, 1129-1139,(2007).

164 Palumbo, R., Galvez, B. G., Pusterla, T., De Marchis, F., Cossu, G., Marcu, K. B. & Bianchi, M. E. Cells migrating to sites of tissue damage in response to the danger signal HMGB1 require NF-kappaB activation. *Journal of Cell Biology* **179**, 33-40,(2007).

165   Riuzzi, F., Sorci, G. & Donato, R. The amphoterin (HMGB1)/receptor for advanced glycation end products (RAGE) pair modulates myoblast proliferation, apoptosis, adhesiveness, migration, and invasiveness. Functional inactivation of RAGE in L6 myoblasts results in tumor formation in vivo. *Journal of Biological Chemistry* **281**, 8242-8253,(2006).

166   Taguchi, A., Blood, D. C., del Toro, G., Canet, A., Lee, D. C., Qu, W., Tanji, N., Lu, Y., Lalla, E. *et al.* Blockade of RAGE-amphoterin signalling suppresses tumour growth and metastases. *Nature* **405**, 354-360,(2000).

167   Rubartelli, A. & Lotze, M. T. Inside, outside, upside down: damage-associated molecular-pattern molecules (DAMPs) and redox. *Trends in Immunology* **28**, 429-436,(2007).

168   Sahu, D., Debnath, P., Takayama, Y. & Iwahara, J. Redox properties of the A-domain of the HMGB1 protein. *FEBS Letters* **582**, 3973-3978,(2008).

169   Kazama, H., Ricci, J. E., Herndon, J. M., Hoppe, G., Green, D. R. & Ferguson, T. A. Induction of immunological tolerance by apoptotic cells requires caspase-dependent oxidation of high-mobility group box-1 protein. *Immunity* **29**, 21-32,(2008).

170   Tang, D., Billiar, T. A. & Lotze, M. T. A Janus Tale of Two Active HMGB1 Redox States. *Molecular Medicine*,(2012).

171   Venereau, E., Casalgrandi, M., Schiraldi, M., Antoine, D. J., Cattaneo, A., De Marchis, F., Liu, J., Antonelli, A., Preti, A. *et al.* Mutually exclusive redox forms of HMGB1 promote cell recruitment or proinflammatory cytokine release. *Journal of Experimental Medicine* **209**, 1519-1528,(2012).

172   Liu, A., Fang, H., Dirsch, O., Jin, H. & Dahmen, U. Oxidation of HMGB1 causes attenuation of its pro-inflammatory activity and occurs during liver ischemia and reperfusion. *PLoS One* **7**, e35379,(2012).

173   Yang, H., Lundback, P., Ottosson, L., Erlandsson-Harris, H., Venereau, E., Bianchi, M. E., Al-Abed, Y., Andersson, U., Tracey, K. J. *et al.* Redox modification of cysteine residues regulates the cytokine activity of HMGB1. *Molecular Medicine* **18**, 250-259,(2012).

174   Kozlowski, J. M., Fidler, I. J., Campbell, D., Xu, Z. L., Kaighn, M. E. & Hart, I. R. Metastatic behavior of human tumor cell lines grown in the nude mouse. *Cancer Research* **44**, 3522-3529,(1984).

175   Schanda, P., Kupce, E. & Brutscher, B. SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *Journal of Biomolecular NMR* **33**, 199-211,(2005).

176   Pervushin, K., Riek, R., Wider, G. & Wuthrich, K. Attenuated T2 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 of the United States of America* **94**, 12366-12371,(1997).

177   Clore, G. M. & Gronenborn, A. M. Determining the structures of large proteins and protein complexes by NMR. *Trends in Biotechnology* **16**, 22-34,(1998).

178   Knapp, S., Muller, S., Digilio, G., Bonaldi, T., Bianchi, M. E. & Musco, G. The long acidic tail of high mobility group box 1 (HMGB1) protein forms an extended and flexible structure that interacts with specific residues within and between the HMG boxes. *Biochemistry* **43**, 11992-11997,(2004).

179   Stott, K., Watson, M., Howe, F. S., Grossmann, J. G. & Thomas, J. O. Tail-mediated collapse of HMGB1 is dynamic and occurs via differential binding of the acidic tail to the A and B domains. *Journal of Molecular Biology* **403**, 706-722,(2010).

180   Watson, M., Stott, K. & Thomas, J. O. Mapping intramolecular interactions between domains in HMGB1 using a tail-truncation approach. *Journal of Molecular Biology* **374**, 1286-1297,(2007).

181   Ek, M., Popovic, K., Harris, H. E., Naucler, C. S. & Wahren-Herlenius, M. Increased extracellular levels of the novel proinflammatory cytokine high mobility group box chromosomal protein 1 in minor salivary glands of patients with Sjogren's syndrome. *Arthritis and Rheumatism* **54**, 2289-2294,(2006).

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

183   Logsdon, C. D., Fuentes, M. K., Huang, E. H. & Arumugam, T. RAGE and RAGE ligands in cancer. *Current Molecular Medicine* **7**, 777-789,(2007).

184   Lotze, M. T. & DeMarco, R. A. Dealing with death: HMGB1 as a novel target for cancer therapy. *Current opinion in investigational drugs* **4**, 1405-1409,(2003).

185   Tang, D., Kang, R., Zeh, H. J., 3rd & Lotze, M. T. High-mobility group box 1 and cancer. *Biochimica et Biophysica Acta* **1799**, 131-140,(2010).

