

Copyright
by
Dan Nguyen
2018

**The Dissertation Committee for Dan Nguyen Certifies that this is the approved
version of the following dissertation:**

**Biophysical Analyses of Intermolecular Ion-Pairs in Protein-DNA
Complexes**

Committee:

Junji Iwahara, PhD, Supervisor

Marc C. Morais, PhD, Chair

Kay Choi, PhD

Whitney Yin, PhD

Chin-Yo Lin, PhD

Dean, Graduate School

**Biophysical Analyses of Intermolecular Ion-Pairs in Protein-DNA
Complexes**

by

Dan Nguyen, 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

June, 2018

Dedication

To Dad, my best friend, and Mom, the strongest woman I know. Lastly, this is for Ông Ngoại and Bà Nội. Thank you for looking out from above. Rest easy.

Acknowledgements

I would like to thank my mentor, Dr. Junji Iwahara, for constantly challenging me to become a better scientist. He is a model scientist with a tremendous amount of knowledge and obvious passion for his research work.

I would also like to thank my supervisory committee members Dr. Marc C. Morais, Dr. Kay Choi, Dr. Whitney Yin, and Dr. Chin-Yo Lin for their support and guidance in my PhD work.

I would like to thank my previous lab members Drs. Alex Esadze and Levani Zandarashvili, who have always been just an e-mail or phone call away when I needed some friendly advice and guidance. I would also like to thank my colleague Dr. Catherine Kemme, and my current lab members Esmeralda Martinez and Ross Luu, for making the long days seem just a bit shorter than they really were.

A special thank you to Dr. James Lee who first challenged my grit during my graduate school interview, and then spent countless hours guiding a very challenging student.

Lastly, I would like to thank God, my parents, my siblings, my family and my friends. The best thing about coming home from a rough week is knowing you are not alone.

**Biophysical Analyses of Intermolecular Ion-Pairs in Protein-DNA
Complexes**

Publication No. _____

Dan Nguyen, PhD

The University of Texas Medical Branch, 2018

Supervisor: Junji Iwahara

Abstract

For biological processes, one of the most fundamental interactions is the formation of the ion pair. In protein-DNA complexes, ion pairs form between basic side chains, such as lysine and arginine, and the phosphate backbone of DNA. Moreover, ion pairs are found in dynamic equilibrium between two major states. While great strides have been made to improve the understanding of ion pair in protein-DNA interactions, much remains to be understood on the role of ion pair dynamics in protein-DNA interactions. Thus, in this work we investigate the role of ion pair dynamics in protein-DNA interactions.

First, we investigate side chain dynamics of the Antennapedia homeodomain (Antp) for both lysine and arginine side chains. Using NMR spectroscopy, we determine the mobility of lysine and arginine side chains in the free and DNA-bound state of Antp through ^{15}N relaxation experiments. We show that side chain mobility is retained for both Arg and Lys side chains that form an ion pair with the phosphate backbone.

Second, we investigate the role of ion pair dynamics of the Antp homeodomain through the oxygen-to-sulfur substitution (dithioation) of a single phosphate that forms an ion pair with a lysine side chain. We show an enhancement in affinity for the dithioated DNA compared to the unmodified DNA by the Antp homeodomain using fluorescence assays. Using ITC, we show that the enhancement in affinity is due to an entropic contribution. Through NMR spectroscopy, we show that the lysine side chain interacting with the dithioated phosphate has a higher mobility compared to the unmodified phosphate and confirm these observations using X-ray crystallography. We further investigate the stereospecific effects of this modification.

Lastly, we establish a system to study kinetic parameters of the Antp homeodomain using Ficoll, as a macromolecular crowder, to investigate the role of ion-pair dynamics in DNA target search. Using fluorescence-based stopped flow kinetics experiments, we show an increase in sliding length for the Antp homeodomain in the presence of Ficoll. Furthermore, through NMR spectroscopy-based diffusion experiments, we show that 3-D diffusion is slower for both the Antp homeodomain and DNA duplex in the presence of Ficoll.

TABLE OF CONTENTS

List of Tables	xii
List of Figures	xiv
List of Abbreviations	xxv
iii	
CHAPTER 1.....	29
Introduction.....	29
1.1 Background	29
1.1.1 Ion pairs of biomolecules.....	29
1.1.2 Importance of ion pairs in DNA recognition by proteins.....	30
1.1.3 The contact ion pair (CIP) and the solvent-separated ion pair (SIP)	31
1.1.4 Potentials of Mean Force and the CIP-SIP transitions	32
1.1.5 Thermodynamics of ion pairs	34
1.1.6 Thermodynamics of protein-DNA association	35
1.1.7 Role of conformational entropy in molecular association	36
1.2 Goals and questions to be addressed.....	39
1.2.1 How do side chain dynamics differ in the free protein and the DNA- bound state?	39
1.2.2 What role do ion-pair dynamics play in enhancing protein-DNA association?	40
1.2.3 What is the role of ion-pairs in protein target search kinetics?.....	41
1.3 Overview of methods and model system	43
1.3.1 Model system for research.....	43
1.3.2 Phosphorothioate: the oxygen-to-sulfur substitution	45
1.3.3 NMR spectroscopy methods to determine side-chain mobility of ion pairs	47

CHAPTER 2.....	50
Internal Motions of Basic Side Chains of the Antennapedia Homeodomain in the Free and DNA-Bound States	50
2.1 Background	50
2.2 Materials and Methods.....	53
2.2.1 Protein and DNA.....	53
2.2.2 NMR relaxation analysis for Lys and Arg side chains.....	54
2.2.3 Determination of order parameters for Lys and Arg side chains...	60
2.3 Results.....	60
2.3.1 Internal motions of Arg side-chain and N_ϵ - H_ϵ groups.....	60
2.3.2 Change in mobility of Arg side chains upon protein-DNA association.....	64
2.3.3 Internal motions of Lys side-chain NH_3^+ groups.....	66
2.3.4 Different motional characteristics of Arg and Lys side chains.....	70
2.3.5 Retained mobility of basic side chains forming ion pairs with DNA phosphates.....	72
2.4 Discussion	73
CHAPTER 3.....	74
Using the Sulfur Substitution in Ion Pairs to Improve Understanding of Ion Pair Dynamics	74
3.1 Introduction	74
3.2 Materials and Methods.....	77
3.2.1 Protein Preparation.....	77
3.2.2 DNA Preparation.....	78
3.2.3 Preparation of R_p and S_p diastereomers of monothioated DNA	78
3.2.4 Fluorescence-based affinity measurements	79
3.2.5 ITC measurements.....	81
3.2.6 NMR Spectroscopy for the unmodified and dithioated protein-DNA complexes	81
3.2.7 NMR Spectroscopy for the Protein-DNA complexes of the R_p and S_p diastereomers of monothioate.....	83
3.2.8 X-ray Crystallography for the unmodified and dithioated protein-DNA complexes.....	84

3.2.9 X-ray Crystallography for the protein-DNA complexes of the R_p and S_p diastereomers of monothioate.....	85
3.3 Results.....	86
3.3.1 Isolation of R_p and S_p diastereomers of monothioated DNA	88
3.3.2 Affinity enhancement by dithioation of DNA phosphate.....	90
3.3.3 Thermodynamic impact of phosphorodithioate on association	91
3.3.4 Impact of R_p and S_p monothioation on protein-DNA binding affinity	97
3.3.5 NMR of interfacial Lys NH_3^+ groups	99
3.3.6 NMR evidence for ionic hydrogen bonds.....	99
3.3.7 Mobility of interfacial Lys NH_3^+ groups	102
3.3.8 Impact of R_p and S_p monothioation on ion-pair dynamics	105
3.3.9 Crystal structures of the complexes	109
3.3.10 Lys57 – DNA interactions in the crystal structures	112
3.3.11 Crystal structures of the R_p and S_p monothioated DNA complexes	113
3.4 Discussion	116
3.4.1 Mobilization of ion pair by oxygen-to-sulfur substitution in DNA phosphate	116
3.4.2 Entropic gain due to mobilization of the intermolecular ion pair ..	117
3.4.3 Role of ion-pair dynamics in protein-DNA interactions	118
3.4.4 Entropic gain due to mobilization of the intermolecular ion pair ..	119
3.4.5 Implications of transcription factor decoys for therapeutic applications	120
CHAPTER 4.....	122
Towards Elucidation of the Relationship Between Ion-Pair Dynamics and Molecular Kinetics	122
4.1 Introduction	122
4.2 Materials and Methods.....	125
4.2.1 Preparation of the Antp Homeodomain protein.....	125
4.2.2 Preparation of fluorescence-labeled probe DNA	125
4.2.3 Competitor DNA.....	126
4.2.4 Stopped-flow fluorescence-based target association kinetics experiments	126

4.2.5 NMR sample preparation.....	128
4.2.6 NMR Spectroscopy	129
4.3 Results.....	130
4.3.1 Target association kinetics and determination of sliding length....	131
4.3.2 Impact of Ficoll on the Antp Homeodomain sliding length	133
4.3.3 Impact of Ficoll on translational diffusion	133
4.3.4 Impact of Ficoll on rotational diffusion.....	136
4.4 Discussion	138
4.5 Future Directions	141
4.5.1 Further delineating the role of ion pair dynamics on Protein-DNA search kinetics.....	141
CHAPTER 5.....	143
Conclusions.....	143
5.1 Major Findings in this Work	143
5.2 Future Perspectives	146
5.2.1 Implications on transcription factor decoys for therapeutic applications	146
5.3 Final Remarks: Potential Impact of this Work.....	146
Appendix.....	150
Bibliography	153
VITA 167	
PUBLICATIONS.....	168

List of Tables

Table 2.1:	^{15}N NMR relaxation data for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups of the Antp homeodomain in the free state at pH 5.8 and 25°C.....	62
Table 2.2:	^{15}N NMR relaxation data for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups of the Antp homeodomain in the specific complex with 15-bp DNA at pH 5.8 and 25°C.....	62
Table 2.3:	Molecular rotational diffusion parameters for the Antp homeodomain in the free and DNA-bound states. ^{a)}	63
Table 2.4:	NMR order parameters S^2_{axis} determined for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups of the Antp homeodomain in the free and DNA-bound states at pH 5.8 and 25°C.....	63
Table 2.5:	^{15}N NMR relaxation data for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups of the Antp homeodomain in the free state at pH 4.5 and 2°C.	69
Table 2.6:	^{15}N NMR relaxation data for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups of the Antp homeodomain in the specific complex with 15-bp target DNA at pH 5.8 and 2°C.	69
Table 2.7:	NMR order parameters S^2_{axis} determined for Lys NH_3^+ groups of the Antp homeodomain in the free and DNA-bound states. ^{a)}	69
Table 3.1:	Thermodynamic parameters on protein-DNA association measured for the Antp homeodomain-DNA complexes at 25°C. ^{a)}	96

Table 3.2:	^{15}N relaxation and dynamics parameters measured for the Lys side-chain NH_3^+ groups of the Antp homeodomain-DNA complexes at 25°C and pH 5.8. ^{a)}	104
Table 3.3:	NMR data for Lys side-chain NH_3^+ groups in the complexes of the Antp homeodomain with R_P and S_P -monothioated DNA. ^{a)}	108
Table 3.4:	Crystallographic data collection and refinement statistics.	111
Table 3.5:	Crystallographic data collection and refinement statistics for R_p and S_p -monothioated DNA complexes.....	115

List of Figures

Figure 1.1: Illustration of the Contact ion pair (left) as opposed to the Solvent-separated ion pair (right).	31
Figure 1.2: Example of PMFs for intermolecular ion pairs of the Antp homeodomain-DNA complex and Egr-1 zinc finger-DNA complex. .	33
Figure 1.3: Determination of S^2 order parameters from observation of internal motions of Lys NH_3^+ , and observation of CIP by Hydrogen-bond scalar J-coupling.	38
Figure 1.4: Illustration of a transcription factor breaking all of its CIPs with the DNA in order to translocate.	42
Figure 1.5: Crystal structure of the Antp Homeodomain-DNA complex (PDB 4XID).	45
Figure 1.6: Chemical composition of a typical DNA phosphate (left) and a dithioated phosphate with the oxygen-to-sulfur substitution (right)....	46
Figure 1.7: Schematic for determination of S^2 order parameters from various NMR experiments.....	49
Figure 2.1: The basic side chains in the Antp homeodomain-DNA complex. (a) Sequences of the Antp homeodomain C39S mutant and the 15-bp DNA duplex used in this study. (b) A crystal structure of the Antp homeodomain-DNA complex (PDB 4XID).	52

Figure 2.2: Pulse sequences for the ^{15}N relaxation measurement for lysine side-chain NH_3^+ groups. The key elements in the current work are indicated in red. Thin and bold bars in black represent hard rectangular 90° and 180° pulses, respectively. Water-selective half-Gaussian (2.1 ms) and soft-rectangular (1.2 ms) 90° pulses are represented by half-bell and short-bold shapes, respectively. Unless indicated otherwise, pulse phases are along x , and the carrier position for ^1H was set to the position of the water resonance. The RF strength for hard rectangular ^1H and ^{15}N pulses were 40 kHz and 6.6 kHz, respectively. The RF strengths of ^1H WALTZ-16 decoupling⁸⁵ during the ^{15}N evolution period was 4.2 kHz. ^{15}N carrier position was set to 33.1 ppm. ^{15}N -decoupling during the t_2 detection period was performed with WALTZ-16 (RF strength, 1.0 kHz). A gray bell-shape for ^{15}N represents an r-SNOB⁸⁰ 180° pulse (1.0 ms) for selective inversion and refocusing of Lys side-chain $^{15}\text{N}_\zeta$ nuclei. The delays τ_a and τ_b were 2.7 ms and 1.3 ms, respectively. Quadrature detection in the t_1 domain was achieved using States-TPPI, incrementing the phase φ_1 . Pulsed field gradients (PFGs) were optimized to minimize the water signal. **(a)** ^{15}N R_1 measurement. Although it is not essential owing to negligible CSA-DD cross correlation for NH_3^+ , a ^1H 180° pulse, which does not affect H_2O resonance, was applied every 10 ms during the delay T for longitudinal relaxation. Phase cycles: $\varphi_1 = (2y, 2(-y))$, $\varphi_2 = (y, -y)$, $\varphi_3 = (4x, 4(-x))$, $\varphi_4 = (8y, 8(-y))$, and receiver = $(x, -x, -x, x, x, 2(-x, x, x, -x), x, -x, -x, x)$. **(b)** ^{15}N $R_{2,ini}$ measurement. The RF strength for ^{15}N pulses for the CPMG scheme was 5.4 kHz. ^1H carrier position was shifted to 7.8 ppm right after the PFG g_4 and set back to the position of water resonance right after the PFG g_5 . The RF strength $\omega_{CW}/2\pi$ of ^1H CW during the CPMG was set to 4.3 kHz, which was

adjusted to satisfy $\omega_{CW}/2\pi = 2k_{CPMG}$ (k , integer).⁸¹ The delays ξ_1 and ξ_2 are for alignment of ^1H magnetization and given by $\xi_1 = 1/\omega_{CW} - (4/\pi)\tau_{90H}$ and $\xi_2 = \tau_{90N} - (2/\pi)\tau_{90H}$,^{81, 86} in which τ_{90} represents a length of a relevant 90° pulse. Phase cycles: $\varphi_1 = (4y, 4(-y))$, $\varphi_2 = (8y, 8(-y))$, $\varphi_3 = x$, $\varphi_4 = (x, -x)$, $\varphi_5 = (2y, 2(-y))$, $\varphi_6 = (2x, 2(-x))$, $\varphi_7 = (2(-y), 2y)$, and receiver = $(x, -x, x, -x, 2(-x, x, -x, x), x, -x, x, -x)$. **(c)** Heteronuclear ^1H - ^{15}N NOE measurement. Measurement with ^1H saturation (5 s) was performed with a train of $180^\circ x$ and $180^\circ(-x)$ pulses (RF strength, 11 kHz) with an interval of 10 ms. ^1H carrier position was at 7.8 ppm during the ^1H saturation period. The reference spectrum was measured without the scheme in the bracket. The recycle delay (including the saturation period) was set to 18 s for a 750-MHz spectrometer. Phase cycles: $\varphi_1 = (y, -y)$, $\varphi_2 = (4x, 4y, 4(-x), 4(-y))$, $\varphi_3 = (2x, 2(-x))$, and receiver = $(x, -x, -x, x, -x, x, x, -x)$. **(d)** Efficiency in coherence transfers as a function of the delay τ_b calculated using Eqs. 1 and 2 with $|^1J_{NH}| = 74$ Hz and ^1H 180° pulse length of 20 μs . Results for the N_y and $4N_y H_z H_z$ terms are shown in solid and dotted lines, respectively. Red and green arrows indicate the values of the delay τ_b in the current and previous pulses sequences, respectively. Due to the different spin systems and ^{15}N chemical shift ranges, the experiments for Lys and Arg side chains should be conducted separately. 59

Figure 2.3: NMR investigations on internal motions of Arg side-chain N_ϵ - H_ϵ moieties of the Antp homeodomain in the free and DNA-bound states. **(a)** Overlaid ^1H - ^{15}N HISQC spectra⁶⁴ recorded for the Arg side chains in the free protein (blue) and in the complex with 15-bp DNA (red). The region indicated by a dotted box is expanded in the inset. **(b)** NMR-derived order parameters S^2 for Arg N_ϵ - H_ϵ bonds in the free protein (blue) and in the complex (red). 64

Figure 2.4: NMR investigations on internal motions of Lys side-chain NH_3^+ moieties of the Antp homeodomain in the free and DNA-bound states. **(a)** Lys NH_3^+ -selective ^1H - ^{15}N HISQC spectra⁶⁴ recorded for the free protein (blue) and the complex with 15-bp DNA (red). **(b)** NMR-derived order parameters S^2_{axis} for Lys side chains in the free protein (blue) and in the complex (red). **(c)** Dynamic equilibrium between the CIP and SIP states. 67

Figure 2.5: Data indicating that the Antp homeodomain at pH 4.5 and 2°C remains folded. Shown here are ^1H - ^{15}N TROSY spectra recorded for the backbone NH groups of the free Antp homeodomain at pH 5.8 and 25°C (left) and at pH 4.5 and 2°C (right). 68

Figure 2.6: Changes in order parameter and entropy for the Arg/Lys cationic groups of the Antp homeodomain upon binding to DNA. **(a)** Difference between the order parameters S^2 for the free and DNA-bound states. $\Delta S^2 = S^2(\text{complex}) - S^2(\text{free})$. Individual values of the S^2 parameters are reported in Tables S4 and S7. **(b)** Entropic changes due to changes in mobility upon the protein-DNA association. The values were determined from $S^2(\text{complex})$ and $S^2(\text{free})$ using Eq. 24 in Yang and Kay.⁹⁵ The overall entropic change for the Arg/Lys cationic groups was calculated to be -48 ± 5 J/mol/K. 71

Figure 3.1: Enhancement of binding affinity for the Antp homeodomain by dithioation of the DNA phosphate at the Lys57 interaction site. **(a)** Dithioation of DNA phosphate. **(b)** Binding isotherm as measured by the fluorescence anisotropy-based titration experiment. Fluorescence anisotropy for TAMRA-labeled 15-bp duplexes (red, unmodified DNA; blue, dithioated DNA) was measured as a function of the concentration of the Antp homeodomain. The probe DNA concentration was 3.3 nM in this experiment. Dissociation constants determined from the three replicates together with Eq. 1 were $K_d = 11 \pm 1$ nM for the unmodified DNA and $K_d = 2.9 \pm 0.2$ nM for the dithioated DNA **(c)** Competition assay data used to indirectly determine K_d constants. The total concentrations of the probe DNA and the protein were 10 and 50 nM, respectively in this experiment. The measurement was replicated three times. The fitting calculations using Eqs 2 and 3 together with the K_d constant for the competitor gave $K_d = 15.3 \pm 0.8$ nM for the unmodified DNA and $K_d = 4.0 \pm 0.5$ nM for the dithioated DNA. The solid lines represent best-fit curves. 87

Figure 3.2: Oxygen-to-sulfur substitution in the DNA phosphate at the Lys57 interaction site in the Antp homeodomain-DNA complex. (A) Two diastereomers (Rp and Sp) of monothioate were used in the current study, whereas a phosphorodithioate was used in our previous study. In Rp and Sp monothioations, a sulfur atom substitutes for OP2 and OP1, respectively. (B) Mass spectrometry data for the monothioated DNA strand. (C) Chromatogram in Mono-Q anion-exchange chromatography for 15-bp DNA duplexes containing a phosphoromonothioate at the Lys57 site. A diastereomeric mixture of the phosphoromonothioate of the DNA duplex gives two major peaks corresponding to the Rp and Sp diastereomers, which were confirmed by crystallography (see Figure 3.10). 89

Figure 3.3: Competition assay data used to assess unlabeled specific DNA (a) and nonspecific DNA (b) duplexes. The total concentrations of the probe DNA and the Antp homeodomain were 10 and 50 nM, respectively. The fitting calculations using Eqs 2 and 3 together with the K_d constant for the probe gave $K_d = 10 \pm 2$ nM for the specific DNA and $K_d = (7.9 \pm 0.8) \times 10^3$ nM for the nonspecific DNA..... 92

Figure 3.4: Thermodynamic impact of dithioation of DNA phosphate at the Lys57 interaction site. **(a)** ITC data for the specific DNA duplexes with 21 injections of the Antp homeodomain, leading to the final molar ratio of 5.3. The DNA concentration was 9 μM . Because the sample concentrations were comparable to the K_d for nonspecific association and the DNA duplexes are long enough to nonspecifically interact with additional protein molecules, heat from the nonspecific interactions was also observed after the target site was saturated. **(b)** ITC data for the nonspecific DNA duplex with 21 injections of the Antp homeodomain, leading to the final molar ratio of 6.6. The DNA concentration was 8 μM DNA. The heat effect of the first and second nonspecific associations were opposite in sign. **(c)** ITC data for the specific DNA duplexes with 21 injections of the Antp homeodomain leading to the final molar ratio of 1.8. The DNA concentration was 17 μM . Green dotted lines correspond to the heat effect arising from association of the Antp homeodomain with the target site on the DNA duplexes. **(d)** Changes in binding free energy ($\Delta\Delta G$), enthalpy ($\Delta\Delta H$) and entropic terms ($-T\Delta\Delta S$) upon the oxygen-to-sulfur substitution of the DNA phosphate. The increase in binding entropy $\Delta\Delta S$ was determined to be $2.5 \pm 0.5 \text{ cal K}^{-1} \text{ mol}^{-1}$. See also Table 1. 95

Figure 3.5: Impact of oxygen-to-sulfur substitution at Lys57 interaction site on binding affinity for the Antp homeodomain-DNA complex. (A) Competition assay data used to determine K_d constants for R_P and S_P -monothioated DNA duplexes. The total concentrations of the probe DNA and the protein were 10 and 50 nM, respectively in this experiment. The measurement was repeated four times. The solid lines represent best-fit curves. (B) Change in the binding free energy upon modifications of the DNA phosphate group at the Lys57 interaction site. In this case, $\Delta\Delta G$ is defined as ΔG (unmodified) – ΔG (modified). ... 98

Figure 3.6: NMR evidence for the hydrogen bonds between the Lys NH₃⁺ and DNA phosphate / phosphorodithioate groups. (a) Overlaid HISQC spectra recorded for the Lys NH₃⁺ groups in the unmodified (red) and dithioated (blue) DNA complexes of the Antp homeodomain. (b) The H3(N)P correlation spectrum recorded for the dithioated DNA complex. This spectrum shows 1H-31P cross peaks arising from coherence transfers via hydrogen-bond scalar coupling h3JNP between Lys 15N and DNA 31P nuclei. The signals represent the direct evidence for ionic hydrogen bonds with DNA phosphate (31P ~-2 ppm) or phosphorodithioate (~108 ppm)..... 101

Figure 3.7: Mobilization of the Lys57 NH_3^+ group by dithioation of the interacting DNA phosphate group. (a) Lys NH_3^+ ^{15}N R_1 relaxation (^1H frequency, 600 MHz) for the unmodified (red) and dithioated (blue) DNA complexes. (b) Lys NH_3^+ ^{15}N R_2 relaxation (^1H frequency, 800 MHz) for the unmodified (red) and dithioated (blue) DNA complexes. (c) Changes in order parameters S^2_{axis} for the symmetry axis of Lys NH_3^+ groups upon dithioation of the DNA phosphate. 103

Figure 3.8: Impact of the oxygen-to-sulfur substitutions on the intermolecular ion pair with Lys57. (A) Overlaid Lys side-chain NH_3^+ -selective ^1H - ^{15}N HISQC spectra⁶⁴ recorded for the four complexes of the Antp homeodomain and DNA. These complexes chemically differ only in the DNA phosphate at the Lys57 interaction site (black, unmodified; blue, R_P -monothioated; red, S_P -monothioated; and magenta, dithioated). (B) Change in NMR-derived order parameters S^2_{axis} for Lys NH_3^+ groups upon oxygen-to-sulfur substitution in DNA phosphate at the Lys57 interaction site. These data show that sulfur substitution of the proximal phosphate oxygen (i.e., O_{P2}) atom significantly mobilizes the Lys side-chain NH_3^+ group. 106

Figure 3.9: Structural impact of dithioation of the DNA phosphate at Antp Lys57 interaction site. (a) Superposition of the crystal structures of the Antp homeodomain-DNA complexes with (blue) and without (red) dithioation. The phosphate and phosphorodithioate groups at the modification site are indicated by an arrow. (b) Major and minor groove widths of the DNA double helices of the complexes. Solid and dotted lines show the data for Structures 1 and 2, respectively, which are two independent structures in the asymmetry unit (see the main text). The values were calculated with the CURVES+ program (44). Horizontal arrows indicate values for the canonical B-form. (c, d) Electron density maps together with the structures of the Lys57 side chain – DNA phosphate (Panel c) / phosphorodithioate (Panel d) ion pairs. Two independent structures in the asymmetric unit are shown for each complex. (e) DNA backbone torsion angles relevant to the phosphorodithioate and phosphate groups at the Lys57 interaction site.

..... 110

Figure 3.10: Crystal structures of the Antp homeodomain-DNA complexes containing R_P phosphoromonothioate (Panel A) or S_P phosphoromonothioate (Panel B) at the Lys57 interaction site. The crystal structures of the R_P and S_P -monothioated DNA complexes were determined at 2.09 and 2.75 Å resolutions, respectively. For each panel, superposition of the crystal structures of the modified (green) and unmodified (purple) complexes is shown. The modification sites are shown in ball-stick representation. (indicated by arrows). For each modified complex, the electrostatic density map, of the phosphoromonothioate at Lys57 interaction site, for two structures in the asymmetric unit is shown. For the R_P -monothioated DNA complex, the tip (i.e., N ζ , C ϵ , and C δ groups) of the Lys57 side chain was not resolved in the crystallographic electron density map. 114

Figure 3.11: Possible reason for the mobilization of Lys57 side chain due to oxygen-to-sulfur substitution in DNA phosphate forming an intermolecular ion pair. 117

Figure 4.1: Experimental design of the using stopped-flow fluorescence kinetic assays. (A) Stopped-flow fluorescence assay for target search kinetics in the absence of Ficoll. (B) Stopped-flow fluorescence assay for target search kinetics in the presence of Ficoll (75 mg/ml). 124

- Figure 4.2: DNA duplexes used for kinetic studies on the Antp homeodomain. **(A)** Probe and competitor DNA duplexes used for the current study. The sequence of nucleotides in green represent the target sequence of the Antp Homeodomain. **(B)** PAGE of the purified duplexes, which were visualized by fluorescence arising from the covalently attached FAM. 130
- Figure 4.3: DNA length-dependence of target association kinetics measured for the Antp homeodomain protein in the presence and absence of 75 mg/ml Ficoll ($D_{tot} = 2.5$ nM, $P_{tot} = 50$ nM, $C_{tot} = 1000$ nM). The rate constant k_{obs} is plotted as a function of DNA length. The red line represents the best fit curve. The error bars represent the standard deviations for 8-10 replicates..... 132
- Figure 4.4: Translational diffusion measurements of free Antp homeodomain protein (0.4 mM; 3.0 mg/ml) and free 15-bp nonspecific DNA (0.4 mM; 4.0 mg/ml) in the presence and absence of Ficoll PM70 (75 mg/ml). Signal intensities were plotted as a function of gradient field strength for the Antp homeodomain in the absence **(A)** and presence **(B)** of Ficoll. Signal intensities were also plotted as a function of gradient field strength for 15-bp nonspecific DNA in the absence **(C)** and presence **(D)** of Ficoll. All experiments were run at pH 5.8 and 25°C at the ^1H frequency of 750 MHz. 135

Figure 4.5: Determination of the rotational correlation coefficient for the Antp homeodomain protein in the presence and absence of Ficoll PM70 (75 mg/ml). ^{15}N R_1 and R_2 relaxation experiments used to determine the rotational correlation time τ_c . The relaxation rates are plotted as a function of individual residues. The blue circles represent relaxation rates for residues of the Antp homeodomain in the presence of Ficoll. The red circles represent relaxation rates for residues of the Antp homeodomain in the absence of Ficoll. All experiments were run at pH 5.8 and 25°C at a ^1H frequency of 750 MHz..... 137

Figure 4.6: A schematic of ion pair mutations to study the role of ion pairs in protein sliding kinetics. 142

List of Abbreviations

Antp	Antennapedia Homeodomain
CIP	Contact Ion Pair
SIP	Solvent-separated Ion Pair
NMR	Nuclear Magnetic Resonance spectroscopy
PMF	Potentials of Mean Force
MD	Molecular Dynamics

CHAPTER 1

Introduction

1.1 BACKGROUND

1.1.1 Ion pairs of biomolecules

One of the most important and fundamental interactions in both inorganic and biological processes is the ion pair. Although the ion pair has had a rich history, much of the information was attained through small ions¹. Ion pair formation is crucial in molecular associations of biological macromolecules such as those found in protein-DNA interactions. For protein-DNA interactions, ion pairs form when the cationic moieties of basic residues interact with anionic moieties of DNA phosphates and release counter ions. For example, ion-pair interactions play a major role in DNA binding proteins, such as transcription factors, and their association to the sugar phosphate backbone of DNA. Despite the importance of ion pairs in macromolecular associations, there is still much to be investigated to deepen our understanding about their role in biological processes. X-ray crystallography has provided evidence of ion pairs for protein-protein interactions, as well as for protein-DNA/RNA interactions at functionally important sites²⁻⁴. Crystal structures have also shown ion pairs in protein-drug complexes at functionally important sites⁵⁻⁷. Therefore, it is important to understand the role of ion pairs to improve drug design. While an abundance of structural information can be drawn from crystal structures, information on the dynamic properties of ion pairs are not well understood. Furthermore, the role of ion-pair dynamics could be entropically important for molecular association of

macromolecules. Thus, it is important to understand the role of ion-pair dynamics in interactions of biological macromolecular molecules.

1.1.2 Importance of ion pairs in DNA recognition by proteins

Major cellular processes are governed by the fundamental interactions of biological macromolecules. Included are interactions found between proteins and nucleic acids of RNA or DNA. Over the years, studies of many three-dimensional structures of protein-DNA complexes led to a better understanding of protein binding and recognition^{8,9}. These structural studies provided evidence that DNA binding proteins may recognize DNA sites by base readout (direct), shape readout (indirect), or by some combination of both. It has also been suggested that the shape of the minor groove, in shape readout, may be elucidated by a set of complimentary basic side chains⁹. In fact, probing of the minor groove is established through ion pairing between the basic side chains of the protein and the DNA backbone, exhibiting the importance of ion pairs in DNA recognition.

Not only are ion pairs important in DNA recognition but they are also important in protein-DNA association. In a study of 129 different protein-DNA complexes, Luscombe *et al.* found that ion pairing between the protein and DNA backbone made up for over two-thirds of the complex-stabilizing interactions¹⁰. Interestingly, the majority of these contacts to the DNA are made by lysine and arginine residues of the protein¹⁰, emphasizing the significance of these particular amino acids. Salt dependence experiments by Privalov and co-workers provide important pieces of thermodynamic information for protein-DNA association¹¹⁻¹⁴. They performed experiments on several protein-DNA complexes, including the Antennapedia homeodomain, to observe the effects of salt concentration on protein-DNA association. By plotting the binding constants as a function of salt

concentration, a slope was obtained to estimate the number of counterions released and determine the number of ion pairs formed in the complexes. Furthermore, by running these experiments at various temperatures, they confirmed that salt dependence was completely entropic. In other words, the electrostatic (or ion pairing) component of protein-DNA association is entropic. These structural and thermodynamic data suggest the importance of intermolecular ion pairs for protein-DNA association.

1.1.3 The contact ion pair (CIP) and the solvent-separated ion pair (SIP)

Ion pairs exist in two different major states known as either the contact ion-pair (CIP) or the solvent-separated ion-pair (SIP) (Figure 1.1)^{1, 15-17}. In the CIP state, the cation forms a direct contact with the anion. In contrast, one or more water molecules intervene between the electrostatically interacting cation and anion in the SIP state.

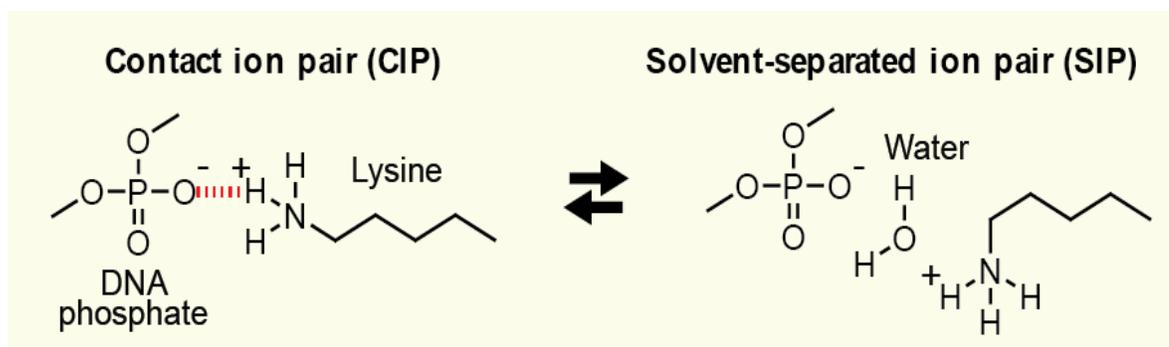


Figure 1.1: Illustration of the Contact ion pair (left) as opposed to the Solvent-separated ion pair (right).

Since the 80's, it has been known that ion pairs can undergo dynamic equilibria between the CIP and SIP states¹⁸⁻²⁰. The preference between the two states largely depends on the type of ions involved. For small ions, the preference of ion pairs toward the CIP or SIP state qualitatively obeys the law of matching water affinity or Collin's Law¹⁵. The empirical rule

proposed by this law states that anions and cations with similar affinities for water will tend to choose the CIP state. Further defined, ions can be categorized into Kosmotropic and Chaotropic ions, where Kosmotropic ions have a stronger affinity for water due to their stronger charge density. Examples of Kosmotropic anions and Chaotropic cations would be phosphate groups in nucleic acids and lysine/arginine side chains of proteins, respectively. According to Collin's Law, the ion pair formed between lysine or arginine with the phosphate backbone should prefer the SIP state. However, X-ray crystallography and recent NMR data showed that actually many of these ion pairs can be found in the CIP state. Apparently, Collin's Law does not represent a general rule for biomolecular ion pairs, and a deeper understanding regarding the physicochemical properties of biomolecular ion pairs should be pursued.

1.1.4 Potentials of Mean Force and the CIP-SIP transitions

The equilibria and dynamic transitions between CIP and SIP states of ion pairs can be understood better if their free energy landscape is available. Potentials of mean force (PMFs) for ion pairs represent the free energy landscape as a function of interionic distance and can be determined for different ion pairs using Monte Carlo (MC) or molecular dynamics (MD) simulations involving solvent atoms²¹⁻²³. Multiple minima are found in the PMFs of ion pairs, the first minimum corresponding to the CIP state and the second minimum corresponding to the SIP state (Figure 1.2)²⁴. The relative populations of the CIP and SIP states can be estimated from the free energies of the states represented by the minima. Furthermore, the peak between the CIP and SIP state in the PMFs represents the energy barrier for the CIP-SIP transitions. For example, a high energy barrier is indicative of slower transitions between the CIP and SIP states.

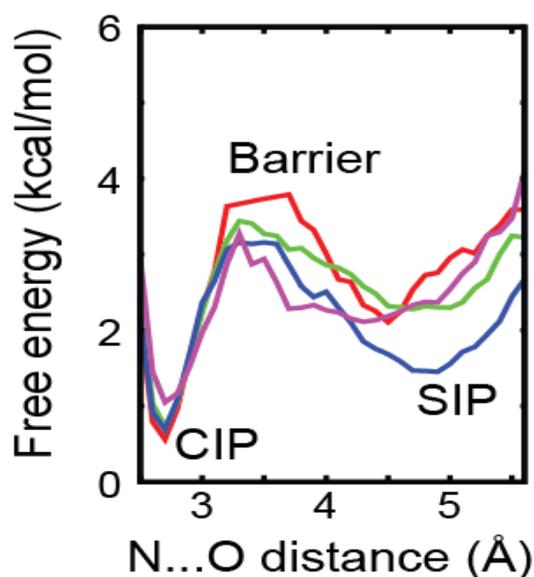


Figure 1.2: Example of PMFs for intermolecular ion pairs of the Antp homeodomain-DNA complex and Egr-1 zinc finger-DNA complex.

We obtained PMFs for intermolecular ion pairs between the protein side chains and DNA phosphates for the Antennapedia homeodomain-DNA complex and Egr-1 zinc-finger-DNA complex using 0.6- μ s MD simulations²⁴. We found that the free energy differences between the CIP and SIP states for the experimentally observed ion-pairs in the CIP states was determined to be 0.8-1.6 kcal/mol at standard temperature. The energy barriers for the transition between the CIP to the SIP state was found to be 2.2-3.2 kcal/mol, consistent with the mean lifetimes of the simulated CIP states. Their PMFs data demonstrate the ability of the ion pair to dynamically transition between the CIP and SIP state due to the relatively low energy barrier.

1.1.5 Thermodynamics of ion pairs

Much of the knowledge of ion pairs has come from studies on small organic molecules. Experimental techniques such as ultrasonic relaxation, vibrational spectroscopy (*i.e.* IR and Raman.), and relaxation spectroscopy were used to investigate thermodynamic properties of ion-pairing by small molecules¹. In fact, Marcus and Hefter described some thermodynamic consequences attributed to ion pairing in their review, and for several ion pairs of small molecules, ion pairing was entropically driven.

Part of the entropic contributions come from the release of solvent molecules upon ion pairing. Consider the local environment of the cationic and anionic moieties that form ion pairs. The strong electric fields of the ions can rearrange and even restrict nearby water molecules. This phenomenon, called solvent binding or electrostriction, leads to a reduction in entropy. When ion pairing occurs, the number of restricted water molecules decreases, in a release known as desolvation causing an increase in entropy of the system. The effect of desolvation on entropy has been well studied for small molecule ions^{1, 25}. Desolvation, although significant, is not the only contributor to the entropic effect. The release of condensed counterions and conformational dynamics also play a major role in entropic forces^{11, 26-29}.

Counterion release as a result of ion pairing in protein-DNA binding has been well studied. In fact, a relationship between the number of counterions released and the number of ion pairs formed was determined from the binding constant^{11, 30}. Work done by Privalov and co-workers showed the significance of ion pairs to entropic contributions of protein-DNA association¹¹. Since it was suggested that ion pairs entropically contribute to protein-DNA association, it is vital to understand the thermodynamic driving forces behind the

formation of ion pairs. In our work, we discuss how NMR spectroscopy can be used to measure conformational dynamics as an entropic contributor. Part of this dissertation work will serve to improve the understanding of ion-pair dynamics through determining the entropic contributions from conformational entropy upon protein-DNA association.

1.1.6 Thermodynamics of protein-DNA association

To fully appreciate the importance of ion pairs for sustaining biological macromolecule interactions, such as protein-DNA complexes, we need to elucidate the energetic parameters of protein-DNA association. Although we alluded that the contribution of ion pair dynamics to protein-DNA association is largely entropic, it is not clear how this occurs. Let us first consider the thermodynamics of protein-DNA association using the following equation:

$$\Delta G = \Delta H - T\Delta S \quad (1)$$

Where the free energy of binding (ΔG) is equal to the difference in the enthalpy of binding (ΔH) and the total entropic contributions (ΔS) in absolute temperature. Typically, the free energy and enthalpy of binding can be experimentally measured through methods such as isothermal titration calorimetry (ITC) and fluorescence anisotropy. Contributions to the change in enthalpy stem from different types of interactions such as electrostatic interactions, hydrogen bonding, and Van der Waals contacts. Entropic contributions to the free energy of binding, on the other hand, can be divided into entropic contributions arising from protein-DNA binding and from the solvent. Examples of entropic contributions from the solvent would occur due to hydrophobic effects or the release of water molecules^{1, 25}. Entropic contributions that result from protein-DNA association are further categorized into contributions from the rotational and translational changes of the protein, internal conformation or configurational changes of the protein, and the release of counterions³⁰. We

further examine the role of conformational entropy in molecular association in the next section.

1.1.7 Role of conformational entropy in molecular association

Previously, we discussed some of the significant entropic contributors to the protein-DNA binding, conformational entropy being one of them. Although conformational entropy of the protein has been considered to play an important role in protein-ligand binding for some time^{26, 31}, experimentally assessing this energetic contribution was challenging. From crystal structures, the importance of basic side chains such as arginine or lysine for protein-DNA interactions was revealed¹⁰. These side chains were observed forming hydrogen bonds with DNA bases as well as electrostatic interactions with the phosphate backbone. NMR spectroscopy has been used to make observation on the dynamics of these side chains possible, and thereby also giving insight into the conformational entropy of the protein.

Most NMR methods developed to observe dynamics have revolved around observation of protein backbone NH or side-chain CH₃ groups. Our group developed NMR experiments to determine the importance of conformational entropy in ion-pair formation and protein-DNA interactions. Using NMR relaxation-based experiments, the mobility of protein side chains can be determined. Through analysis of ¹⁵N relaxation of lysine NH₃⁺ and arginine Nε-Hε, and scalar coupling experiments mentioned above, it is possible to probe the side chain mobility of these basic side chains that form ion pairs at the molecular interface³²⁻³⁴. In short, by analyzing the ¹⁵N relaxation of the lysine NH₃⁺ or arginine Nε-Hε, the *S*² order parameters can be determined. The *S*² order parameter gives a numeric value between zero and one to provide a measure for the angular distribution of the multiple orientations of particular bond vectors^{31, 34, 35}. An *S*² order parameter closer to zero suggests that the side

chain is highly mobile, whereas an S^2 order parameter closer to one implies that the side chain is more immobile (Figure 1.3). S^2 order parameters were determined for the HoxD9 homeodomain-DNA complex and the Early growth response factor-1 (Egr-1) DNA binding domain (free and DNA-bound) previously^{24,36}. These studies allowed for the characterization of the internal motions of the basic side chains in the free and DNA-bound states. In regards to lysine side chains, experiments revealed that the lysine side chains found at the protein-DNA interface retained their mobility when forming an ion pair.

From a thermodynamic perspective, while the release of counter ions has been well studied in consideration of the counter ion condensation theory, our NMR studies revealed that there are also entropic effects as a result of ion pair dynamics. Lysine side chain relaxation experiments suggested that the mobility of the side chain is retained for the most part, which in turn, could reduce the entropic loss upon complex formation. However, there is still much to be uncovered in terms of the role of ion-pair dynamics in protein-DNA interactions.

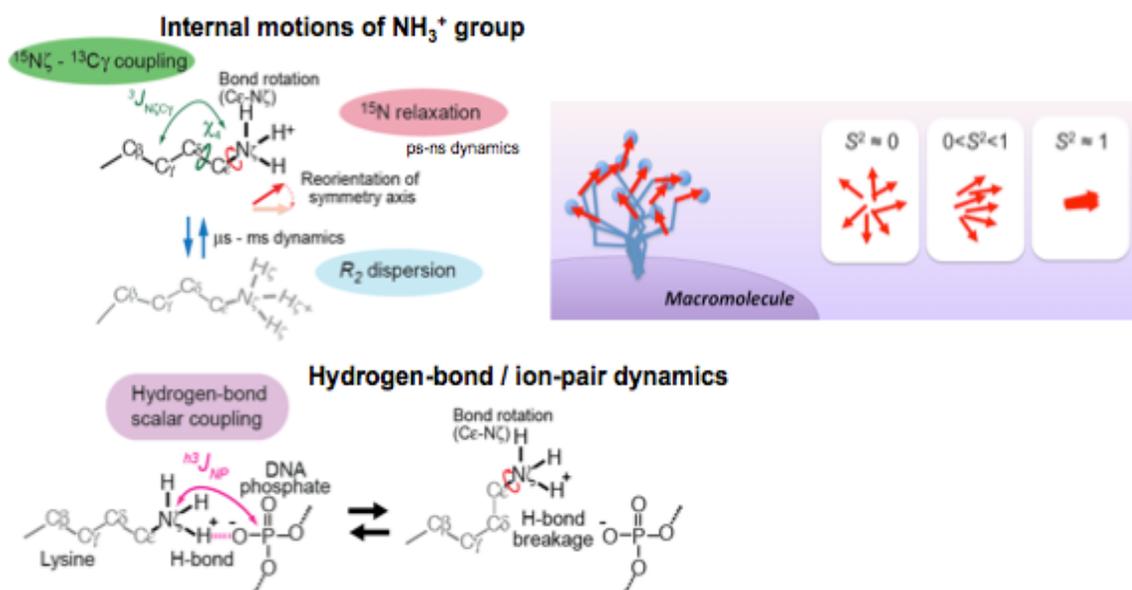


Figure 1.3: Determination of S^2 order parameters from observation of internal motions of Lys NH₃⁺, and observation of CIP by Hydrogen-bond scalar J-coupling.

1.2 GOALS AND QUESTIONS TO BE ADDRESSED

The previous section introduced the ion-pair in protein-DNA interactions and emphasized the importance of understanding ion-pair dynamics in order to delineate their role in protein-DNA interactions. While great strides have been made to improve the overall understanding of the roles of ion-pairs in protein-DNA interactions, much remains to be understood. With that in mind, the **overall goal** of this dissertation work is to determine the role of ion-pair dynamics in protein-DNA interactions. In order to progress toward this goal, we address several questions.

1.2.1 How do side chain dynamics differ in the free protein and the DNA-bound state?

Directly forming hydrogen bonds with the nucleotide bases and through electrostatic interactions with the DNA backbone are just a few of the ways proteins can interact with DNA. Basic side chains such as lysine or arginine are important for such interactions, and form salt bridges or ion pairs with the phosphate backbone of DNA. The formation of these ion pairs is important not only in driving protein association to DNA but also in shape readout or recognition^{9, 10}, in which proteins recognize their target through shape recognition of the DNA. Therefore, basic side chains play a significant role in protein-DNA interactions.