186   Ulloa, L. & Messmer, D. High-mobility group box 1 (HMGB1) protein: friend and foe. *Cytokine and Growth Factor Reviews* **17**, 189-201,(2006).

187   Wake, H., Mori, S., Liu, K., Takahashi, H. K. & Nishibori, M. Histidine-rich glycoprotein inhibited high mobility group box 1 in complex with heparin-induced angiogenesis in matrigel plug assay. *European Journal of Pharmacology* **623**, 89-95,(2009).

188   Wake, H., Mori, S., Liu, K., Takahashi, H. K. & Nishibori, M. High mobility group box 1 complexed with heparin induced angiogenesis in a matrigel plug assay. *Acta Medica Okayama* **63**, 249-262,(2009).

189   Mollica, L., Curioni, A., Andreoni, W., Bianchi, M. E. & Musco, G. The binding domain of the HMGB1 inhibitor carbenoxolone: Theory and experiment. *Chemical Physics Letters* **456**, 236-242,(2008).

190   Mollica, L., De Marchis, F., Spitaleri, A., Dallacosta, C., Pennacchini, D., Zamai, M., Agresti, A., Trisciuoglio, L., Musco, G. *et al.* Glycyrrhizin binds to high-mobility group box 1 protein and inhibits its cytokine activities. *Chemistry and Biology* **14**, 431-441,(2007).

VITA:

Levani Zandarashvili was born on March 10, 1988, in Tbilisi, Georgia, to Tamazi Zandarashvili and Mariami Shamugia. After graduation from specialized physical-mathematical school in Tbilisi, Georgia, with several national awards in high-school physics tournaments, Levani applied to the Tbilisi State University, Georgia, department of physics. During his undergraduate studies, Levani also worked as a laboratory assistant at "EMCoS". Right after graduating with B.S. in physics in 2008, he joined laboratory of Dr. James C. Lee at the University of Texas Medical Branch (UTMB) at Galveston as a research scholar, where for one year he learned different biochemical and biophysical techniques and investigated properties of wtCRP protein. In 2009, he enrolled in the Molecular Biophysics Educational Track (MBET) at UTMB. After the first year of rotations, Levani joined laboratory of Dr. Junji Iwahara. During following years he worked on several different projects. The main focus of his studies was elucidation of relationship between protein dynamics and its functional kinetics. During his studies at UTMB, Levani was awarded with numerous local and national awards. In 2010 he was awarded with ISOTEC Student Sponsorship for 51st Experimental Nuclear Magnetic Resonance Conference (ENC). In 2011 he received Student travel award for 52nd ENC Conference, Barbara Bowman Memorial Award for Research Excellence (UTMB), Edith and Robert Zinn Presidential Scholarship (UTMB) and the poster award at GCC MR conference. In 2012, Protein Science Young Investigator Travel Grant / Protein Society Finn Wold Travel Award, Poster award at the Protein Society Meeting, Irma Mendoza Scholarship (UTMB), Marianne Blum, Ph.D. Endowed Scholarship (UTMB) and James E. Beall II Memorial Scholarship (UTMB). Also, Levani has presented poster at the Annual Structural Biology Symposium, Galveston, Texas in 2010, 2011, 2012 and 2013, at ENC conference in 2011 and 2012, as well as at the Protein Society meeting in 2012.

During his PhD studies, Levani published three papers in reputable journals and one book chapter (see below). In the fall of 2010 and 2011, Levani was a teaching assistant for the "Molecular Biophysics I" course. For different periods of his studies, Levani also served as MBET Student Representative to the Graduate Faculty and as the Treasurer of Biological Chemistry Student Organization. He was also the member of the Structural Biology Symposium organizer committee in 2012, as well as 2013. In the summer 2013 Levani was awarded Kempner Postdoctoral Fellowship and following completion of his dissertation he will continue his scientific work with his mentor, Dr. J. Iwahara at UTMB.

Publications

Levani Zandarashvili, Debashish Sahu, Kwanbok Lee, Yong Sun Lee, Pomila Singh, Krishna Rajarathnam and Junji Iwahara (2013) Oxidation in Extracellular Fluids Studied by in Situ Protein NMR Spectroscopy. 2013; VOL. 288, NO. 17, pp. 11621–11627. *Journal of Biological Chemistry*

Levani Zandarashvili, Alexandre Esadze, and Junji Iwahara (2013) NMR studies on the dynamics of hydrogen bonds and ion pairs involving lysine side chains of proteins. *Adv Protein Chem Struct Biol* 93, in press.

Levani Zandarashvili, Dana Vuzman, Alexandre Esadze, Yuki Takayama, Debashish Sahu, Yaakov Levy, and Junji Iwahara (2012) Asymmetrical roles of zinc fingers in dynamic DNA-scanning process by the inducible transcription factor Egr-1. *Proceedings of the National Academy of Sciences*, 109(26):E1724-32

Levani Zandarashvili, Da-Wei Li, Tianzhi Wang, Rafael Brüschweiler, and Junji Iwahara (2011) Signature of Mobile *Hydrogen* Bonding of Lysine Side Chains from Long-Range ^{15}N-^{13}C scalar J-Couplings and Computation. *Journal of the American Chemical Society* 133, 9192-9195.

<u>Permanent address:</u> 215 Mechanic Str., Apt L209, Galveston, TX, 77550, USA

This dissertation was typed by Levani Zandarashvili.