Work done on the Egr-1 zinc finger protein identified the importance of conformational mobility of basic side chains³⁷. Studies on arginine and lysine were performed to observe whether there were changes in side chain dynamics upon association of Egr-1 to its target sequence. It was observed that arginine side chains interacting with DNA bases were highly immobilized. However, both arginine and lysine side-chains interacting with the DNA

phosphate backbone remained dynamically mobile upon binding, particularly for Lys side-chains³⁷. The results of the experiments observing the side chain mobility of Egr-1 upon protein-DNA association showed that mobility was largely retained for side chains interacting with DNA phosphates, suggesting a reduction in the overall loss of conformational entropy upon protein-DNA association.

We need to consider whether the observed retention of side chain mobility upon protein association to DNA is universal for all ion pairs. Thus, in **Chapter 2**, our goal is to determine whether side chain mobility is retained in protein-DNA association for other systems as well and how this mobility affects conformational entropy. For this, we used the fruit fly Antennapedia (Antp) homeodomain and its target sequence as model system to observe side chain dynamics.

1.2.2 What role do ion-pair dynamics play in enhancing protein-DNA association?

For protein-DNA interactions, hydrogen bonding with the nucleotide bases and electrostatic interactions between basic side chains and the DNA phosphate backbone are particularly important. By X-ray crystallography, many protein-DNA complexes were studied, showing the importance of both protein interactions with the DNA bases as well as the phosphate backbone. However, while crystal structures show the existence of both CIP and SIP ion pairs formed between protein side chains and the DNA backbone, they do not demonstrate how the dynamics of ion pairs impact protein-DNA association overall. Would modifying ion pairs enhance protein-DNA affinity?

To answer this question, we use the oxygen-to-sulfur substitution of the DNA phosphate. This kind of substitution is known as either a di- or monothioate depending on how many non-bridging oxygen atoms of the phosphate are replaced with sulfur. An enhancement in

protein-DNA association with this substitution has been observed for other proteins at particular regions on the DNA³⁸⁻⁴¹. Our group also observed this enhancement when testing this substitution on the HoxD9 homeodomain-DNA complex^{36, 42}. However, not much is understood about why this enhancement occurs.

Therefore, in Chapter 3, we explore the mechanisms behind the enhancement of protein-DNA association by the sulfur substitution of the DNA phosphate through NMR spectroscopy, and other thermodynamic methods such as Isothermal Titration Calorimetry (ITC) and fluorescence anisotropy. Again, the Antennapedia homeodomain is used as a model system, and a single phosphate in the DNA sequence will be modified to include the sulfur substitution. Furthermore, the single substitution of a non-bridging oxygen atom by the oxygen-to-sulfur substitution is studied in order to understand the stereospecific effect of this substitution in ion-pair dynamics. In relation to Chapter 2, we observe whether this substitution shows an enhancement in protein-DNA association, and whether there is an impact on side chain mobility.

1.2.3 What is the role of ion-pairs in protein target search kinetics?

While X-ray crystallography showed some evidence for the importance of ion pair interactions at functionally important sites, it was not until more recent NMR experiments that evidence for the transitions between the CIP to the SIP state were made observable. Our previous work suggested that the dynamic transitions between these two equilibrium states are on the ps-ns timescale^{24, 34, 36, 37}, and these transitions may have an impact on the kinetic properties of protein.

We are trying to elucidate the role of ion pair dynamics in the kinetics of protein-DNA interactions. In a natural cellular environment, the nucleus contains a very high concentration

of DNA (100mg/ml). Transcription factors must efficiently navigate this crowded environment to reach their target DNA sequence. Binding to nonspecific and nonfunctional DNA sites can thwart the protein from reaching its target site sooner. However, the challenge of binding to non-target sites may be overcome by perturbing the dynamic transition between the CIP-SIP states. In order for the protein to dissociate from the nonspecific DNA, all contact ion pairs between the protein and DNA should be broken. If the transitions between the CIP and SIP states are rapid, the time necessary to simultaneously break all CIPs may be shortened and even facilitate sliding along the DNA to reach the target sequence (Figure 1.4). In fact, our previous data for the timescale of transitions between the CIP and SIP states estimated the time required to simultaneously break all CIPs of the homeodomain-DNA complex was $\sim 10^{-7}$ - 10^{-6} s^{24, 36}. This is comparable to the timescale of sliding estimated for the HoxD9 homeodomain⁴³, and suggests the importance of ion pairs in kinetic parameters of DNA binding proteins.

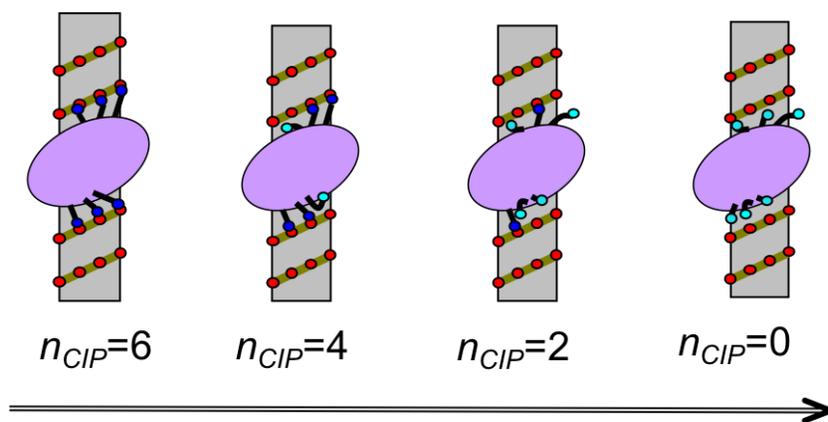


Figure 1.4: Illustration of a transcription factor breaking all of its CIPs with the DNA in order to translocate.

In **Chapter 4**, we establish an experimental approach to determining the kinetic parameters for the Antp homeodomain. With this approach, we will be able to work toward elucidating the role of ion pair dynamics in the kinetics of DNA target search by protein. Using fluorescence-based stopped-flow experiments, we observe the target association kinetics by the Antp homeodomain in the absence and presence of a synthetic macromolecular crowder called Ficoll PM70, which has been used in many crowding studies^{44, 45}. These experiments allow us to determine the 1-dimensional diffusion coefficient and sliding length of the Antp homeodomain. Because the kinetics of molecular association by proteins to DNA are a combination of 1-D and 3-D diffusion, we also use NMR spectroscopy to investigate the impact of Ficoll PM70 on the 3-D diffusion parameters of the free Antp homeodomain and free DNA duplex in solution.

1.3 OVERVIEW OF METHODS AND MODEL SYSTEM

1.3.1 Model system for research

In this work, the Antennapedia homeodomain is used as a model system to study the role of ion-pair dynamics in protein-DNA interactions. Homeodomains were discovered in the early 80's in *Drosophila*, when it was found that mutations in the genes induced very peculiar physiological changes⁴⁶. One example of these mutations is Antennapedia, where it was found that a mutation to this gene caused the growth of legs in place of the antennae.

Antp has been very well characterized biophysically and biochemically and can be easily expressed in *E. coli*^{12, 47}. As seen in figure 1.5, the crystal structures for the Antp homeodomain-DNA complexes have also been solved. The Antp homeodomain binds to DNA as a monomer via a helix-turn-helix motif⁴⁶⁻⁴⁹. In the crystal structure of the specific

complex, up to six intermolecular ion pairs are formed between the Antp homeodomain and DNA.

In this work, a 60-amino acid construct of the Antp homeodomain and 15-bp DNA duplex containing the 6-bp recognition sequence are used. In order to explore the role of ion pair dynamics in this work, we used NMR spectroscopy, and long periods of time are required to complete NMR experiments. The Antp homeodomain does not degrade easily and can remain stable for several months, making this protein ideal for NMR studies. Most importantly, NMR spectra for the Antp homeodomain give well-isolated signals for all interfacial Lys and Arg side chains, allowing quantitative analyses of their ion-pair dynamics. Information from the crystal structures of the Antp homeodomain-DNA are also available⁵⁰, as well as detailed thermodynamic data for the Antp homeodomain-DNA association¹².

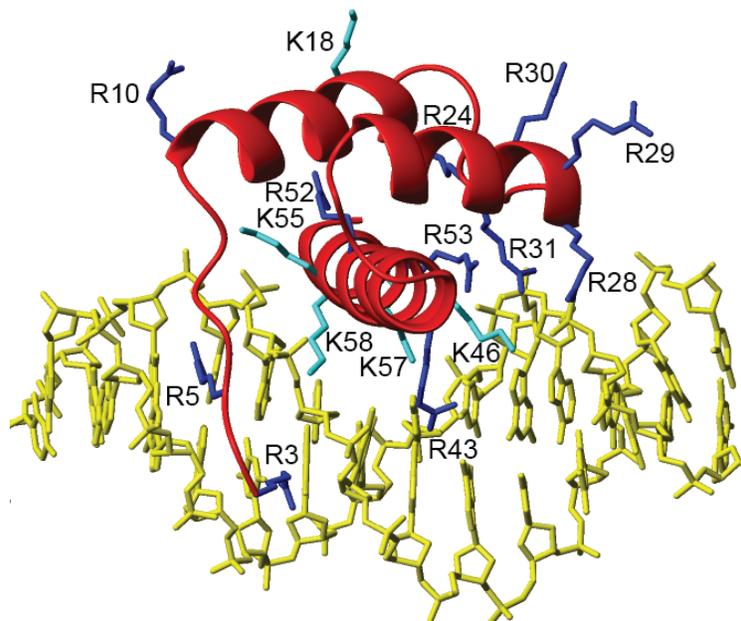


Figure 1.5: Crystal structure of the Antp Homeodomain-DNA complex (PDB 4XID).

1.3.2 Phosphorothioate: the oxygen-to-sulfur substitution

The phosphorothioation of oligonucleotides has been used for some time in oligonucleotide modifications³⁹⁻⁴¹. In this modification, sulfur atoms substitute either one or both of the non-bridging oxygen atoms of the phosphate backbone (Figure 1.6). This results in what is called a mono- or dithioated phosphate. Chemically, this substitution does not cause a deviation in overall charge, and even maintains a tetrahedral covalent geometry. In addition, this kind of chemical modification has been shown to not only improve cellular uptake, but also aids the oligonucleotide in avoiding cellular nucleases⁵¹⁻⁵³. For our

research, the most interesting and useful characteristic of this substitution is its ability to enhance binding affinity^{36, 38, 40, 41, 54}.

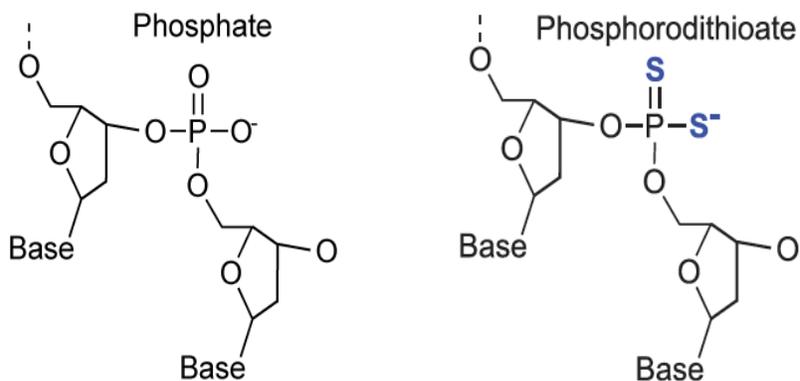


Figure 1.6: Chemical composition of a typical DNA phosphate (left) and a dithioated phosphate with the oxygen-to-sulfur substitution (right).

Our group previously studied the phosphorodithioate modification in the HoxD9-DNA complex, in which a single phosphate at the interface with protein was modified. NMR studies on this modification suggested that the lysine interacting with dithioate at the interfacial region actually showed an increase in side chain mobility. Although this observation was quite interesting, information explaining the enhancement in side chain mobility is still limited. Because we observed the retainment of interfacial lysine mobility upon protein-DNA binding⁵⁵ and hypothesized that this conformational mobility can be attributed to an entropic contribution, the oxygen-to-sulfur substitution can be used as a great tool to assess the entropic role of ion pair dynamics in protein-DNA association. In other words, if our hypothesis holds true, we should be able to observe an enhanced mobility for the Lys side chain interacting with the sulfur-substituted phosphate and observe an enhanced binding affinity for the modified DNA upon protein association.

1.3.3 NMR spectroscopy methods to determine side-chain mobility of ion pairs

We used NMR spectroscopy in order to observe the conformational dynamics of interfacial lysine and arginine side chains of the Antp homeodomain. NMR relaxation-based methods that observe the ^{15}N nuclei of lysine NH_3^+ or arginine $\text{N}\epsilon\text{--H}\epsilon$ were developed by our group and used in this dissertation work^{17, 24, 34, 36, 37, 42, 56}. In fact, this dissertation includes a subsection on improving the methodology behind ^{15}N relaxation experiments for both lysine and arginine side chains. As seen in Figure 1.7, various relaxation experiments for both the side chain and backbone are used to determine side-chain order parameters S^2 . The S^2 order parameter is derived from relaxation-based experiments by using the model-free approach⁵⁷. Lipari and Szabo showed that information from the fast-internal motions of NMR relaxation experiments could be used to determine both the S^2 order parameters and the correlation times for bond reorientations. In this dissertation work, S^2 order parameters and bond correlation times are determined for important lysine and arginine side-chain residues of the Antp homeodomain that form ion pairs with the DNA phosphate backbone. The S^2 order parameter, giving information on mobility, can then be used to estimate conformational entropy as there is a strong correlation between the two^{26, 31, 35}.

In order to observe the dynamics of the torsion angles of the basic amino acid side-chains, which are also important for observation of side-chain dynamics, three-bond scalar coupling constants are quite useful⁵⁸⁻⁶⁰. For example, χ_4 torsion angles were observed for lysine side chain NH_3^+ using coupling between the $^{15}\text{N}_\zeta$ and $^{13}\text{C}_\gamma$ nuclei ($^3J_{\text{N}_\zeta \text{C}_\gamma}$). The average ensemble of $^3J_{\text{N}_\zeta \text{C}_\gamma}$ can also be determined through MD simulations, as well as crystal structures. Our group compared experimentally determined $^3J_{\text{N}_\zeta \text{C}_\gamma}$ from NMR

experiments with both the ${}^3J_{N\zeta C\gamma}$ determined from crystal structures and MD simulations and found that the crystal structures display a more bimodal distribution for two different χ_4 conformers. On the other hand, there was excellent agreement with the ${}^3J_{N\zeta C\gamma}$ constants calculated from MD simulations²⁴, suggesting that the lysine side chains χ_4 really are as dynamic as the MD simulations projected.

With the use of hydrogen-bond scalar coupling experiments, we can detect coupling for ion pairs in the CIP state. Because of relatively slow transverse ${}^{15}\text{N}$ relaxation, these scalar couplings can be observed for Lys side-chain NH_3^+ groups. Our NMR experiments showed direct evidence of side chains observed in the CIP state through these scalar coupling experiments. In fact, we observed evidence of ion pairs in the CIP state in experiments for the HoxD9-DNA complex and the Egr-1 zinc-finger-DNA complex^{24, 37, 42}. We found that only CIPs have an observable chemical shift in hydrogen-bond scalar coupling experiments because only CIPs display a J -coupling. In these experiments, the hydrogen-bond scalar coupling is observed between the lysine side chain ${}^{15}\text{N}$ and ${}^{31}\text{P}$ DNA nuclei at the molecular interface where ion pairs are formed. These NMR experiments will allow us to quantitatively analyze basic side chain mobility for interfacial ion pairs to elucidate the role of ion pairs dynamics on conformational entropy.

Determination of Arg/Lys side-chain order parameters S^2

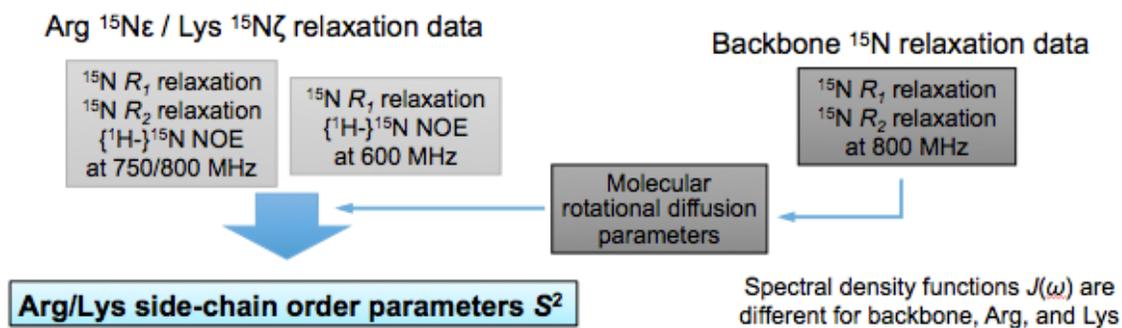


Figure 1.7: Schematic for determination of S^2 order parameters from various NMR experiments.

CHAPTER 2

Internal Motions of Basic Side Chains of the Antennapedia

Homeodomain in the Free and DNA-Bound States¹

2.1 BACKGROUND

In protein-DNA complexes, arginine (Arg) and lysine (Lys) side chains interact with DNA phosphates to form intermolecular ion pairs, serving as crucial constituents of the molecular interfaces.^{10, 61} The electrostatic interactions of basic side chains with DNA are important not only as a major driving force for protein-DNA association but also as a shape-readout mechanism to recognize particular DNA⁶¹. Direct readout through hydrogen bonds between basic side chains and DNA bases are also common.^{10, 61} Thus, basic side chains of proteins play crucial roles in association with and recognition of DNA.

Recent studies have shown that conformational mobility makes significant entropic contribution to the thermodynamics of macromolecular association.^{26, 62} To understand DNA recognition by proteins from a thermodynamic viewpoint, the dynamic properties of basic side chains should be delineated. Some NMR methods have been developed for dynamics investigations of basic side chains.^{32, 42, 63-69} Using these methods, we recently studied changes in the mobility of the basic side chains of the Egr-1 zinc-finger protein upon association with its target DNA.³⁷ Interestingly, the basic side chains that form ion pairs with the DNA phosphates were found to retain high mobility despite the simultaneous

¹ Chapter adapted with permission from American Chemical Society (see Appendix). Nguyen, D., et al. (2017) Internal motions of basic side chains of the Antennapedia homeodomain in the free and DNA-bound states. *Biochemistry* 56, 5866-9.

presence of hydrogen bonds and strong short-range electrostatic interactions. This trend was particularly remarkable for the Lys side chains. In contrast, the Arg side chains that directly interact with the DNA bases were found to lose substantial mobility upon binding. To examine whether or not these are general characteristics of basic side chains, other systems should also be studied.

For this purpose, we study the dynamics of the basic side chains of the fruit fly *Antennapedia* (Antp) homeodomain in the free and DNA-bound states. The Antp homeodomain is comprised of 60 residues, 18 of which are basic side chains (12 Arg and 6 Lys residues) (Figure 2.1). This protein recognizes DNA sequence, TAATGG, and binds to it with a dissociation constant on the order of 10^{-8} M at physiological ionic strength.⁷⁰ Five crystal structures are available for the Antp homeodomain-DNA complexes.^{70, 71} NMR structures for the DNA complex of a 68-residue construct of the Antp homeodomain are also available.^{72, 73} These structures show that R3, R5, R28, R31, R43, K46, R53, K55, and K57 interact with DNA. We previously studied dynamics of Lys side chains of the same Antp homeodomain-DNA complex.²⁴ In the current work, using NMR spectroscopy, we investigate dynamics of the Arg and Lys side chains in the free and DNA-bound states of the Antp homeodomain and analyze changes in side-chain mobility upon the molecular association.

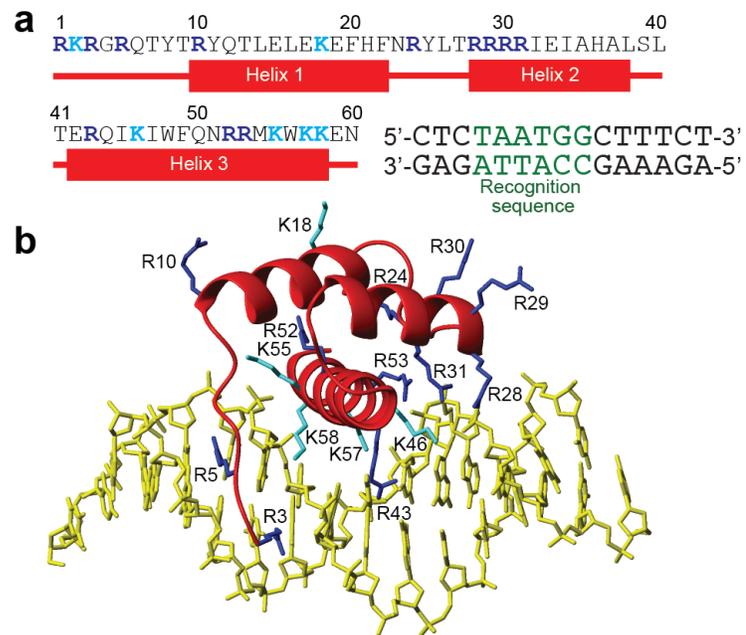


Figure 2.1: The basic side chains in the Antp homeodomain-DNA complex. (a) Sequences of the Antp homeodomain C39S mutant and the 15-bp DNA duplex used in this study. (b) A crystal structure of the Antp homeodomain-DNA complex (PDB 4XID).

2.2 MATERIALS AND METHODS

2.2.1 Protein and DNA

A 60 amino-acid construct of the fruit fly Antp homeodomain with the C39S mutation (the amino-acid sequence shown in Figure 2.1a) was expressed in *E. coli* strain BL21 (DE3) and purified using SP fast-flow cation-exchange, S-100 size-exclusion, and Resource-S cation exchange columns (GE Healthcare) as previously described.^{24, 70, 74} The protein was quantified using UV light absorbance at 280 nm together with an extinction coefficient of 15,470 M⁻¹ cm⁻¹ (<http://web.expasy.org/protparam/>). Individual strands of the 15-bp DNA shown in Figure 1a were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). Each DNA strand was purified using a Mono Q anion-exchange column (GE Healthcare) with a GE ÄKTA purifier system. The 15-bp DNA duplex was formed through annealing of the complimentary strands and isolated by second Mono Q anion-exchange chromatography.^{24, 70, 74}

To prepare the complex of the 15N Antp homeodomain and the unlabeled 15-bp DNA duplex, the protein and DNA solutions at ~500 mM NaCl was mixed at a molar ratio of 1:1.5. Using an Amicon Ultra-4 centrifugal device, the solution buffer was changed to 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 mM NaF (as a preservative). The final concentration of the complex was 0.4 mM. 100 µl of D2O was separately sealed for NMR lock in a co-axial insert (Norell; diameter, 2 mm) to avoid partial deuteration of Lys NH₃⁺ groups (i.e., NDH₂⁺ and ND₂H⁺ species). A 370 µl solution of 0.4 mM 15N-Antp homeodomain or its complex with DNA was sealed in an outer tube (diameter, 5 mm) of the co-axial NMR tube.

The same conditions were also used for the NMR sample to measure Arg side-chain ^{15}N relaxation for the Antp homeodomain in the free state. Only for the sample to measure ^{15}N relaxation of Lys side-chain NH_3^+ groups of the free protein at 2°C , the buffer pH was adjusted to 4.5, and methanol- d_4 was separately sealed instead of D_2O for NMR lock in the inner tube of the co-axial tube because pure D_2O (freezing point, 4°C) can freeze at this temperature. The lower pH and temperature were required to observe signals from Lys side-chain NH_3^+ groups of the Antp homeodomain in the free state.

2.2.2 NMR relaxation analysis for Lys and Arg side chains

The ^{15}N relaxation measurements for Lys side-chain NH_3^+ groups were performed as previously described.^{42, 65, 68, 70, 75} For Lys NH_3^+ groups, ^{15}N R_1 and heteronuclear NOE data were recorded at the ^1H frequencies of 800 and 600 MHz and ^{15}N R_2 data were recorded at the ^1H frequency of 800 MHz. These measurements for the complex were conducted at 15°C and pH 5.8 under the above-mentioned buffer conditions. The Lys ^{15}N relaxation experiments for the free protein were conducted at 2°C and pH 4.5. The lower temperature and pH were necessary to mitigate broadening of the Lys NH_3^+ signals due to rapid hydrogen exchange.^{42, 65} To analyze the contribution of slow dynamics to the observed R_2 relaxation rates, Lys ^{15}N R_2 relaxation dispersion experiment was performed at the ^1H frequency of 800 MHz, as previously described.⁶⁵

The use of co-axial tubes further decreases sensitivity due to a smaller sample volume and multilayer glass walls. Thus, sensitivity improvement would be desirable for NMR experiments on NH_3^+ groups, especially for quantitative experiments such as ^{15}N relaxation measurements. The first step for measuring ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates is to create the ^{15}N in-phase single-quantum term via coherence transfer

from ^1H to ^{15}N nuclei through a refocused INEPT scheme ⁷⁶. The product operator terms N_x , $2N_yH_z$, $4N_xH_zH_z$, and $8N_yH_zH_zH_z$ are generated in the period of $2\tau_b$ in the first refocused INEPT scheme of our pulse sequence for ^{15}N R_1 and R_2 measurements. The $2N_yH_z$ and $8N_yH_zH_zH_z$ terms are eliminated by the pulsed field gradient (PFG) g_4 after the ^1H $90^\circ(-x)$ and ^{15}N $90^\circ(y)$ pulses at the end of the refocused INEPT scheme, but the $4N_xH_zH_z$ term is not eliminated because a PFG alone cannot destroy the homonuclear zero-quantum coherence $4N_zH_yH_y$ ⁷⁷. The coefficients of the transfers from $2N_yH_z$ to N_x and to $4N_xH_zH_z$ are given by $\cos^2\theta\sin\theta$ and $(3\cos^2\theta - 1)\sin\theta$, respectively, in which $\theta = 2\pi^l J_{NH}\tau_b$ and $^l J_{NH}$ represents the one-bond ^1H - ^{15}N scalar coupling constant ⁷⁶. The use of the time τ_b satisfying $3\cos^2\theta - 1 = 0$ thus eliminates the $4N_xH_zH_z$ term but retains the N_x term, and this condition is achieved by $\tau_b = 2.1$ ms in the original pulse sequences ⁶⁵. This approach was also used for ^{13}C R_1 and R_2 relaxation measurements for protein CH_3 groups ^{78,79}. A practical problem of this approach is that it reduces the efficiency of the transfer from $2N_yH_z$ to N_x (i.e., $\cos^2\theta\sin\theta$) and weakens the signals from NH_3^+ groups. We eliminate the adverse effects of the $4N_xH_zH_z$ term in a different manner and maximize $f_{CT}(2N_yH_z \rightarrow N_x)$ to increase sensitivity in ^{15}N relaxation measurements for NH_3^+ groups. As shown in Figure 2.2d, the signal arising from the N_x term should be strongest when $\tau_b = 1.3$ ms. Although this condition increases the $4N_xH_zH_z$ term generated through the refocused INEPT scheme, our pulse sequences shown in Figure 2.2 prevent the undesired $4N_xH_zH_z$ term from becoming

observable in the ^1H detection period t_f . This allows us to use $\tau_b = 1.3$ ms and improve sensitivity without compromising accuracy in ^{15}N relaxation measurements.²

The ^{15}N relaxation experiments for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups were performed at 25°C with the pulse sequences for NH groups together with selective ^{15}N 180° rSNOB pulses⁸⁰ (1.0 ms) in the INEPT schemes. By using ^{15}N carrier position at 85 ppm together with these selective pulses, the Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ resonances were selectively observed in these ^{15}N relaxation experiments. For these Arg $\text{N}_\epsilon\text{-H}_\epsilon$ groups, ^{15}N R_1 and heteronuclear NOE data were recorded at the ^1H frequencies of 750 and 600 MHz and ^{15}N R_2 data were recorded at the ^1H frequency of 750 MHz. The ^{15}N R_2 CPMG relaxation dispersion experiment for Arg $^{15}\text{N}_\epsilon$ nuclei was performed at the ^1H frequency of 750 MHz using the CW-CPMG scheme.⁸¹ The CPMG frequencies (ν_{CPMG}) used in these measurements were 33, 67, 100, 200, 333, 500, 667, 1000, 1333, and 1667 Hz.

The pulse sequences shown in Figure 2.2 can be used to measure ^{15}N R_1 , R_2 , and heteronuclear NOE for Arg side-chain $\text{N}_\epsilon\text{H}_\epsilon$ groups as well, though changes in some parameters are required due to the differences between Lys NH_3^+ and Arg $\text{N}_\epsilon\text{H}$ groups. The caption of Figure 2.2 shows the required changes, many of which are related to the different spin systems (AX_3 vs. AX) and $^1J_{\text{NH}}$ coupling constants (74 Hz vs. 92 Hz) of Lys NH_3^+ and Arg $\text{N}_\epsilon\text{H}$ groups. In the ^{15}N relaxation measurements for Arg $\text{N}_\epsilon\text{H}$ groups, we typically use a ^{15}N carrier position at ~ 84 ppm and a ^{15}N spectral width of ~ 6 -12 ppm. These settings together with the use of ^{15}N rSNOB 180° pulses in the INEPT schemes allow ^{15}N relaxation

² Paragraph adapted from the open access journal *Molecules*. Nguyen, et al. (2017) A unique and simple approach to improve sensitivity in ^{15}N NMR relaxation measurements for NH_2 groups: Application to a protein-DNA complex. *Molecules* 22, E1355 and Nguyen, et al. (2018) NMR methods for characterizing the basic side chains of proteins: electrostatic interactions, hydrogen bonds, and conformational dynamics. Submitted to *Methods in Enzymology*

measurements for Arg side-chain $^{15}\text{N}_\epsilon$ nuclei only. Although simultaneous detection of Arg side-chain $^{15}\text{N}_\epsilon$ and backbone ^{15}N nuclei is possible, separate measurements have the following advantages. First, selective observation of Arg $^{15}\text{N}_\epsilon$ allows the use of a narrower ^{15}N spectral width for a high resolution without any concerns about undesired overlaps with folded signals from backbone NH or Lys NH_3^+ groups. Second, R_2 measurements with the ^{15}N carrier position set to the Arg N_ϵ resonance allow application of on-resonance 180° pulses in the CPMG scheme. Simultaneous R_2 measurements of backbone and Arg side-chain ^{15}N nuclei require the use of a ^{15}N carrier position that introduces significant off-resonance effects to the CPMG scheme⁸² and decrease the data quality for either type of ^{15}N nuclei. Furthermore, selective observation of Arg $\text{N}_\epsilon\text{H}$ groups allows higher sensitivity because the conditions optimized for backbone NH groups considerably weaken the signals from Arg $\text{N}_\epsilon\text{H}$ groups unless broadband ^{15}N 180° pulses are used⁸³. We also considered the presence of two geminal couplings to $^{15}\text{N}_{\eta 1}$ and $^{15}\text{N}_{\eta 2}$ nuclei, which can in principle affect measurements of transverse relaxation of Arg $^{15}\text{N}_\epsilon$ nuclei. However, our recent study shows that the impact of the $^2J_{NN}$ couplings on CPMG-based transverse relaxation measurements for Arg $^{15}\text{N}_\epsilon$ nuclei is typically negligible⁸⁴. For these reasons, we conduct the ^{15}N relaxation experiments for backbone NH groups, Lys NH_3^+ groups, Arg $\text{N}_\epsilon\text{H}$ group separately^{84,3}.

To determine the rotational diffusion parameters, the backbone ^{15}N R_1 and R_2 relaxation rates at the ^1H frequency of 800 MHz were measured. These measurements for the free

³ Nguyen, et al. (2018) NMR methods for characterizing the basic side chains of proteins: electrostatic interactions, hydrogen bonds, and conformational dynamics. Submitted to Methods in Enzymology

protein were conducted at 25°C and pH 5.8 as well as at 2°C and pH 4.5. The measurements for the complex were conducted at 25°C and 5°C (both at pH 5.8).

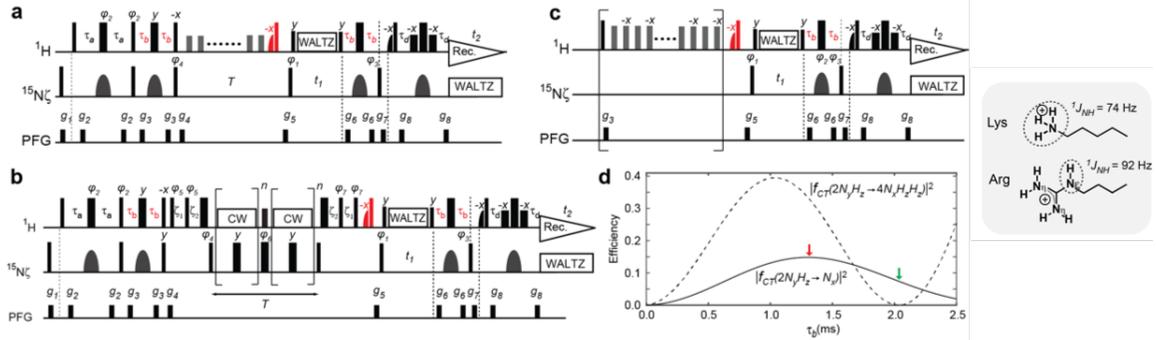


Figure 2.2: Pulse sequences for the ^{15}N relaxation measurement for lysine side-chain NH_3^+ groups. The key elements in the current work are indicated in red. Thin and bold bars in black represent hard rectangular 90° and 180° pulses, respectively. Water-selective half-Gaussian (2.1 ms) and soft-rectangular (1.2 ms) 90° pulses are represented by half-bell and short-bold shapes, respectively. Unless indicated otherwise, pulse phases are along x , and the carrier position for ^1H was set to the position of the water resonance. The RF strength for hard rectangular ^1H and ^{15}N pulses were 40 kHz and 6.6 kHz, respectively. The RF strengths of ^1H WALTZ-16 decoupling⁸⁵ during the ^{15}N evolution period was 4.2 kHz. ^{15}N carrier position was set to 33.1 ppm. ^{15}N -decoupling during the t_2 detection period was performed with WALTZ-16 (RF strength, 1.0 kHz). A gray bell-shape for ^{15}N represents an r-SNOB⁸⁰ 180° pulse (1.0 ms) for selective inversion and refocusing of Lys side-chain $^{15}\text{N}_\zeta$ nuclei. The delays τ_a and τ_b were 2.7 ms and 1.3 ms, respectively. Quadrature detection in the t_1 domain was achieved using States-TPPI, incrementing the phase φ_1 . Pulsed field gradients (PFGs) were optimized to minimize the water signal. **(a)** ^{15}N R_1 measurement. Although it is not essential owing to negligible CSA-DD cross correlation for NH_3^+ , a ^1H 180° pulse, which does not affect H_2O resonance, was applied every 10 ms during the delay T for longitudinal relaxation. Phase cycles: $\varphi_1 = (2y, 2(-y))$, $\varphi_2 = (y, -y)$, $\varphi_3 = (4x, 4(-x))$, $\varphi_4 = (8y, 8(-y))$, and receiver = $(x, -x, -x, x, 2(-x, x, x, -x), x, -x, -x, x)$. **(b)** ^{15}N $R_{2,ini}$ measurement. The RF strength for ^{15}N pulses for the CPMG scheme was 5.4 kHz. ^1H carrier position was shifted to 7.8 ppm right after the PFG g_4 and set back to the position of water resonance right after the PFG g_5 . The RF strength $\omega_{CW}/2\pi$ of ^1H CW during the CPMG was set to 4.3 kHz, which was adjusted to satisfy $\omega_{CW}/2\pi = 2k\nu_{CPMG}$ (k , integer).⁸¹ The delays ξ_1 and ξ_2 are for alignment of ^1H magnetization and given by $\xi_1 = 1/\omega_{CW} - (4/\pi)\tau_{90H}$ and $\xi_2 = \tau_{90N} - (2/\pi)\tau_{90H}$,^{81, 86} in which τ_{90} represents a length of a relevant 90° pulse. Phase cycles: $\varphi_1 = (4y, 4(-y))$, $\varphi_2 = (8y, 8(-y))$, $\varphi_3 = x$, $\varphi_4 = (x, -x)$, $\varphi_5 = (2y, 2(-y))$, $\varphi_6 = (2x, 2(-x))$, $\varphi_7 = (2(-y), 2y)$, and receiver = $(x, -x, x, -x, 2(-x, x, -x, x), x, -x, x, -x)$. **(c)** Heteronuclear ^1H - ^{15}N NOE measurement. Measurement with ^1H saturation (5 s) was performed with a train of $180^\circ x$ and $180^\circ(-x)$ pulses (RF strength, 11 kHz) with an interval of 10 ms. ^1H carrier position was at 7.8 ppm during the ^1H saturation period. The reference spectrum was measured without the scheme in the bracket. The recycle delay (including the saturation period) was set to 18 s for a 750-MHz spectrometer. Phase cycles: $\varphi_1 = (y, -y)$, $\varphi_2 = (4x, 4y, 4(-x), 4(-y))$, $\varphi_3 = (2x, 2(-x))$, and receiver = $(x, -x, -x, x, -x, x, x, -x)$. **(d)** Efficiency in coherence transfers as a function of the delay τ_b calculated using Eqs. 1 and 2 with $|^1J_{NH}| = 74$ Hz and ^1H 180° pulse length of 20 μs . Results for the N_y and $4N_y H_z H_z$ terms are shown in solid and dotted lines, respectively. Red and green arrows indicate the values of the delay τ_b in the current and previous pulses sequences, respectively. Due to the different spin systems and ^{15}N chemical shift ranges, the experiments for Lys and Arg side chains should be conducted separately.

2.2.3 Determination of order parameters for Lys and Arg side chains

Rotational diffusion parameters (D_{\parallel} , D_{\perp} , and two polar angles for the main principal axis) for the axially symmetric diffusion model⁸⁷ were determined from the backbone ^{15}N relaxation data using a C program together with GNU Scientific Library, as described.^{88, 89} The effective rotational correlation time $\tau_{r,eff}$ and the anisotropy of the rotational diffusion r are given by $(2D_{\parallel} + 4D_{\perp})^{-1}$ and $D_{\parallel} / D_{\perp}$, respectively.⁸⁷ Using MATLAB software, the order parameters for Arg $\text{N}_{\epsilon}\text{-H}_{\epsilon}$ groups were calculated from the relaxation data at the two magnetic fields. The ^{15}N chemical shift anisotropy parameter ($\sigma_{\parallel} - \sigma_{\perp}$) for arginine side-chain $^{15}\text{N}_{\epsilon}$ nuclei was set to -114 ppm and the $\text{N}_{\epsilon}\text{-H}_{\epsilon}$ distance was set to 1.04 Å according to Trbovic et al.³² Four spectral density functions were tested for each Arg $\text{N}_{\epsilon}\text{-H}_{\epsilon}$ group: two of them were the model-free functions of Lipari and Szabo (Eqs. 35 and 43 in Ref. ⁹⁰) and the others were the extended model-free functions of Clore et al. (Eqs. 2 and 4 in Ref. ⁹¹ multiplied by 2/5). The best model among the four spectral density functions was selected using Akaike's information criterion (AIC) calculated for each model.⁹² The order parameters for Lys side-chain NH_3^+ groups were calculated from the ^{15}N relaxation data at the two magnetic fields, as previously described in detail by Esadze et al.⁶⁵

2.3 RESULTS

2.3.1 Internal motions of Arg side-chain and $\text{N}_{\epsilon}\text{-H}_{\epsilon}$ groups

We compared the internal motions of the Arg guanidinium groups of the Antp homeodomain in the free and DNA-bound states under the same buffer conditions and temperature. Figure 2.3a shows the $^1\text{H}\text{-}^{15}\text{N}$ heteronuclear in-phase single quantum coherence (HISQC) spectra⁶⁴ recorded for the Arg side-chain $\text{N}_{\epsilon}\text{-H}_{\epsilon}$ moieties of the Antp homeodomain in the free and DNA-bound states at 25°C. Both samples were dissolved in

a buffer of 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 mM NaF (as a preservative). The Arg side chains exhibited well-isolated signals in the ^1H - ^{15}N heteronuclear correlation spectra for both states. For the Arg side-chain N_ϵ nuclei, we measured ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates and heteronuclear NOE at the ^1H frequency of 750 MHz. Heteronuclear NOE and ^{15}N R_1 data were collected at 600 MHz as well. These relaxation data are shown in Tables 2.1 and 2.2. Using these data, we determined the generalized order parameters (S^2)⁹⁰ for the Arg N_ϵ - H_ϵ bonds of the Antp homeodomain in the free and DNA-bound states, as described.³⁷ The S^2 parameter satisfies the inequalities of $0 \leq S^2 \leq 1$, and represents a measure of the degree of spatial restriction of internal motion.⁹⁰ Changes in order parameters upon molecular association are related to changes in conformational entropy.²⁶ The molecular rotational diffusion parameters (Table 2.3) were determined from backbone ^{15}N relaxation rates R_1 and R_2 .³⁷ The data of the generalized order parameters (S^2) determined for Arg N_ϵ - H_ϵ bonds in the free protein and in the complex are shown in Figure 2.3b as well as Table 2.4.

Table 2.1: ^{15}N NMR relaxation data for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups of the Antp homeodomain in the free state at pH 5.8 and 25°C.

Arg $\text{N}_\epsilon\text{-H}_\epsilon$	$^{15}\text{N}_\epsilon R_1$ (s^{-1})	$^{15}\text{N}_\epsilon R_1$ (s^{-1})	$^{15}\text{N}_\epsilon R_2$ (s^{-1})	$\{^1\text{H}\}\text{N}_\epsilon\text{NOE}$	$\{^1\text{H}\}\text{N}_\epsilon\text{NOE}$
	600 MHz	750 MHz	750 MHz	600 MHz	750 MHz
R1	0.34 ± 0.02	0.33 ± 0.01	2.09 ± 0.17	-2.31 ± 0.16	-1.51 ± 0.08
R3	0.44 ± 0.01	0.39 ± 0.01	1.31 ± 0.06	-1.76 ± 0.07	-1.34 ± 0.04
R5	0.40 ± 0.01	0.34 ± 0.01	1.65 ± 0.08	-1.89 ± 0.08	-1.51 ± 0.05
R10	0.57 ± 0.01	0.49 ± 0.01	1.93 ± 0.06	-1.15 ± 0.05	-0.91 ± 0.03
R24	1.11 ± 0.01	0.93 ± 0.01	3.80 ± 0.07	0.04 ± 0.03	0.13 ± 0.03
R28	0.82 ± 0.02	0.68 ± 0.01	2.08 ± 0.05	-0.64 ± 0.04	-0.46 ± 0.02
R29	0.58 ± 0.01	0.51 ± 0.01	2.38 ± 0.05	-1.07 ± 0.04	-0.79 ± 0.02
R30	1.47 ± 0.02	1.15 ± 0.01	4.95 ± 0.05	0.51 ± 0.03	0.58 ± 0.02
R31	1.42 ± 0.03	1.13 ± 0.01	5.39 ± 0.07	0.43 ± 0.04	0.45 ± 0.03
R43	0.60 ± 0.03	0.60 ± 0.02	2.30 ± 0.21	-1.20 ± 0.13	-0.69 ± 0.07
R52	1.58 ± 0.03	1.20 ± 0.01	5.47 ± 0.07	0.56 ± 0.04	0.62 ± 0.04
R53	1.56 ± 0.03	1.19 ± 0.01	5.67 ± 0.10	0.47 ± 0.05	0.63 ± 0.04

Table 2.2: ^{15}N NMR relaxation data for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups of the Antp homeodomain in the specific complex with 15-bp DNA at pH 5.8 and 25°C.

Arg $\text{N}_\epsilon\text{-H}_\epsilon$	$^{15}\text{N}_\epsilon R_1$ (s^{-1})	$^{15}\text{N}_\epsilon R_1$ (s^{-1})	$^{15}\text{N}_\epsilon R_2$ (s^{-1})	$\{^1\text{H}\}\text{N}_\epsilon\text{NOE}$	$\{^1\text{H}\}\text{N}_\epsilon\text{NOE}$
	600 MHz	750 MHz	750 MHz	600 MHz	750 MHz
R1	0.70 ± 0.01	0.59 ± 0.01	2.64 ± 0.02	-0.68 ± 0.01	-0.49 ± 0.01
R3	0.96 ± 0.01	0.76 ± 0.01	5.44 ± 0.03	0.04 ± 0.01	0.09 ± 0.01
R5	1.04 ± 0.02	0.77 ± 0.01	9.24 ± 0.05	0.52 ± 0.01	0.52 ± 0.01
R10	0.58 ± 0.01	0.53 ± 0.01	2.27 ± 0.02	-1.05 ± 0.01	-0.78 ± 0.01
R24	0.92 ± 0.01	0.73 ± 0.01	5.75 ± 0.03	0.04 ± 0.01	0.11 ± 0.01
R28	0.92 ± 0.01	0.78 ± 0.01	3.67 ± 0.02	-0.33 ± 0.01	-0.24 ± 0.01
R29	0.53 ± 0.01	0.49 ± 0.01	3.09 ± 0.02	-1.08 ± 0.01	-0.81 ± 0.01
R30	0.99 ± 0.01	0.73 ± 0.01	9.48 ± 0.06	0.51 ± 0.01	0.53 ± 0.01
R31	1.17 ± 0.02	0.88 ± 0.01	9.03 ± 0.08	0.53 ± 0.01	0.53 ± 0.02
R43	0.87 ± 0.01	0.71 ± 0.01	5.18 ± 0.03	-0.12 ± 0.01	-0.05 ± 0.01
R52	1.00 ± 0.02	0.70 ± 0.01	10.9 ± 0.1	0.72 ± 0.02	0.73 ± 0.02
R53	1.11 ± 0.07	0.73 ± 0.02	13.4 ± 0.5	0.79 ± 0.05	0.66 ± 0.06

Table 2.3: Molecular rotational diffusion parameters for the Antp homeodomain in the free and DNA-bound states. ^{a)}

Molecules	$\tau_{r,eff}$ (ns) ^{b)}	$D_{\parallel} / D_{\perp}$ ^{b)}
Free protein		
at 25°C	4.86 ± 0.01	1.15 ± 0.03
at 2°C	8.82 ± 0.02	1.12 ± 0.01
Complex		
at 25°C	10.43 ± 0.03	1.30 ± 0.02
at 15°C	16.67 ± 0.07	1.33 ± 0.03

^{a)} Determined from backbone ¹⁵N R_1 and R_2 relaxation rates for the rigid regions.

^{b)} Parameters for the axially symmetric rotational diffusion model. The effective molecular rotational correlation time $\tau_{r,eff}$ and the anisotropy r are given by $(2 D_{\parallel} + 4 D_{\perp})^{-1}$.

Table 2.4: NMR order parameters S^2_{axis} determined for Arg side-chain N_{ϵ} - H_{ϵ} groups of the Antp homeodomain in the free and DNA-bound states at pH 5.8 and 25°C

Arg N_{ϵ} - H_{ϵ}	S^2 (free)	S^2 (complex)
R1	0.088 ± 0.005	0.147 ± 0.001
R3	0.110 ± 0.004	0.360 ± 0.002
R5	0.109 ± 0.004	0.679 ± 0.009
R10	0.175 ± 0.005	0.126 ± 0.001
R24	0.471 ± 0.025	0.391 ± 0.003
R28	0.223 ± 0.010	0.207 ± 0.001
R29	0.235 ± 0.002	0.204 ± 0.002
R30	0.665 ± 0.030	0.712 ± 0.005
R31	0.751 ± 0.009	0.630 ± 0.012
R43	0.172 ± 0.016	0.347 ± 0.003
R52	0.751 ± 0.031	0.804 ± 0.015
R53	0.731 ± 0.035	0.967 ± 0.015

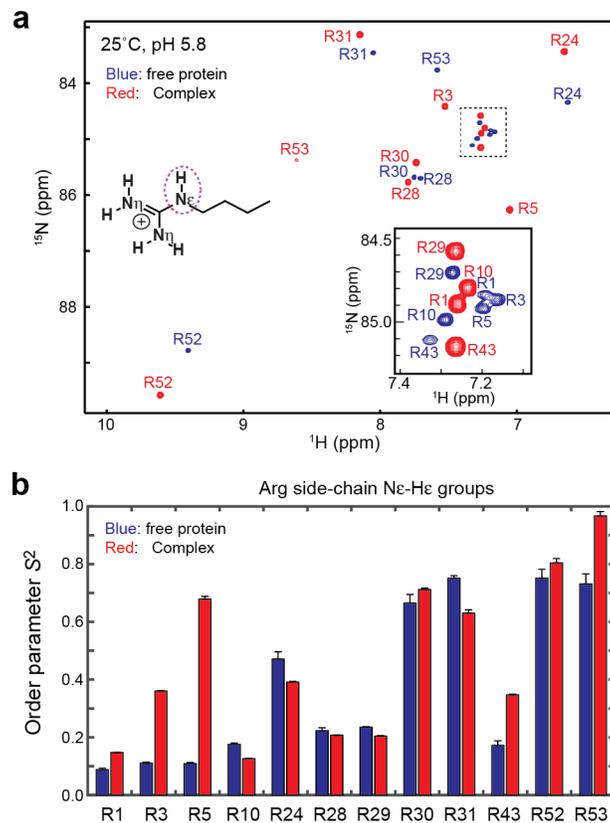


Figure 2.3: NMR investigations on internal motions of Arg side-chain N_ϵ - H_ϵ moieties of the Antp homeodomain in the free and DNA-bound states. **(a)** Overlaid ^1H - ^{15}N HISQC spectra⁶⁴ recorded for the Arg side chains in the free protein (blue) and in the complex with 15-bp DNA (red). The region indicated by a dotted box is expanded in the inset. **(b)** NMR-derived order parameters S^2 for Arg N_ϵ - H_ϵ bonds in the free protein (blue) and in the complex (red).

2.3.2 Change in mobility of Arg side chains upon protein-DNA association

The S^2 data allow us to assess the change in dynamics of the Arg side chains upon DNA-binding. This is straightforward because these Arg order parameters were determined under identical conditions for the free and DNA-bound states. Among the Arg side chains, R5 showed the largest increase in order parameter S^2 , indicating that the mobility of this side chain becomes severely hampered upon the protein-DNA association. This side chain is deeply buried in DNA minor groove (see Figure 2.1b) and makes hydrogen bonds with the

O2 atom of the first thymine base of the recognition sequence TAATGG. The large increase in the S^2 parameter for the R5 N $_{\epsilon}$ -H $_{\epsilon}$ bond can be attributed to immobilization arising from hydrogen-bonding to the thymine base in the narrow space of DNA minor groove. R5 is extremely well conserved among the homeodomain proteins and plays an important role in DNA shape recognition by these proteins.^{93, 94} The side chain of R3, another residue important for DNA shape recognition,⁹³ also exhibited a larger increase in the S^2 parameter. In some crystal structures, the side chains of R3, R28, R31, R43, and R53 form contact ion pairs (CIPs) with DNA phosphates. Despite the short-range electrostatic interactions with DNA, R3, R28, and R43 side chains retain substantial mobility, exhibiting $S^2 < 0.5$. Upon molecular association, the order parameters S^2 of R3, R43, and R53 N $_{\epsilon}$ -H $_{\epsilon}$ bonds were found to increase by ~ 0.2 , suggesting that the ion pair formation with DNA phosphate restricts the internal motions of these Arg side chains. Interestingly, R28 and R31 did not exhibit such immobilization. The S^2 value for the R31 side chain was found to decrease by 0.12, indicating that this side chain becomes more mobile upon protein-DNA association. The R31 side chain in the free state exhibited a relatively large S^2 value (= 0.77), probably due to hydrogen-bonding to E42 O $_{\epsilon}$ as seen in some crystal structures of free homeodomains (e.g., PDB 1ENH and 1P7I). This interaction is broken due to attraction of the R31 guanidino cation to a DNA phosphate in the crystal structures of the Antp homeodomain-DNA complexes. Competition between E42 carboxylate and DNA phosphate for the R31 guanidino cation might mobilize the R31 side chain in the complex. The retained or enhanced mobility of Arg side chains upon protein-DNA association would be entropically favorable for binding.

2.3.3 Internal motions of Lys side-chain NH_3^+ groups

We previously determined the order parameters of the Lys side-chain NH_3^+ moieties in the Antp homeodomain-DNA complexes at pH 5.8 and 25°C.^{24, 70} Under these conditions, the side-chain NH_3^+ moieties of the interfacial Lys residues, K46, K55, K57, and K58, in the complexes exhibited signals in ^1H - ^{15}N HISQC spectra. In our previous studies, intermolecular CIPs were confirmed for K46, K55, and K57 by observation of the hydrogen-bond scalar couplings ($^hJ_{NP}$) between Lys side-chain ^{15}N and DNA phosphate ^{31}P nuclei.^{24, 70, 74} Although it would be ideal to compare the internal motions of the free and DNA-bound states under the identical conditions, unfortunately, the Antp homeodomain in the free state showed no observable signals from Lys NH_3^+ groups at pH 5.8 and 25°C due to rapid hydrogen exchange. However, the free Antp homeodomain at pH 4.5 and 2°C exhibited ^1H - ^{15}N HISQC signals from all Lys NH_3^+ moieties (Figure 2.4a). Backbone ^1H - ^{15}N TROSY spectra show that the protein remains folded under these conditions (Figure 2.5). As previously described,^{65, 69} we measured ^{15}N relaxation at the ^1H -frequencies of 800 and 600 MHz, and determined order parameters of Lys side-chain NH_3^+ moieties of the free protein under these conditions. Because the complex was found to partially aggregate at the low pH and low temperature (data not shown), we conducted the Lys side-chain ^{15}N relaxation analysis for the complex at pH 5.8 and 15°C, in addition to our previous analysis at 25°C.²⁴ The ^{15}N relaxation parameters measured for the free protein and the complex in the current study are shown in Tables 2.5 and 2.6. Figure 2.4b shows the order parameters (S^2_{axis}) determined for Lys side-chain NH_3^+ moieties in the free protein and in the complex (values are shown in Table 2.7).

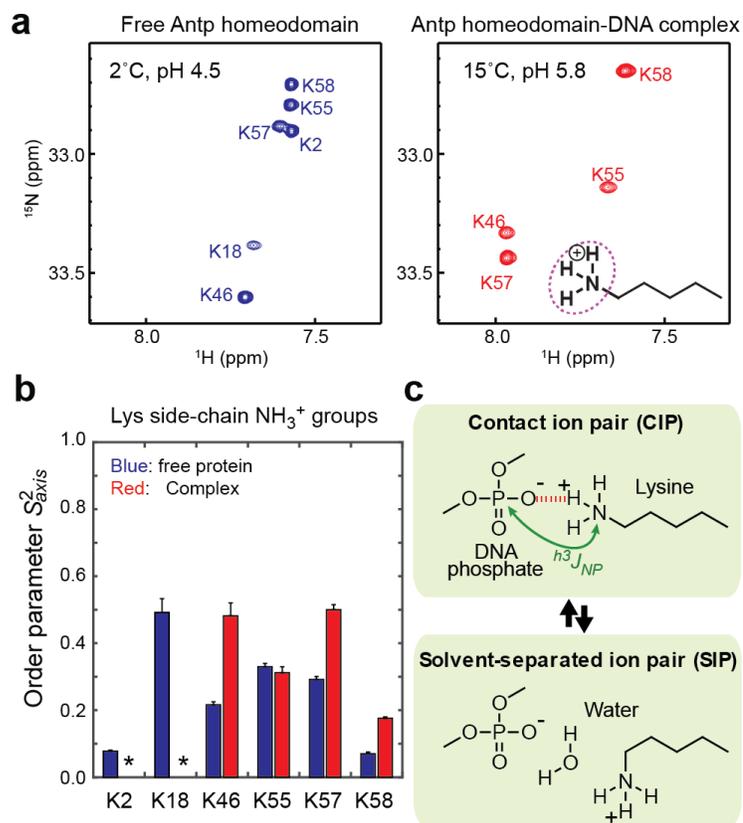


Figure 2.4: NMR investigations on internal motions of Lys side-chain NH_3^+ moieties of the Antp homeodomain in the free and DNA-bound states. **(a)** Lys NH_3^+ -selective ^1H - ^{15}N HSQC spectra⁶⁴ recorded for the free protein (blue) and the complex with 15-bp DNA (red). **(b)** NMR-derived order parameters S^2_{axis} for Lys side chains in the free protein (blue) and in the complex (red). **(c)** Dynamic equilibrium between the CIP and SIP states.

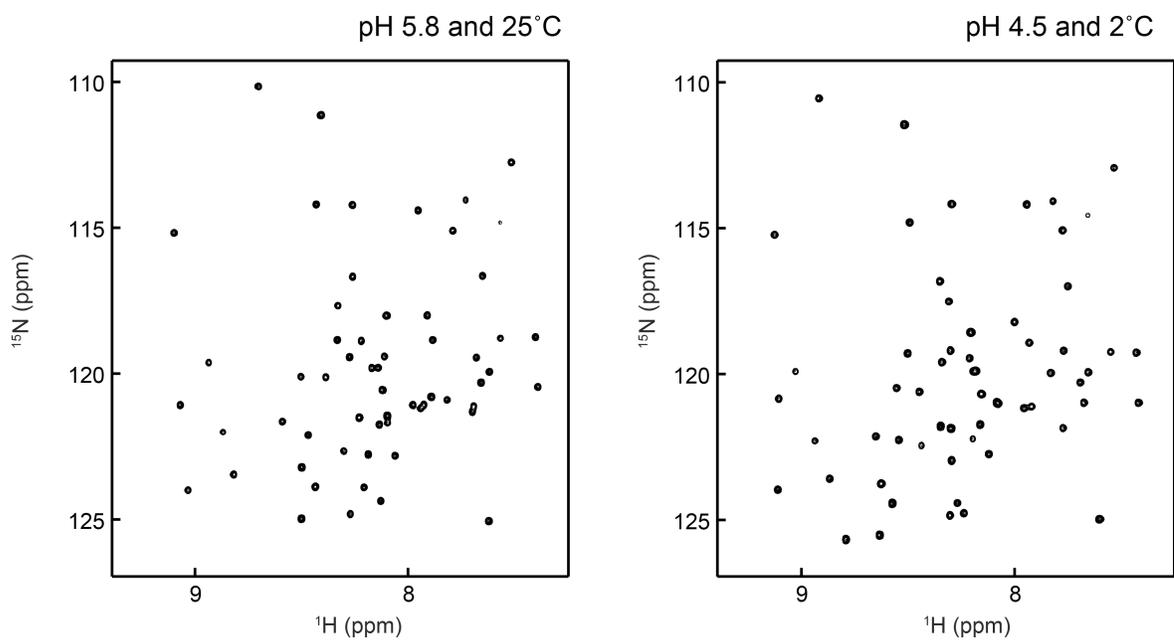


Figure 2.5: Data indicating that the Antp homeodomain at pH 4.5 and 2°C remains folded. Shown here are ^1H - ^{15}N TROSY spectra recorded for the backbone NH groups of the free Antp homeodomain at pH 5.8 and 25°C (left) and at pH 4.5 and 2°C (right).

Table 2.5: ^{15}N NMR relaxation data for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups of the Antp homeodomain in the free state at pH 4.5 and 2°C .

Lys NH_3^+	$^{15}\text{N}_\zeta R_1$ (s^{-1})	$^{15}\text{N}_\zeta R_1$ (s^{-1})	$^{15}\text{N}_\zeta R_{2,\text{ini}}$ (s^{-1}) ^{a)}	$\{^1\text{H}\}\text{N}_\zeta \text{NOE}$	$\{^1\text{H}\}\text{N}_\zeta \text{NOE}$
	600 MHz	800 MHz	800 MHz	600 MHz	800 MHz
K2	0.400 ± 0.001	0.296 ± 0.001	0.53 ± 0.01	-2.86 ± 0.02	-3.00 ± 0.02
K18	0.701 ± 0.016	0.580 ± 0.016	1.92 ± 0.13	-2.90 ± 0.13	-2.86 ± 0.12
K46	0.594 ± 0.002	0.514 ± 0.002	1.10 ± 0.02	-2.77 ± 0.03	-2.59 ± 0.03
K55	0.741 ± 0.002	0.630 ± 0.002	1.47 ± 0.03 ^{b)}	-3.06 ± 0.03	-2.67 ± 0.03
K57	0.473 ± 0.003	0.380 ± 0.003	1.18 ± 0.03	-2.88 ± 0.05	-2.72 ± 0.04
K58	0.425 ± 0.001	0.390 ± 0.001	0.58 ± 0.01	-2.97 ± 0.02	-2.55 ± 0.02

^{a)} The initial rate for intrinsically biexponential ^{15}N transverse relaxation of NH_3^+ .

^{b)} Corrected based on the ^{15}N R_2 relaxation dispersion data for the NH_3^+ groups. The ^{15}N R_2 relaxation dispersion experiment for Lys side chains showed a significant contribution from motions on a μs – ms timescale only for this residue.

Table 2.6: ^{15}N NMR relaxation data for Arg side-chain $\text{N}_\epsilon\text{-H}_\epsilon$ groups of the Antp homeodomain in the specific complex with 15-bp target DNA at pH 5.8 and 2°C .

Lys NH_3^+	$^{15}\text{N}_\zeta R_1$ (s^{-1})	$^{15}\text{N}_\zeta R_1$ (s^{-1})	$^{15}\text{N}_\zeta R_{2,\text{ini}}$ (s^{-1}) ^{a)}	$\{^1\text{H}\}\text{N}_\zeta \text{NOE}$	$\{^1\text{H}\}\text{N}_\zeta \text{NOE}$
	600 MHz	800 MHz	800 MHz	600 MHz	800 MHz
K2	n.d. ^{b)}	n.d. ^{b)}	n.d. ^{b)}	n.d. ^{b)}	n.d. ^{b)}
K18	n.d. ^{b)}	n.d. ^{b)}	n.d. ^{b)}	n.d. ^{b)}	n.d. ^{b)}
K46	1.367 ± 0.021	1.156 ± 0.011	3.75 ± 0.06	-2.53 ± 0.22	-2.29 ± 0.05
K55	0.730 ± 0.009	0.590 ± 0.006	2.21 ± 0.09	-2.83 ± 0.23	-2.81 ± 0.05
K57	1.255 ± 0.013	1.054 ± 0.004	3.68 ± 0.04	-2.83 ± 0.16	-2.60 ± 0.03
K58	0.435 ± 0.002	0.417 ± 0.001	1.31 ± 0.02	-3.02 ± 0.11	-2.68 ± 0.02

^{a)} The initial rate for intrinsically biexponential ^{15}N transverse relaxation of NH_3^+ .

^{b)} No data available due to rapid hydrogen exchange with water.

Table 2.7: NMR order parameters S^2_{axis} determined for Lys NH_3^+ groups of the Antp homeodomain in the free and DNA-bound states. ^{a)}

Lys NH_3^+	$S^2_{\text{axis}}(\text{free})$ ^{b)}	$S^2_{\text{axis}}(\text{complex})$ ^{c)}
K2	0.078 ± 0.003	n.d.
K18	0.492 ± 0.041	n.d.
K46	0.216 ± 0.009	0.482 ± 0.038
K55	0.330 ± 0.010	0.312 ± 0.018
K57	0.292 ± 0.009	0.500 ± 0.015
K58	0.070 ± 0.005	0.176 ± 0.004

^{a)} The S^2_{axis} parameters are the order parameters for the symmetry axis (i.e., the $\text{C}_\alpha\text{-N}_\zeta$ bond) of the tetrahedral NH_3^+ groups

^{b)} Determined at pH 4.5 and 2°C .

^{c)} Determined at pH 5.8 and 15°C .

2.3.4 Different motional characteristics of Arg and Lys side chains

The order parameters S^2_{axis} of Lys NH_3^+ groups are defined for the C_3 symmetry axis, which corresponds to the $C_\epsilon\text{-N}_\zeta$ bond. Although the same number of rotatable bonds (i.e., four covalent bonds) are involved between Lys C_α and C_ϵ atoms and between Arg C_α and N_ϵ atoms, the ranges of the measured order parameters were remarkably different for Lys $C_\epsilon\text{-N}_\zeta$ and Arg $N_\epsilon\text{-H}_\epsilon$ bonds. Compared to Lys $C_\epsilon\text{-N}_\zeta$ bonds, Arg $N_\epsilon\text{-H}_\epsilon$ bonds exhibited a wider range of order parameters (see Figures 2.3b vs. 2.4b) as well as larger changes in order parameters upon DNA binding (Figure 2.6). The same characteristic differences between Arg and Lys side chains were also seen in our previous study on the Egr-1 zinc-finger protein in the free and DNA-bound states.³⁷ Judging from our current and previous studies, the mobility of Arg side chains seems to be more sensitive to the surrounding environment, compared to Lys side chains.

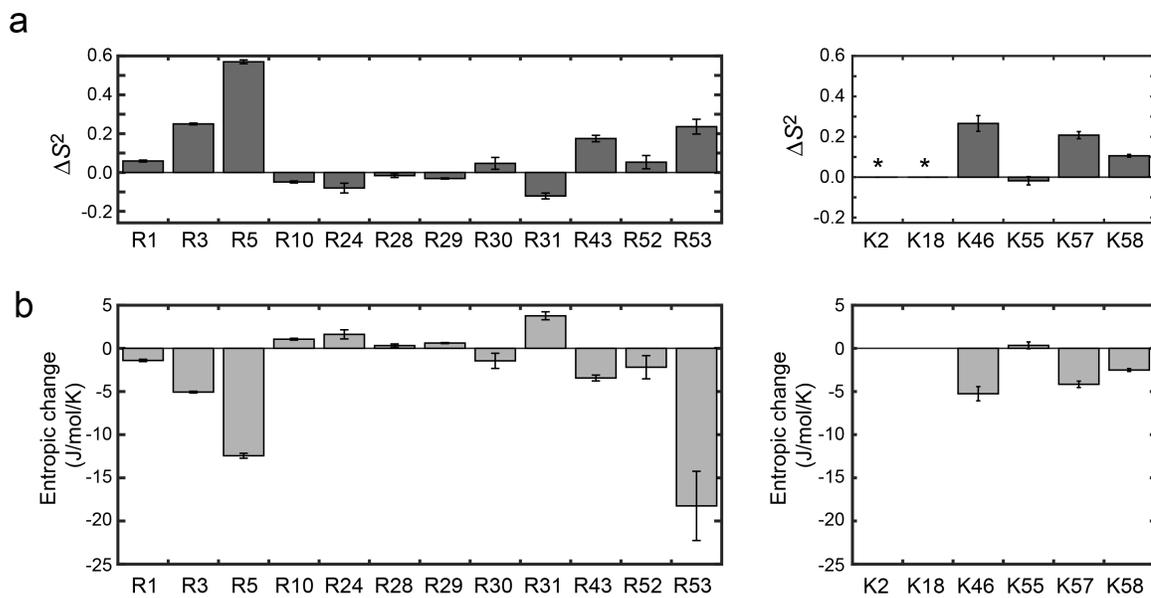


Figure 2.6: Changes in order parameter and entropy for the Arg/Lys cationic groups of the Antp homeodomain upon binding to DNA. **(a)** Difference between the order parameters S^2 for the free and DNA-bound states. $\Delta S^2 = S^2(\text{complex}) - S^2(\text{free})$. Individual values of the S^2 parameters are reported in Tables S4 and S7. **(b)** Entropic changes due to changes in mobility upon the protein-DNA association. The values were determined from $S^2(\text{complex})$ and $S^2(\text{free})$ using Eq. 24 in Yang and Kay.⁹⁵ The overall entropic change for the Arg/Lys cationic groups was calculated to be -48 ± 5 J/mol/K.

2.3.5 Retained mobility of basic side chains forming ion pairs with DNA phosphates

Our previous study on the Egr-1 zinc-finger protein in the free and DNA-bound states showed that the basic side chains forming ion pairs with DNA phosphates tend to retain substantial mobility.³⁷ This trend was confirmed in the current study on the Antp homeodomain. The side chains R3, R28, R43, K46, K55, and K57 in the DNA-bound state exhibited order parameters S^2 between 0.19 and 0.55, indicating substantial mobility, although they form intermolecular ion pairs with DNA. Although our current study suggests that the overall entropic change is negative for the Arg/Lys cationic moieties (see Figure 2.6b), their high mobility in the complex should mitigate the entropic loss. The high mobility of the interfacial basic side chains may also help proteins adaptively recognize DNA, even while the DNA undergoes significant conformational fluctuations such as B_I–B_{II} transitions.

Despite the simultaneous presence of the hydrogen bonds and strong short-range electrostatic interactions, why are these basic side chains so mobile? As we previously discussed based on theoretical and computational investigations,^{17, 24, 37} this high mobility can be attributed to the dynamic equilibria between the CIP and solvent-separated ion-pair (SIP) states (Figure 2.4c). The CIP and SIP states can be distinguished in terms of N...O distance between cationic and anionic groups.²⁴ For Lys side-chain NH₃⁺–DNA phosphate ion pairs, the free energy differences between CIP and SIP states were estimated to be 0.8–1.6 kcal/mol and the energy barriers for CIP→SIP transitions were estimated to be 2.2–3.2 kcal/mol.²⁴ With the relatively small energy difference and barrier, transitions between the CIP and SIP states rapidly occur, making the ion pairs highly dynamic in a ps–ns timescale. Interestingly, a computational study showed that the free energy difference and barriers

between CIP and SIP states of an Arg⁺–Glu⁻ pair are significantly larger than those of a Lys⁺–Glu⁻ pair,⁹⁶ which seems to be qualitatively consistent with our experimental observation that Lys side chains tends to be more mobile than Arg side chains. It should also be noted that Lys and Arg side chains differ not only in number of hydrogen bonds, but also in desolvation energy, as discussed by Rohs et al.⁹⁷ These differences may account for the higher mobility of Lys side chains.

2.4 DISCUSSION

Despite strong short-range electrostatic interactions and hydrogen bonds, the S^2 order parameters determined from our experiments suggests that the majority of the Lys and Arg side chains forming an ion pair with the DNA phosphate backbone displayed a retained mobility. This retained mobility was observed for both the Antp homeodomain-DNA complex as well as the Egr-1 Zinc finger complex³⁷. Upon further observation, ion-pairs involving Lys side chains showed a tendency to be more mobile when compared to ion-pairs involving Arg side chains. Our previous computational and theoretical work indicated that the retained mobility is a result of rapid transitions between the CIP-SIP equilibria states, and the difference in dynamic properties of the basic side chains could be a result of intrinsic differences in their structures. These observations were quite fascinating and infer that a set of general trends could be observed for ion-pairs of protein-DNA complexes. Although these studies also implied that the retained side chain mobility could mitigate the overall entropic loss of protein-DNA binding, further studies should be done to assess whether ion-pair dynamics can reduce the entropic cost for binding and how. We address these questions in Chapter 3 of this dissertation work.

CHAPTER 3

Using the Sulfur Substitution in Ion Pairs to Improve Understanding of Ion Pair Dynamics⁴

3.1 INTRODUCTION

Chemically modified oligonucleotides have drawn considerable interest as potential therapeutic reagents⁹⁸⁻¹⁰⁰. Monothioate and dithioate-derivatives of phosphate are commonly used in oligonucleotides designed for potential therapeutic applications.¹⁰¹⁻¹⁰⁷ These derivatives, in which non-bridging oxygen is replaced with sulfur, retain the overall charge and similar tetrahedral covalent geometry of DNA phosphate, but increase resistance to nucleases and improve cell penetration properties.^{51, 108, 109} It was recently discovered that phosphorothioate is also naturally present in some bacterial genomes^{110, 111}. The role of phosphoromonothioate *in vivo* remains to be elucidated, though a potential role as an antioxidant was suggested.¹¹² Interestingly, compared to unmodified DNA, thioated DNA often exhibits stronger binding affinity, at least for some proteins and in some positions^{36, 38, 39, 41, 104}. Due to these properties, short DNA duplexes containing phosphoromonothioate or phosphorodithioate groups can effectively serve as decoy molecules that inhibit particular transcription factors involved in pathogenesis¹¹³⁻¹¹⁵.

⁴ Chapter adapted with permission from John Wiley and Sons (see Appendix). Nguyen, D., et al. (2016) Stereospecific effects of oxygen-to-sulfur substitution in DNA phosphate on ion-pair dynamics and protein-DNA affinity. *ChemBioChem* 17, 1636-42. And with permission from Elsevier (see Appendix). Zandarashvili, L. and Nguyen, D., et al. (2015) Entropic enhancement of protein-DNA affinity by oxygen-to-sulfur substitution in DNA phosphate. *Biophys J* 109, 1026-37.

From a physicochemical point of view, however, the protein-DNA affinity enhancement by the oxygen-to-sulfur substitution in DNA phosphate may appear counterintuitive, especially given the following two facts. First, sulfur atoms in organic compounds tend to serve as relatively poor hydrogen bond acceptors compared with oxygen atoms. For example, the boiling point of methanethiol (CH_3SH) is lower than that of methanol (CH_3OH) by 59°C ; and the boiling point of 2-mercaptoethanol ($\text{HO-CH}_2\text{-CH}_2\text{-SH}$) is lower than that of ethylene glycol ($\text{HO-CH}_2\text{-CH}_2\text{-OH}$) by 39°C ¹¹⁶. Second, the electronegativity of the sulfur atom is weaker than that of the oxygen atom (2.58 vs. 3.44 by Pauling scale)¹¹⁶.

Recently, we gained important insight into this question. In our previous NMR studies of the HoxD9 homeodomain-DNA complexes⁶⁸, we found that the mobility of the Lys sidechain NH_3^+ group is enhanced upon the oxygen-to-sulfur substitution of the DNA phosphate group, which forms an intermolecular ion pair at the molecular interface. The entropic impact of this mobilization on the binding free energy was estimated from the changes in NMR order parameters and bond-rotation correlation times. The data suggested that the mobilization of the intermolecular ion pair can at least partially account for affinity enhancement by the oxygen-to-sulfur substitution of DNA phosphate. However, this was indecisive because neither structural detail around the ion pairs nor thermodynamic (i.e., enthalpic and entropic) data on binding was available for the HoxD9-DNA complexes.

In the current work, we resolve this issue and further examine the role of ion-pair dynamics in affinity enhancement by the oxygen-to-sulfur substitution of DNA phosphate. For this purpose, using fluorescence spectroscopy, isothermal titration calorimetry (ITC),

NMR spectroscopy, and x-ray crystallography, we characterize the sequence specific interactions of the Antennapedia (Antp) homeodomain with unmodified and dithioated DNA. The Antp homeodomain is practically more useful than the HoxD9 homeodomain due to higher solubility and stability of the free state under physiological conditions. Furthermore, in previous studies by other research groups, the Antp homeodomain has been extensively characterized by biochemical methods^{47, 117} as well as by biophysical methods such as ITC¹², NMR^{73, 118, 119}, and x-ray crystallography⁵⁰. Using this well-suited system, we investigate how the oxygen-to sulfur substitution in a DNA phosphate group influences protein-DNA association in terms of thermodynamics, internal motions, and structure.

To gain more mechanistic insight, we also investigated the impacts of stereospecific sulfur substitutions for single non-bridging oxygen atoms in the phosphate group at the Lys57 interaction site of the Antp homeodomain–DNA complex. Although the monothioation involves RP and SP diastereomers, they could be readily isolated by chromatography. For each diastereomer, we examined the impacts of the oxygen-to-sulfur substitution on binding affinity, structure, and dynamics by using fluorescence-based assays, NMR spectroscopy, and X-ray crystallography. The results from these investigations further support the relationship between ion pair dynamics and binding affinity.

3.2 MATERIALS AND METHODS

3.2.1 Protein Preparation

A synthetic gene encoding the 60 amino-acid residues of the fruit fly Antp homeodomain (RKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALSLTERQIKIWFQNRRMKWKKEN) with C39S mutation⁵⁰ was sub-cloned into the NdeI/ HindIII sites of the pET-49b vector (Novagen). *E. coli* strain BL21(DE3) was transformed with this plasmid and cultured at 37°C in 4 L of M9 minimal media containing kanamycin (30 µg/ml) as well as ammonium chloride and glucose as the sole nitrogen and carbon sources, respectively. At $OD_{600} \approx 0.8$, protein expression was induced by adding isopropyl β-D-thiogalactopyranoside (0.4 mM) to the media, and the *E. coli* culture was continued at 18 °C for additional 16 hours. The cells were harvested and disrupted by sonication in a buffer containing 20mM Tris•HCl (pH 7.5), 1mM EDTA, 500mM NaCl, 2mM DTT, 5% glycerol and a protease inhibitor cocktail (Roche). The supernatant of the lysate was loaded on a SP-FF column (GE Healthcare) equilibrated with 50mM Phosphate buffer (pH 7.5) and 500mM NaCl, and the protein was eluted using the gradient of 500-1500 mM NaCl. Fractions containing the Antp homeodomain were pooled and concentrated to ~10 ml and then loaded onto a S100 size-exclusion column (GE Healthcare) equilibrated with a buffer of 50mM Tris•HCl (pH 7.5), 1mM EDTA, and 400mM NaCl. The protein solution was loaded onto a Resource-S cation-exchange column (GE Healthcare) equilibrated with a buffer of 50mM Tris•HCl (pH 7.5), 1mM EDTA, and 600mM NaCl, and then eluted with gradient of 600-1400 mM NaCl. Purity of the protein was confirmed to be >95%. The protein was quantified using UV absorbance at 280 nm together with an extinction coefficient of 15,470 M⁻¹ cm⁻¹.

3.2.2 DNA Preparation

The sequences of DNA strands used in this work are shown in the figures together with the data. The DNA strand containing a phosphorodithioate group was synthesized on an ABI Expedite 8909 DNA synthesizer with standard dA/dC/dG/dT-phosphoramidites and dC-thiophosphoramidite (AM Biotechnologies / Glen Research), and purified via anion-exchange chromatography using a Mono-Q column as described⁵⁴. All the other DNA strands were purchased from Integrated DNA Technologies, Inc. and purified by anion-exchange chromatography. For preparation of double-stranded DNA, complementary strands were annealed and minor single-stranded DNA excess due to the uncertainty in measuring single strand concentrations was removed by anion-exchange chromatography.

3.2.3 Preparation of R_p and S_p diastereomers of monothioated DNA

The DNA strands used in this study are shown in Figure 3.2A. Unmodified DNA strands were purchased from Integrated DNA Technologies. DNA strands containing a phosphoromonothioate group at the Lys57 interaction site were purchased from Sigma Aldrich. The incorporation of the phosphoromonothioate group, which is chiral at the phosphorous atom, was not stereospecific in the chemical synthesis and produced a mixture of the R_P and S_P diastereomers. Single-stranded DNA was purified via Mono-Q anion-exchange chromatography. Electrospray mass spectrometry measurement for the monothioated DNA strand was conducted by the Integrated DNA Technologies. Double-stranded DNA was formed by mixing equimolar amounts of the complementary stands followed by annealing from ~ 85 °C to room temperature over 1-2 hours. The DNA duplexes of the R_P and S_P diastereomers were isolated through the following two runs of Mono-Q anion-exchange chromatography. In the first run, minor amounts of excess single

stranded DNA due to uncertainties in the concentrations of individual strands were removed using a buffer of 50 mM Tris•HCl (pH 7.5) and 1 mM EDTA in a gradient of 0-675 mM NaCl. Elution fractions containing the DNA duplexes of R_P and S_P species were separately collected and subjected to the second run for further isolation via a gentler gradient of 480-510 mM NaCl over 100 ml.

3.2.4 Fluorescence-based affinity measurements

The affinities of the Antp homeodomain for the unmodified and dithioated target DNA duplexes were determined using two different methods. One is based on fluorescence anisotropy as a function of protein concentration (0.1 – 500 nM). Fluorescence arising from tetramethylrhodamine (TAMRA) attached to the 3'-terminus of DNA (3.3 nM) was measured using an ISS PC-1 spectrofluorometer. Excitation and emission wavelengths used were 533 and 580 nm, respectively. The titration experiments were performed at 25 °C using a buffer of 10 mM sodium phosphate (pH 5.8) and 150 mM NaCl. The dissociation constant K_d was calculated from the anisotropy data via nonlinear least-squares fitting with:

$$A_{obs} = A_{free} + (A_{bound} - A_{free}) \left(\frac{P + D + K_d - \sqrt{(P + D + K_d)^2 - 4PD}}{2D} \right) \quad [1],$$

where A_{obs} is the observed anisotropy; A_{bound} and A_{free} are those of protein-bound DNA and free DNA; and P and D are total concentrations of the protein and the probe DNA, respectively. For each DNA, the affinity measurements were repeated three times.

We also measured K_d values using a fluorescence-based competition assay. In this assay, solutions of 10nM 3'-TAMRA-labeled target DNA, 50nM Antp homeodomain, and unlabeled competitor DNA (0.5 – 6400 nM) were made and fluorescence anisotropy was

measured as a function of the competitor concentration. We used the following equation for analysis of the competition assay data:

$$r = \frac{CK_{d,p} + K_{d,p}K_{d,c} - PK_{d,p} + 2PK_{d,c} - K_{d,p}\sqrt{(C + K_{d,c} - P)^2 + 4PK_{d,c}}}{2\{CK_{d,p} - (K_{d,p} - K_{d,c})(K_{d,p} + P)\}} \quad [2],$$

$$A_{obs} = (1-r)A_{free} + rA_{bound} \quad [3],$$

where $K_{d,c}$ and $K_{d,p}$ are the dissociation constants for the competitor and probe DNA duplexes, respectively; and C is the concentration of competitor DNA. When a value of either $K_{d,c}$ or $K_{d,p}$ is known, the other dissociation constant can be determined from the competition assay data via nonlinear least-squares fitting. Eq. 2 assumes $D \ll P$. When $P \ll C$ is also satisfied (e.g., for the data for the nonspecific competitor DNA), Eq. 2 becomes equivalent to a popular form based on binding polynomial¹²⁰:

$$r = \frac{P/K_{d,p}}{1 + C/K_{d,c} + P/K_{d,p}} \quad [4]$$

The dissociation constant K_d was calculated from the fluorescence anisotropy data via nonlinear least-squares fitting using MATLAB (MathWorks, Inc.). The affinities of the Antp homeodomain for the R_P and S_P diastereomers were also determined with competitive binding assays using fluorescence anisotropy as a function of unlabeled competitor DNA. The measurements were carried out four times for each monothioated DNA.

3.2.5 ITC measurements

Using a MicroCal VP-ITC microcalorimeter, the ITC experiments were carried out at 25 °C for three different DNA duplexes. Two of these DNA duplexes include the Antp target sequence and one of them is dithioated at the Lys57 interaction site. The third DNA duplex is nonspecific 15-bp DNA. Each solution of DNA or protein was extensively dialyzed to a buffer of 10 mM sodium phosphate (pH 5.8) and 150 mM NaCl. Each titration experiment consists of one 5 µl and twenty 12 µl injections of Antp homeodomain into the cell, which initially contained a 1.41 ml solution of DNA. Concentrations of the molecular components are given in the figure legends. The interval between the injections was 4 minutes. For both unmodified and dithioated DNA, the ITC experiment was repeated three times. To account for the heat of protein dilution, a control experiment was also performed using titration of the protein into the buffer under the identical experimental conditions. Heat of the control titration was subtracted from the original titration data. To avoid influence of heat from nonspecific association, data with relatively low concentrations of the Antp homeodomain with DNA in excess were used to measure the binding enthalpies for the specific complex formation. The ITC data was analyzed with the Origin 7.0 software. The affinity data from the fluorescence experiment together with the ITC data was used to determine binding entropy as described previously¹².

3.2.6 NMR Spectroscopy for the unmodified and dithioated protein-DNA complexes

All NMR experiments for the Antp homeodomain-DNA complexes were performed using Bruker Avance III spectrometers operated at a ¹H frequency of 600, 750, or 800 MHz. A 280 µl solution of 0.8 mM Antp homeodomain (¹⁵N- or ¹³C/¹⁵N-labeled) and 1.2 mM DNA in a buffer of 10 mM sodium phosphate (pH 5.8) and 20 mM NaCl was sealed into

the inner tube (outer diameter 4.1 mm) of a 5-mm co-axial NMR tube (Shigemi). D₂O for NMR lock was separately sealed in the outer layer of the co-axial tube to avoid deuteration of NH₃⁺ groups (i.e., NDH₂⁺ and ND₂H⁺ species)⁴². Backbone ¹H/¹³C/¹⁵N resonances were assigned using 2D HSQC spectra and 3D HNC(O), HN(CA)CO, HNCA, HN(CO)CA, CBCA(CO)NH, HNCACB, and HBHA(CO)NH spectra¹²¹. Side-chain ¹H/¹³C resonances were assigned using 3D H(CCO)NH, C(CO)NH, HCCH-TOCSY, and HCCH-COSY spectra¹²¹. Lys side-chain NH₃⁺ ¹H/¹⁵N resonances were assigned using Lys-selective 2D HISQC⁶⁴, (H₂C)N(CC)H-TOCSY⁵⁶, and H₂(C)N¹²² spectra and 3D H₃NCECD⁶⁴, HDHE(CDCE)NH₃⁶⁴, and H₃NCG⁵⁶ spectra. Scalar coupling between lysine side-chain ¹⁵N and DNA phosphate/phosphorodithioate ³¹P nuclei across hydrogen bonds (^{*h*3}*J*_{NP}) was analyzed by 2D H₃(N)P and spin-echo ^{*h*3}*J*_{NP}-modulation difference constant-time HISQC experiments with a Bruker cryogenic QCI ¹H/¹³C/¹⁵N/³¹P probe at the ¹H-frequency of 600 MHz as described previously³⁶. Backbone ¹⁵N longitudinal and transverse relaxation rates (denoted by R₁ and R₂, respectively) were measured at 25 °C, from which the molecular rotational correlation time and anisotropy of rotational diffusion of the Antp homeodomain-DNA complexes were determined. For Lys side-chain NH₃⁺ groups, ¹⁵N R₁ and R₂ relaxation rates, and heteronuclear NOE were measured at 25 °C as described previously³⁴. For the NH₃⁺ groups, the generalized order parameters *S*²_{axis} and the reorientation correlation time τ_{*i*} for the symmetry axis, and the bond-rotation correlation time τ_{*r*} were determined using Mathematica as described previously^{34,55}.

3.2.7 NMR Spectroscopy for the Protein-DNA complexes of the R_p and S_p diastereomers of monothioate

Again, all NMR experiments for the Antp homeodomain-DNA complexes were performed with Bruker Avance III spectrometers operated at a ^1H frequency of either 600 or 800 MHz. Each experiment was performed at 25 °C using a cryogenic probe. A 280 μl solution of 0.8 mM ^{15}N -Antp homeodomain and 1.2 mM monothioated DNA in a buffer of 10 mM sodium phosphate (pH 5.8) and 20 mM NaCl was sealed into the inner tube (outer diameter 4.1 mm) of a 5-mm co-axial NMR tube (Shigemi). D_2O for the NMR lock was separately sealed in the outer layer of the co-axial tube to avoid deuteration of NH_3^+ groups (i.e., NDH_2^+ and ND_2H^+ species). Protein backbone and side-chain ^1H , ^{13}C , and ^{15}N resonances of the complex were assigned using the unmodified and dithioated protein-DNA complexes.^{24, 70} Backbone ^{15}N longitudinal and transverse relaxation rates (R_1 and R_2 , respectively) were measured at 25 °C, from which the molecular rotational diffusion parameters for each complex were determined. ^{15}N R_1 and heteronuclear NOE for the Lys side-chain NH_3^+ groups of the complexes were measured at the ^1H frequencies of 600 and 800 MHz, as described.^{42, 65} ^{15}N R_2 rates for the Lys side-chain NH_3^+ groups of the complexes were measured at the ^1H frequency of 800 MHz, as described.^{42, 65} All NMR data were processed with NMRPipe program¹²³ and analyzed with NMRView program¹²⁴. For each Lys NH_3^+ group, the generalized order parameter S^2_{axis} , the rotational correlation time for the symmetry axis τ_i , and the $\text{C}\epsilon\text{-N}\zeta$ bond-rotation correlation time τ_f were determined using Mathematica, as previously described.^{42, 65} The NMR pulse programs and analysis scripts for investigating the dynamics of Lys side-chain NH_3^+ groups are available at www.scsb.utmb.edu/labgroups/iwahara/software.

3.2.8 X-ray Crystallography for the unmodified and dithioated protein-DNA complexes

The Antp homeodomain-DNA complexes with and without dithioation for DNA phosphate at the Lys57 interaction sites were crystallized using the conditions described by Fraenkel and Pabo⁵⁰. Crystals were grown at 17 °C over several days using the sitting-drop vapor diffusion method. For the unmodified DNA complex, X-ray diffraction data were collected on a Rigaku FR-E++DW with an R-AXIS-IV++ image plate detector using Cu radiation from a single crystal. Two hundred 1/2-degree width frames were collected. The diffraction images were processed and scaled using HKL3000, using an I/σ of 1.0 resolution cutoff criteria. The data for the dithioated DNA complex were collected at the APS beamline 19ID at a wavelength of 0.97921 Å using a Quantum 315 detector. In order to collect complete data without overloads two data sets were collected at different resolutions, exposure times, frame widths (1° and 1/2°), and beam attenuation, from the same crystal. Images were processed and scaled using HKL2000, using the $CC^* > 0.5$ criteria for resolution cutoff. Refinement was performed using Phenix¹²⁵, with TLSMD¹²⁶ determined TLS parameters, weight optimization, and DNA restraints. The crystallographic phase was determined using the molecular replacement method. Model building and validation was performed in Coot¹²⁷. The atomic coordinates of the crystal structures of the unmodified and dithioated complexes have been deposited to Protein Data Bank (PDB accession codes **4XID** and **4XIC**, respectively).

3.2.9 X-ray Crystallography for the protein-DNA complexes of the R_p and S_p diastereomers of monothioate

The Antp homeodomain bound to either the R_p or S_p -monothioated DNA was crystalized using the sitting-drop vapor diffusion method under the same conditions as mentioned above for the native and dithioated DNA complexes. For each of the two protein-DNA complexes, a solution containing ~1.8 mM DNA and 0.8 mM protein in a buffer of 5mM bis-Tris propane was made. For the reservoir buffers, 20 mM bis-Tris propane (pH 6.0-7.5), 10 mM NiCl_2 , and 4-6% 2-methyl-2,4-pentandiol (MPD) were used. The concentrations of the complexes and MPD and the pH were varied to optimize the crystallization. For each well, 1 μl of sample was mixed with 1 μl of reservoir buffer. Crystals were grown at 19°C over several days. For the R_p -monothioated DNA complex, X-ray diffraction data were collected on a Bruker TXS Cu source with a D8 goniometer and Photon 100 detector. Twelve hundred and twenty five $\frac{1}{4}^\circ$ -wide frames were collected from a single crystal, in two omega sweeps with different crystal orientations. The diffraction images were processed and scaled using Proteum/SAINT, with an I/σ of 1.8 resolution cutoff criteria. The data for the S_p monothioated DNA complex were collected at the APS beamline 21-ID-F at a wavelength of 0.97921 Å using a Rayonix MX225 detector. A full hemisphere of data was collected with $\frac{1}{2}^\circ$ -frames from a single crystal. Images were processed and scaled using HKL2000 to 2.75 Å, using the $\text{CC}^* > 0.5$ criteria for resolution cutoff. The crystallographic phase was determined by the molecular replacement method using PDB model **4XID** in Phaser¹²⁸. Refinement was performed using Phenix¹²⁵, with TLSMD¹²⁶ determined TLS parameters, weight optimization, and DNA restraints. Model building and validation were performed in Coot¹²⁷. The covalent parameters from the quantum chemical calculations by Florián et al.¹²⁹ were used for the

phosphoromonothioate. The atomic coordinates of the crystal structures of the R_P and S_P -monothioated complexes were deposited to the Protein Data Bank (PDB accession codes **5JLW** and **5JLX**, respectively).

3.3 RESULTS

We compare the two DNA complexes of the Antp homeodomain with 15-bp DNA duplexes with identical base sequences: one with no chemical modification, the other with phosphorodithioate at the Lys57 interaction site (i.e., two sulfur atoms at the positions of the O_{P1} and O_{P2} atoms) (Figure 3.1a). We conducted the same set of biophysical experiments for these two complexes to study the impacts of the oxygen-to-sulfur substitution on protein-DNA association. We also use two diastereomers (R_P and S_P) of the monothioated phosphate. In the R_P phosphoromonothioate, the O_{P2} atom is replaced with sulfur, whereas in the S_P phosphoromonothioate, the O_{P1} atom is replaced. The Lys57 side-chain NH_3^+ group is more proximal to the O_{P2} atom.

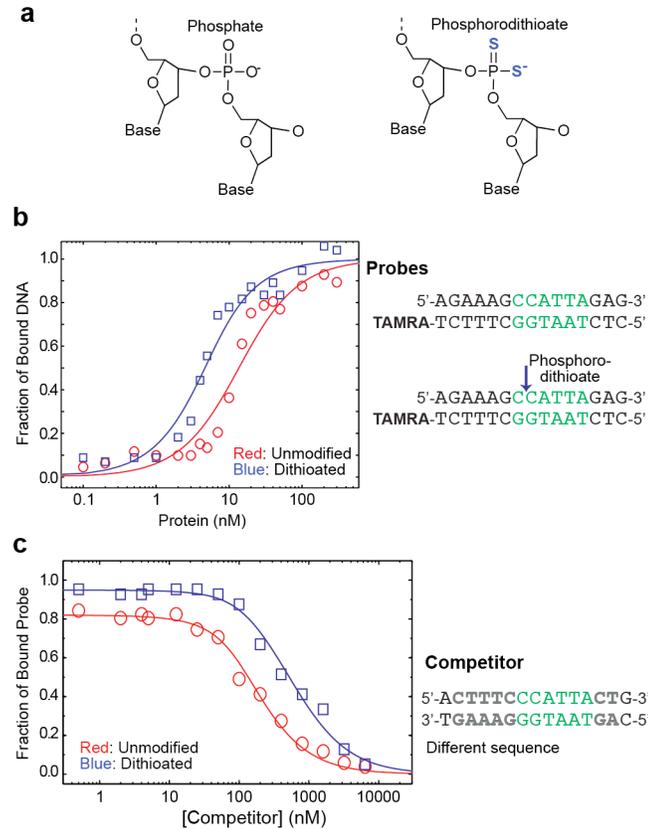


Figure 3.1: Enhancement of binding affinity for the Antp homeodomain by dithioation of the DNA phosphate at the Lys57 interaction site. **(a)** Dithioation of DNA phosphate. **(b)** Binding isotherm as measured by the fluorescence anisotropy-based titration experiment. Fluorescence anisotropy for TAMRA-labeled 15-bp duplexes (red, unmodified DNA; blue, dithioated DNA) was measured as a function of the concentration of the Antp homeodomain. The probe DNA concentration was 3.3 nM in this experiment. Dissociation constants determined from the three replicates together with Eq. 1 were $K_d = 11 \pm 1$ nM for the unmodified DNA and $K_d = 2.9 \pm 0.2$ nM for the dithioated DNA **(c)** Competition assay data used to indirectly determine K_d constants. The total concentrations of the probe DNA and the protein were 10 and 50 nM, respectively in this experiment. The measurement was replicated three times. The fitting calculations using Eqs 2 and 3 together with the K_d constant for the competitor gave $K_d = 15.3 \pm 0.8$ nM for the unmodified DNA and $K_d = 4.0 \pm 0.5$ nM for the dithioated DNA. The solid lines represent best-fit curves.

3.3.1 Isolation of R_p and S_p diastereomers of monothioated DNA

The DNA strand containing a phosphoromonothioate at the Lys57 interaction site was initially purified by Mono-Q anion-exchange chromatography. Though the R_P and S_P diastereomers could not be isolated at this stage, the expected mass (4,650 Da) of the monothioated strand (Figure 3.2B) was confirmed. By annealing an equimolar mixture of this and complementary strands, we prepared the double-stranded DNA. We further purified the double-stranded DNA by Mono-Q anion-exchange chromatography and obtained two major peaks of nearly the same magnitude (Figure 3.2C). We attributed these peaks to the R_p and S_p diastereomers because Frederiksen and Piccirilli reported such a separation for double-stranded RNA containing a phosphoromonothioate for which suitable resolution of the diastereomers could not be achieved for individual strands.¹³⁰ Their empirical rule suggested that the first and second peaks correspond to R_P and S_P diastereomers, respectively. Our crystallographic data confirmed that this was correct, as described below. The isolation of the R_p and S_p -monothioated DNA duplexes allowed us to characterize the stereospecific impacts of the oxygen-to-sulfur substitutions on the ion pair with the Lys57 side chain in the Antp homeodomain-DNA complex.

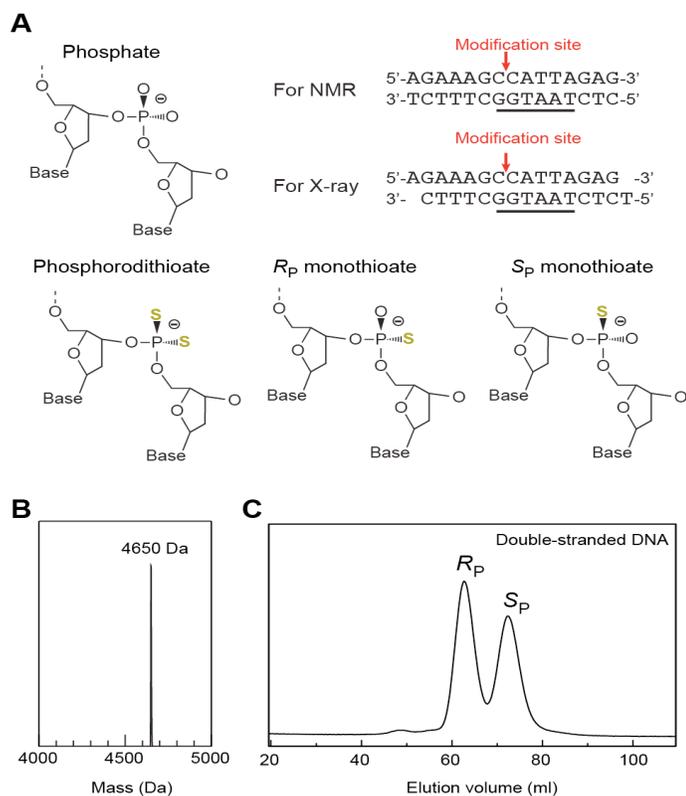


Figure 3.2: Oxygen-to-sulfur substitution in the DNA phosphate at the Lys57 interaction site in the Antp homeodomain-DNA complex. (A) Two diastereomers (R_P and S_P) of monothioate were used in the current study, whereas a phosphorodithioate was used in our previous study. In R_P and S_P monothioations, a sulfur atom substitutes for OP2 and OP1, respectively. (B) Mass spectrometry data for the monothioated DNA strand. (C) Chromatogram in Mono-Q anion-exchange chromatography for 15-bp DNA duplexes containing a phosphoromonothioate at the Lys57 site. A diastereomeric mixture of the phosphoromonothioate of the DNA duplex gives two major peaks corresponding to the R_P and S_P diastereomers, which were confirmed by crystallography (see Figure 3.10).

3.3.2 Affinity enhancement by dithioation of DNA phosphate

Using two different fluorescence-based assays with TAMRA-labeled DNA as a fluorescent probe, we examined the influence of the dithioation on target DNA association of the Antp homeodomain at 150 mM NaCl. By protein titration assays, in which TAMRA fluorescence anisotropy is monitored as a function of the protein concentration, we directly measured the dissociation constants K_d for the complexes with 15 bp DNA duplexes (Figure 3.1b). From the protein titration data, we determined values of K_d of the Antp homeodomain to be 11 ± 1 nM for the unmodified DNA and 2.9 ± 0.2 nM for the dithioated DNA.

Because the affinities were close to the limit of the measurable range in the protein titration assay and the probe concentration used was comparable to K_d values, we also analyzed affinities by competition assays that allow for K_d determination for high affinity systems. These assays involve competitor DNA (15-bp) that contains the same 6-bp target sequence but differs from the probe DNA in the other parts. In this competition assay, the difference in sequence between the probe and competitor DNA is important because otherwise, transfer of the dithioated strand to competitor DNA can occur. We used 50 nM protein and 10 nM probe DNA and varying concentrations of the competitor DNA. The fluorescence anisotropy changes as the unlabeled competitor increasingly outcompetes the probe DNA (Figure 3.1c). From these data together with the affinity of the competitor DNA, we determined the K_d values for the unmodified and dithioated complexes to be 15.3 ± 0.8 and 4.0 ± 0.5 nM, respectively. Thus, both datasets indicate that the oxygen-to-sulfur substitution in the DNA phosphate enhances the binding affinity for the Antp homeodomain by a factor of ~ 4 . From the K_d data, the change in binding free energy $\Delta\Delta G$

upon the oxygen-to-sulfur substitution in the DNA phosphate was calculated to be -0.8 ± 0.1 kcal/mol for the Antp-DNA complexes.

3.3.3 Thermodynamic impact of phosphorodithioate on association

For the thermodynamic characterization, we adopted the approach of Dragan et al.¹² that combines the fluorescence and ITC methods. In this approach, the binding enthalpy ΔH is obtained from the ITC data and the binding free energy ΔG is obtained from the fluorescence anisotropy-based titration data. Because ITC allows for direct observation of the heat from association, ΔH can be measured directly even if the dissociation constant K_d is too small to determine by ITC. For this approach to be valid, however, the fluorescent probe (i.e., TAMRA) should not perturb the binding properties of DNA. We confirmed this by a competition assay using TAMRA-labeled and unlabeled DNA with the identical sequence: the obtained K_d of the unlabeled DNA was virtually the same within experimental uncertainties (Figure 3.3a).

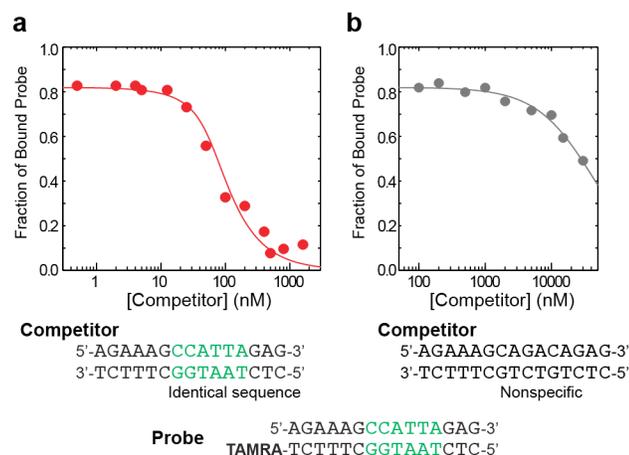


Figure 3.3: Competition assay data used to assess unlabeled specific DNA (a) and nonspecific DNA (b) duplexes. The total concentrations of the probe DNA and the Antp homeodomain were 10 and 50 nM, respectively. The fitting calculations using Eqs 2 and 3 together with the K_d constant for the probe gave $K_d = 10 \pm 2$ nM for the specific DNA and $K_d = (7.9 \pm 0.8) \times 10^3$ nM for the nonspecific DNA.

For the 15-bp DNA duplexes with and without the dithioation at the Lys57 interaction site, we conducted the ITC experiments in which the Antp homeodomain was injected into the DNA solutions to the final molar ratio of 5.3 (Figure 3.4a). When the molar ratio is larger than 1, we observed non-monotonic change of the heat effect over a wide range of the molar ratio, suggesting the presence of weak association events. Judging from the fluorescence-based K_d data, however, it is very unlikely that this effect is due to the specific association with the target site. These ITC data clearly suggest the presence of the multiple-binding sites^{131, 132}, most likely due to nonspecific association of DNA with additional protein molecules¹³². In fact, the crystal structures of the Antp homeodomain-DNA complexes suggest that two additional protein molecules can nonspecifically interact with

DNA regions that are not covered by the protein bound to the target site. Using the fluorescence-based competition assay, we measured the apparent affinity for a nonspecific 15-bp DNA duplex that does not contain the core recognition sequence for Antp (Figure 3.3b). The apparent K_d constant of this nonspecific DNA was $(7.9 \pm 0.8) \times 10^3$ nM, which is comparable to the sample concentrations used in the ITC experiment. With this affinity, influence of the nonspecific association on the fluorescence data for the specific DNA duplexes (Figure 3.1) is virtually negligible because the sample concentrations are lower than the apparent K_d constant for nonspecific complexes. We also conducted the ITC experiment for the same 15-bp nonspecific DNA duplex (Figure 3.4b). The ITC data suggest that nonspecific association of the first and second protein molecules generate opposite heat effects. These data suggest that the abnormal ITC profiles at molar ratios higher than 1 (Figure 3.4a) are due to nonspecific association.

When the molar ratio is significantly less than 1, the contribution of the nonspecific association to the heat effect is virtually negligible, because the affinity for the target site is stronger than that for nonspecific DNA by a factor >500 . To analyze the binding enthalpy for the specific association, we conducted additional ITC experiments for the specific DNA duplexes using a larger number of injections in a narrower range of molar ratio (up to 1.8) (Figure 3.4c). In this second set of ITC data, the heat effects at molar ratio less than 0.9 were virtually constant. Using these data at low molar ratio, we obtained the enthalpies of the specific association. Interestingly, the binding enthalpies were virtually identical for the unmodified and dithioated DNA. Table 3.1 summarizes the obtained binding free energy, enthalpy, and entropy for each DNA. Differences in the binding free energy ($\Delta\Delta G$) and its

enthalpic ($\Delta\Delta H$) and entropic ($-T\Delta\Delta S$) terms clearly indicate that the affinity enhancement by the oxygen-to-sulfur substitution in DNA phosphate is entropy-driven (Figure 3.4d).

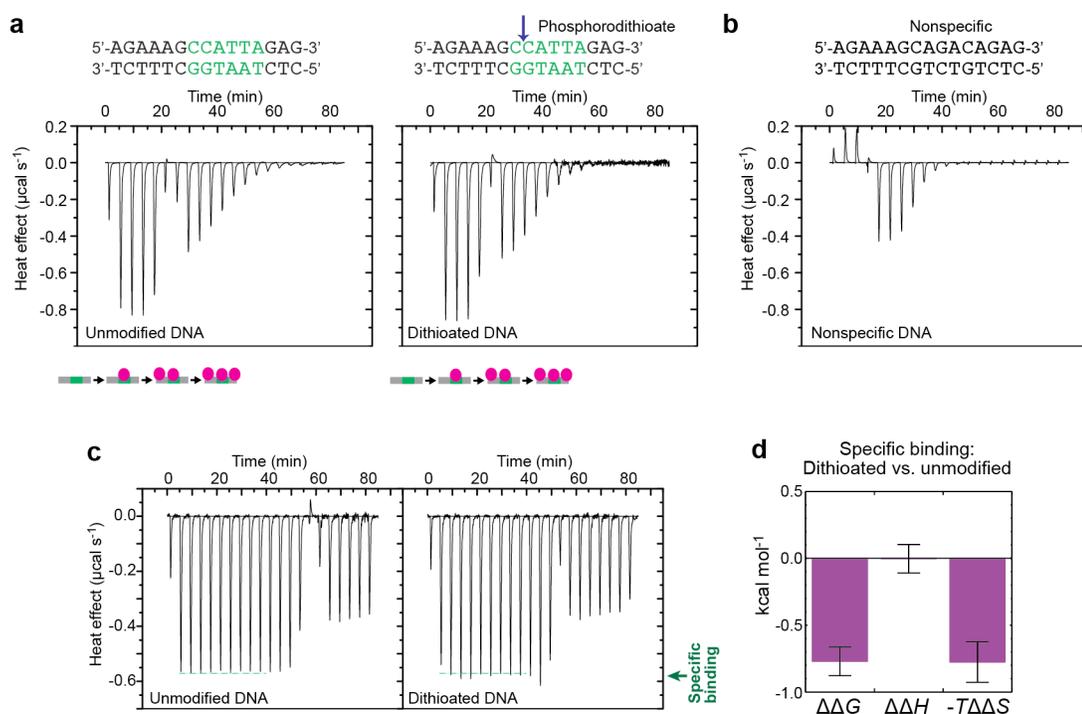


Figure 3.4: Thermodynamic impact of dithioation of DNA phosphate at the Lys57 interaction site. **(a)** ITC data for the specific DNA duplexes with 21 injections of the Antp homeodomain, leading to the final molar ratio of 5.3. The DNA concentration was 9 μM . Because the sample concentrations were comparable to the K_d for nonspecific association and the DNA duplexes are long enough to nonspecifically interact with additional protein molecules, heat from the nonspecific interactions was also observed after the target site was saturated. **(b)** ITC data for the nonspecific DNA duplex with 21 injections of the Antp homeodomain, leading to the final molar ratio of 6.6. The DNA concentration was 8 μM DNA. The heat effect of the first and second nonspecific associations were opposite in sign. **(c)** ITC data for the specific DNA duplexes with 21 injections of the Antp homeodomain leading to the final molar ratio of 1.8. The DNA concentration was 17 μM . Green dotted lines correspond to the heat effect arising from association of the Antp homeodomain with the target site on the DNA duplexes. **(d)** Changes in binding free energy ($\Delta\Delta G$), enthalpy ($\Delta\Delta H$) and entropic terms ($-T\Delta\Delta S$) upon the oxygen-to-sulfur substitution of the DNA phosphate. The increase in binding entropy $\Delta\Delta S$ was determined to be $2.5 \pm 0.5 \text{ cal K}^{-1} \text{ mol}^{-1}$. See also Table 1.

Table 3.1: Thermodynamic parameters on protein-DNA association measured for the Antp homeodomain-DNA complexes at 25°C. ^a

Complex	ΔG [kcal/mol] ^b	ΔH [kcal/mol] ^c	ΔS [cal K ⁻¹ mol ⁻¹] ^d
Unmodified DNA ^e	-10.89 ± 0.08	-8.64 ± 0.06	7.5 ± 0.4
Dithioated DNA ^f	-11.66 ± 0.08	-8.68 ± 0.07	10.0 ± 0.4

^a Each uncertainty represents the standard error of the mean.

^b Measured from the fluorescence anisotropy-based titration data.

^c Measured with the ITC data (Figure 3.4c).

^d Calculated by $\Delta G = \Delta H - T\Delta S$.

^e Complex with unmodified DNA.

^f Complex with DNA containing phosphorodithioate at the Lys57 interaction site

3.3.4 Impact of R_p and S_p monothioation on protein-DNA binding affinity

Using fluorescence anisotropy-based competitive binding assays, we measured the binding affinities of the R_p and S_p -monothioated DNA duplexes for the Antp homeodomain (Figure 3.5A). The dissociation constants K_d at 150 mM NaCl were determined to be 3.0 ± 0.7 nM for the R_p -monothioated DNA and 7.3 ± 0.8 nM for the S_p -monothioated DNA. The uncertainties in the measured K_d values were calculated as the standard errors of the means for four independent measurements. In our previous work, the dissociation constants for the unmodified and dithioated DNA duplexes were determined to be 15.3 ± 0.8 and 4.0 ± 0.5 nM, respectively, by the same method under the same conditions.⁷⁰ Thus, the R_p -monothioated DNA duplex showed significantly stronger affinity than that of the unmodified DNA and virtually the same as that of the dithioated DNA. Figure 3.5B shows the thermodynamic impact of the modifications of the DNA phosphate at the Lys57 sites in terms of the free energies of the protein-DNA association. These data show that the sulfur substitution of the O_{P2} atom, which is closer to the Lys57 NH_3^+ group, yields a stronger impact on the affinity enhancement.

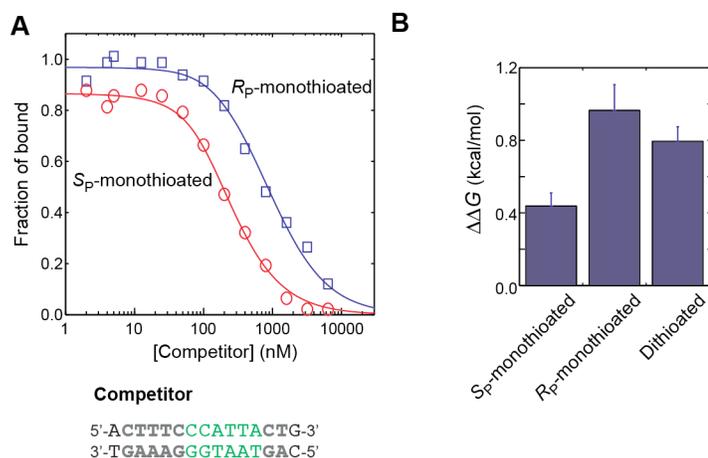


Figure 3.5: Impact of oxygen-to-sulfur substitution at Lys57 interaction site on binding affinity for the Antp homeodomain-DNA complex. (A) Competition assay data used to determine K_d constants for R_P and S_P -monothioated DNA duplexes. The total concentrations of the probe DNA and the protein were 10 and 50 nM, respectively in this experiment. The measurement was repeated four times. The solid lines represent best-fit curves. (B) Change in the binding free energy upon modifications of the DNA phosphate group at the Lys57 interaction site. In this case, $\Delta\Delta G$ is defined as ΔG (unmodified) – ΔG (modified).

3.3.5 NMR of interfacial Lys NH₃⁺ groups

To investigate the impact of the oxygen-to-sulfur substitution in the DNA phosphate on the intermolecular ion pairs with the protein side chain, we extensively characterized the Lys side-chain NH₃⁺ groups of the Antp homeodomain-DNA complexes by NMR. Figure 4a shows the Lys NH₃⁺-selective HISQC⁶⁴ spectra recorded for the complexes with unmodified or dithioated 15-bp DNA at 15 °C and pH 5.8. Four out of six Lys side-chain NH₃⁺ groups of the protein clearly showed the ¹H-¹⁵N cross peaks. By Lys-specific^{42, 56, 64, 122} and general¹³³ triple resonance experiments, these NH₃⁺ signals were assigned to Lys46, Lys55, Lys57, and Lys58. The NH₃⁺ groups of Lys46, Lys55, and Lys57 form ion pair with DNA phosphate⁵⁰. The NH₃⁺ groups of Lys2 and Lys18 were not observed, presumably due to their rapid hydrogen exchange.

3.3.6 NMR evidence for ionic hydrogen bonds

A remarkable feature of NMR investigations on Lys NH₃⁺ groups is that quantitative measurements of relatively small (< 1 Hz) scalar couplings between ¹⁵N and other nuclei is feasible owing to extremely slow intrinsic ¹⁵N transverse relaxation of NH₃⁺ groups^{34, 42, 66}. For interfacial Lys NH₃⁺ groups forming a contact ion pair (CIP) with DNA, hydrogen-scalar coupling ^h*J*_{NP} between protein side-chain NH₃⁺ ¹⁵N and DNA ¹³P nuclei across a hydrogen bond could be detectable^{42, 68}. In fact, the H3(N)P spectrum³⁶ recorded for the Antp homeodomain – dithioated DNA complex (Figure 3.6b) clearly shows ¹H-³¹P cross peaks that arise from heteronuclear ¹⁵N-³¹P scalar couplings across a hydrogen bond between Lys NH₃⁺ and DNA phosphate / phosphorodithioate groups. Importantly, this spectrum shows a cross peak from the Lys57 NH₃⁺ group with ³¹P resonance at 107.7 ppm, a typical ³¹P chemical shift for phosphorodithioate, which is substantially different from

that for DNA phosphate (\sim -2 ppm) ^{134, 135}. By the spin-echo $^3J_{NP}$ -modulation constant-time HISQC experiment³⁶, the $^3J_{NP}$ constant was measured to be 0.23 Hz for the Lys57 – phosphorodithioate ion pair. The $^3J_{NP}$ constants for the Lys46 and Lys55, which form a CIP with DNA phosphate, were measured to be 0.62 and 0.41, respectively. Lys58 did not exhibit $^3J_{NP}$ coupling, which is reasonable because this residue does not form a CIP in the crystal structures. The $^3J_{NP}$ data represent direct evidence for the hydrogen bonds between the Lys side-chain NH_3^+ and DNA phosphate / phosphorodithioate groups in the complexes.

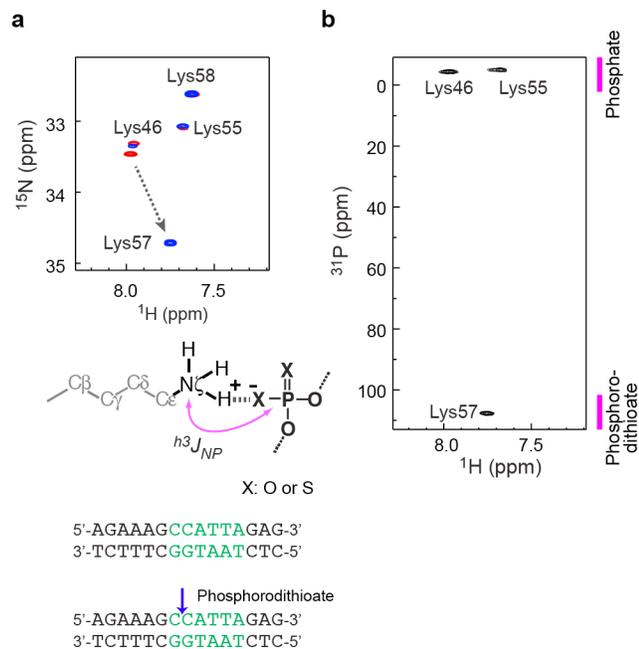


Figure 3.6: NMR evidence for the hydrogen bonds between the Lys NH_3^+ and DNA phosphate / phosphorodithioate groups. (a) Overlaid HISQC spectra recorded for the Lys NH_3^+ groups in the unmodified (red) and dithioated (blue) DNA complexes of the Antp homeodomain. (b) The H3(N)P correlation spectrum recorded for the dithioated DNA complex. This spectrum shows ^1H - ^{31}P cross peaks arising from coherence transfers via hydrogen-bond scalar coupling h^3J_{NP} between Lys ^{15}N and DNA ^{31}P nuclei. The signals represent the direct evidence for ionic hydrogen bonds with DNA phosphate (^{31}P ~-2 ppm) or phosphorodithioate (~108 ppm).

3.3.7 Mobility of interfacial Lys NH₃⁺ groups

To investigate the impact of the oxygen-to-sulfur substitution on the dynamics of ionic interactions between protein side chain and DNA phosphate, we conducted ¹⁵N relaxation analysis for the Lys side-chain NH₃⁺ groups of the two Antp-homeodomain DNA complexes at 25 °C. Using 800-MHz and 600-MHz NMR spectrometers, we measured heteronuclear ¹H-¹⁵N NOE and ¹⁵N *R*₁ and *R*₂ relaxation rates for the Lys NH₃⁺ groups of these complexes. The measured values of these parameters are given in Table 3.2. Upon dithioation of the DNA phosphate, only the Lys57 NH₃⁺ group exhibited significant changes in ¹⁵N relaxation parameters (Figure 3.7a and b; see also Table 3.2). These mostly likely reflect a change in mobility rather than in covalent geometry, because NH₃⁺ groups are known to retain almost ideal tetrahedral geometry even in an ion pair with an acidic group¹³⁶. From the ¹⁵N relaxation data, we determined the order parameters S^2_{axis} and reorientation correlation times τ_i for symmetry axes of NH₃⁺ groups and C ϵ -N ζ bond-rotation correlation times τ_f (Table 3.2). As we observed previously for the HoxD9 homeodomain, the interfacial Lys NH₃⁺ groups (Lys46, Lys55, and Lys58) exhibited high mobility with $S^2_{axis} < 0.6$ despite the ionic interaction with DNA. Among the four NH₃⁺ groups analyzed, only the Lys57 NH₃⁺ group exhibited statistically different S^2_{axis} values for the two complexes (Figure 3.7c), which is reasonable because this is the only Lys residue close to the oxygen-to-sulfur substitution site. The S^2_{axis} value for the complex with dithioated DNA was significantly smaller than that for the complex with unmodified DNA (0.34 vs. 0.54). These results indicate that the intermolecular ion pair of Lys57 gets mobilized upon the oxygen-to-sulfur substitution in DNA phosphate, although our $^hJ_{NP}$

data clearly indicate the presence of a hydrogen bond between the NH_3^+ and phosphorodithioate groups.

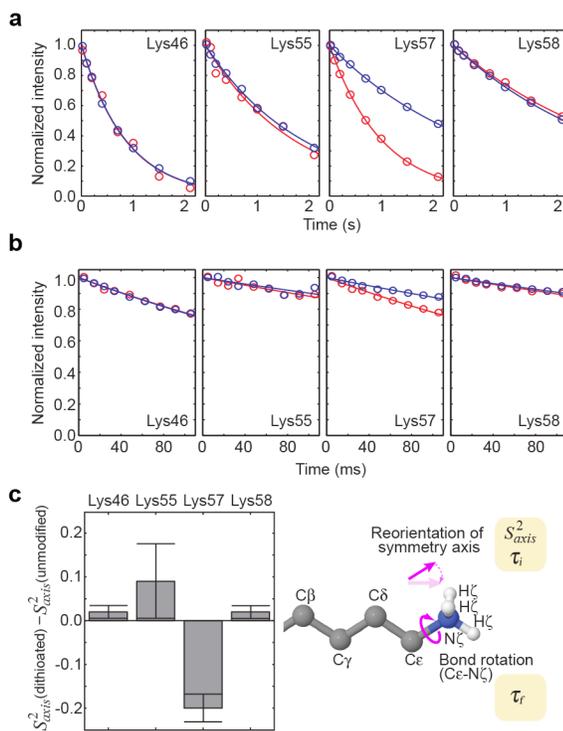


Figure 3.7: Mobilization of the Lys57 NH_3^+ group by dithioation of the interacting DNA phosphate group. (a) Lys NH_3^+ ^{15}N R_1 relaxation (^1H frequency, 600 MHz) for the unmodified (red) and dithioated (blue) DNA complexes. (b) Lys NH_3^+ ^{15}N R_2 relaxation (^1H frequency, 800 MHz) for the unmodified (red) and dithioated (blue) DNA complexes. (c) Changes in order parameters S^2_{axis} for the symmetry axis of Lys NH_3^+ groups upon dithioation of the DNA phosphate.

Table 3.2: ^{15}N relaxation and dynamics parameters measured for the Lys side-chain NH_3^+ groups of the Antp homeodomain-DNA complexes at 25°C and pH 5.8.^{a)}

	Lys46 NH_3^+	Lys55 NH_3^+	Lys57 NH_3^+	Lys58 NH_3^+
Complex with unmodified DNA				
-800 MHz-				
$^{15}\text{N } R_1$ (s^{-1})	1.01 ± 0.02	0.49 ± 0.03	0.84 ± 0.01	0.28 ± 0.01
$^{15}\text{N } R_{2,ini}$ (s^{-1}) ^{b)}	2.62 ± 0.06	1.60 ± 0.25	2.53 ± 0.04	1.05 ± 0.04
$^1\text{H-}^{15}\text{N}$ NOE	-2.34 ± 0.08	-2.40 ± 0.23	-2.67 ± 0.03	-2.47 ± 0.06
-600MHz-				
$^{15}\text{N } R_1$ (s^{-1})	1.17 ± 0.03	0.58 ± 0.05	0.98 ± 0.01	0.30 ± 0.01
$^1\text{H-}^{15}\text{N}$ NOE	-2.93 ± 0.10	-2.90 ± 0.46	-2.84 ± 0.20	-3.09 ± 0.20
-Dynamics-^{c)}				
S^2_{axis}	0.49 ± 0.01	0.34 ± 0.07	0.54 ± 0.03	0.24 ± 0.01
τ_f (ps)	265 ± 11	60 ± 84	77 ± 53	5.4 ± 0.6
τ_i (ps)	0.0 ± 0.1	326 ± 170	550 ± 272	217 ± 14
Complex with dithioated DNA				
-800 MHz-				
$^{15}\text{N } R_1$ (s^{-1})	1.02 ± 0.01	0.50 ± 0.01	0.33 ± 0.01	0.31 ± 0.01
$^{15}\text{N } R_{2,ini}$ (s^{-1}) ^{b)}	2.74 ± 0.07	1.83 ± 0.17	1.39 ± 0.02	1.10 ± 0.03
$^1\text{H-}^{15}\text{N}$ NOE	-2.49 ± 0.05	-2.75 ± 0.09	-2.35 ± 0.02	-2.60 ± 0.03
-600MHz-				
$^{15}\text{N } R_1$ (s^{-1})	1.17 ± 0.02	0.53 ± 0.02	0.35 ± 0.01	0.32 ± 0.01
$^1\text{H-}^{15}\text{N}$ NOE	-2.88 ± 0.07	-3.00 ± 0.12	-2.63 ± 0.02	-2.93 ± 0.03
-Dynamics-^{c)}				
S^2_{axis}	0.51 ± 0.01	0.43 ± 0.05	0.34 ± 0.01	0.26 ± 0.01
τ_f (ps)	243 ± 9	27 ± 12	8.4 ± 0.2	7.7 ± 0.3
τ_i (ps)	0.0 ± 0.01	212 ± 70	248 ± 6	211 ± 9

^{a)} Signals from the Lys2 and Lys18 NH_3^+ groups were not observed due to rapid hydrogen exchange with water.

^{b)} The initial rate for intrinsically bi-exponential ^{15}N transverse relaxation of NH_3^+ (32).

^{c)} Symbols are defined in Fig. 5c. The molecular rotational correlation time and anisotropy were determined to be 9.9 ns and 1.3, respectively, from backbone ^{15}N relaxation data.

3.3.8 Impact of R_p and S_p monothioation on ion-pair dynamics

To investigate the impact of monothioation on the dynamics of the intermolecular ion pair with Lys57, we used the NMR methods for the Lys side-chain NH_3^+ groups. Figure 3.8A shows an overlay of the NH_3^+ -selective ^1H - ^{15}N heteronuclear in-phase single-quantum coherence (HISQC)⁶⁴ spectra recorded for Lys side chains of the Antp homeodomain in the complexes with the unmodified and modified DNA complexes. As seen in Figure 3.6a, the dithioation of DNA phosphate significantly increased the ^{15}N chemical shift of the interacting Lys NH_3^+ group. The monothioation of the interacting phosphate also significantly increased the ^{15}N chemical shift of the Lys57 NH_3^+ group but to a lesser degree. The ^1H - ^{15}N correlation signals from the Lys57 NH_3^+ group of the complexes with the R_p and S_p -monothioated DNA were observed at different positions.

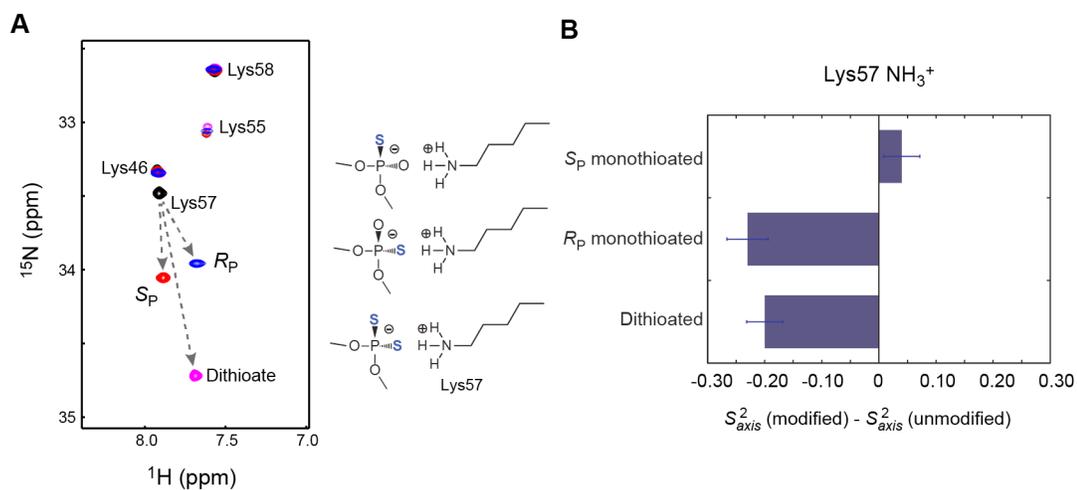


Figure 3.8: Impact of the oxygen-to-sulfur substitutions on the intermolecular ion pair with Lys57. (A) Overlaid Lys side-chain NH_3^+ -selective ^1H - ^{15}N HISQC spectra⁶⁴ recorded for the four complexes of the Antp homeodomain and DNA. These complexes chemically differ only in the DNA phosphate at the Lys57 interaction site (black, unmodified; blue, R_P -monothioated; red, S_P -monothioated; and magenta, dithioated). (B) Change in NMR-derived order parameters S^2_{axis} for Lys NH_3^+ groups upon oxygen-to-sulfur substitution in DNA phosphate at the Lys57 interaction site. These data show that sulfur substitution of the proximal phosphate oxygen (i.e., $\text{O}_{\text{P}2}$) atom significantly mobilizes the Lys side-chain NH_3^+ group.

To analyze the internal motions of the interfacial Lys NH₃⁺ groups, we collected ¹⁵N relaxation data at the ¹H frequencies of 600 and 800 MHz. Using the ¹⁵N relaxation data, we determined the order parameters S^2_{axis} , bond-rotation correlation times τ_f , and reorientation correlation times τ_i for the Lys side-chain NH₃⁺ groups of the *R_p* and *S_p*-monothioated DNA complexes, as described in our previous papers.^{24, 42, 55, 65, 68, 70} The values of these relaxation and dynamic parameters are shown in Table 3.3. Only Lys57 exhibited significant changes in the NH₃⁺ order parameters upon *R_p* monothioation and dithioation of the phosphate. This is reasonable because only this Lys side chain is located at the modification site. We examined the changes in the order parameters upon the oxygen-to-sulfur substitution in the phosphate group at the Lys57 site (Figure 3.8B). The sulfur substitution of the O_{P2} atom, which is more proximal to the Lys 57 NH₃⁺ group, via *R_p* monothioation or dithioation caused a significant increase in the mobility of the Lys57 NH₃⁺ group, whereas the sulfur substitution of the O_{P1} atom via *S_p* monothioation did not. Importantly, the increase in the mobility of the intermolecular ion pair was associated with the increase in the protein-DNA affinity.

Table 3.3: NMR data for Lys side-chain NH_3^+ groups in the complexes of the Antp homeodomain with R_P and S_P -monothioated DNA.^{a)}

	Lys46 NH_3^+	Lys55 NH_3^+	Lys57 NH_3^+	Lys58 NH_3^+
R_P-monothioated				
DNA complex				
-800 MHz-				
$^{15}\text{N } R_l$ (s^{-1})	1.02 ± 0.01	0.50 ± 0.01	0.35 ± 0.01	0.31 ± 0.01
$^{15}\text{N } R_{2,ini}$ (s^{-1}) ^{b)}	2.74 ± 0.07	1.83 ± 0.17	1.38 ± 0.08	1.10 ± 0.03
$^1\text{H-}^{15}\text{N}$ NOE	-2.49 ± 0.05	-2.75 ± 0.09	-2.53 ± 0.06	-2.60 ± 0.03
-600MHz-				
$^{15}\text{N } R_l$ (s^{-1})	1.16 ± 0.02	0.53 ± 0.02	0.40 ± 0.01	0.32 ± 0.01
$^1\text{H-}^{15}\text{N}$ NOE	-2.88 ± 0.07	-3.00 ± 0.12	-2.87 ± 0.09	-2.93 ± 0.03
-Dynamics-				
S^2_{axis}	0.51 ± 0.01	0.41 ± 0.05	0.31 ± 0.02	0.25 ± 0.01
τ_f (ps)	240 ± 8	26 ± 11	12 ± 1	7.5 ± 0.3
τ_i (ps)	0.26 ± 0.93	243 ± 57	265 ± 19	214 ± 8
S_P-monothioated				
DNA complex				
-800 MHz-				
$^{15}\text{N } R_l$ (s^{-1})	1.01 ± 0.01	0.51 ± 0.01	0.74 ± 0.01	0.29 ± 0.01
$^{15}\text{N } R_{2,ini}$ (s^{-1}) ^{b)}	2.67 ± 0.06	1.60 ± 0.18	2.60 ± 0.02	1.41 ± 0.04
$^1\text{H-}^{15}\text{N}$ NOE	-2.49 ± 0.04	-2.74 ± 0.08	-2.89 ± 0.02	-2.53 ± 0.02
-600MHz-				
$^{15}\text{N } R_l$ (s^{-1})	1.21 ± 0.02	0.56 ± 0.03	0.85 ± 0.01	0.32 ± 0.01
$^1\text{H-}^{15}\text{N}$ NOE	-2.80 ± 0.05	-2.92 ± 0.08	-2.97 ± 0.02	-2.98 ± 0.03
-Dynamics-				
S^2_{axis}	0.51 ± 0.01	0.34 ± 0.05	0.58 ± 0.01	0.31 ± 0.01
τ_f (ps)	249 ± 7	40 ± 53	49 ± 1	8.5 ± 0.3
τ_i (ps)	0.07 ± 0.29	285 ± 89	348 ± 14	130 ± 10

^{a)} Measured at 25 °C and pH 5.8. Under these conditions, the Signals from the Lys2 and Lys18 NH_3^+ groups were not observed due to rapid hydrogen exchange with water.

^{b)} The initial rate for intrinsically bi-exponential ^{15}N transverse relaxation of NH_3^+ .⁶⁵

3.3.9 Crystal structures of the complexes

For structural investigation, we crystallized the Antp homeodomain-DNA complexes with and without the dithioation for that phosphate. DNA duplexes with 5'-overhangs were used for the crystallization (Figure 3.9). X-ray diffraction data for both complexes were collected to 2.7 Å. The crystals of these two complexes gave the same space group P43212 and virtually the same cell dimensions. Although we obtained the crystals under the conditions described by Fraenkel and Pabo, the space group of our crystals was different from the previous one, P2221⁵⁰. This might be due to slight difference in the N-termini of the protein constructs: Ours contains a Met residue (from the initial codon) prior to the first residue of the 60-amino-acid Antp homeodomain, whereas the corresponding N-terminal addition is Met-Glu in the construct of Fraenkel and Pabo⁵⁰. The asymmetric unit of our crystals contained two complexes. Refinement produced structures with good geometry and no Ramachandran outliers with free-R values of 28.7% for the unmodified DNA complex and 27.0% for the dithioated DNA complex. Table 3.4 summarizes the crystallographic data and statistics. Figure 3.9a shows a superposition of our crystal structures of the dithioated (PDB ID: 4XIC) and unmodified (PDB ID 4XID) DNA complexes of the Antp homeodomain. The overall backbone structures of these complexes were virtually the same with RMSD being only 0.45 Å for the backbone atoms. DNA groove widths/depths calculated with the CURVES+ program¹³⁷ were virtually the same for the unmodified and dithioated DNA complexes (Figure 3.9b). These results indicated that dithioation of a single phosphate does not impact the overall structure of the protein-DNA complex.

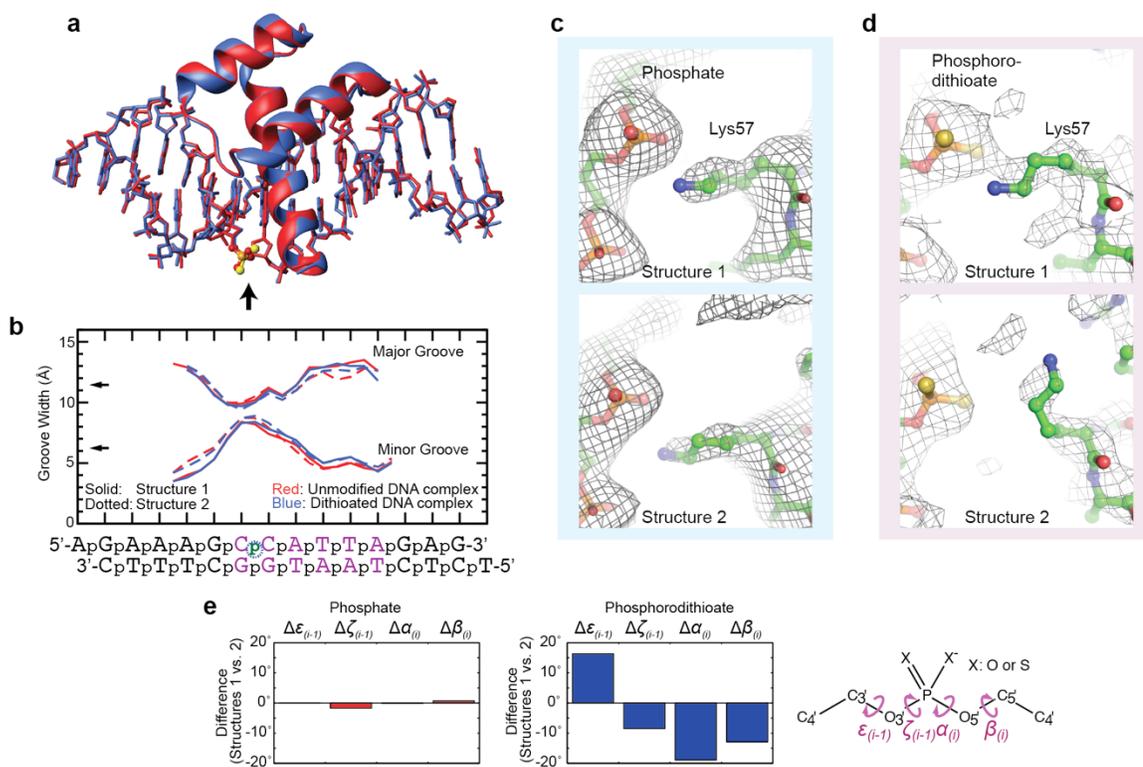


Figure 3.9: Structural impact of dithioation of the DNA phosphate at Antp Lys57 interaction site. (a) Superposition of the crystal structures of the Antp homeodomain-DNA complexes with (blue) and without (red) dithioation. The phosphate and phosphorodithioate groups at the modification site are indicated by an arrow. (b) Major and minor groove widths of the DNA double helices of the complexes. Solid and dotted lines show the data for Structures 1 and 2, respectively, which are two independent structures in the asymmetry unit (see the main text). The values were calculated with the CURVES+ program (44). Horizontal arrows indicate values for the canonical B-form. (c, d) Electron density maps together with the structures of the Lys57 side chain – DNA phosphate (Panel c) / phosphorodithioate (Panel d) ion pairs. Two independent structures in the asymmetric unit are shown for each complex. (e) DNA backbone torsion angles relevant to the phosphorodithioate and phosphate groups at the Lys57 interaction site.

Table 3.4: Crystallographic data collection and refinement statistics.

	Unmodified DNA complex (PDB ID: 4XID)	Dithioated DNA complex (PDB ID: 4XIC)
Crystallographic data collection		
X-ray Source	Rigaku FRE++	APS BL-19ID
Wavelength (Å)	1.5418	0.9792
Space Group	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2
Unit cell parameters (Å, °)	$\alpha = 96.44$, $b = 96.44$, $c = 90.14$, $\alpha = \beta = \gamma = 90$	$\alpha = 96.54$, $b = 96.54$, $c = 89.55$, $\alpha = \beta = \gamma = 90$
Sample temperature (K)	85	90
Resolution range (Å)	37.6 – 2.7	37.4 – 2.7
Total reflections	22043	18507
Non-anomalous reflections	12006	10913
Completeness (%)	99.0	99.8
Multiplicity	7.8	8.9
R_{merge}	0.054	0.127
R_{pim}	0.024	0.043
Refinement		
R_{work} (%)	23.3	21.6
R_{free} (%)	28.7	27.2
Bond RMSD from ideal values (Å)	0.005	0.006
Angle RMSD from ideal values (°)	0.746	0.823
Ramachandran plot: Favored (%)	98.2	100
Allowed (%)	1.8	0
Outliers (%)	0	0
No. non-H atoms: Protein	1055	1074
DNA	1222	1218
Water	15	14

3.3.10 Lys57 – DNA interactions in the crystal structures

We compared the structural details of the interactions between Lys57 side chain and DNA phosphate / phosphorodithioate for the crystal structures of the Antp homeodomain-DNA complexes. Fortunately, owing to the presence of two independent molecules per asymmetric unit, our crystallographic data provide some insight into the structural dynamics relevant to the oxygen-to-sulfur substitution. The two structures of the unmodified DNA complex show similar ion-pairing interactions between the Lys57 side chain and the DNA phosphate involving an ionic hydrogen bond (i.e., CIP) (Fig 3.9c). This observation is consistent with the crystal structure of Fraenkel and Pabo⁵⁰ as well as with our NMR hydrogen-bond scalar coupling $^hJ_{NP}$ data. In contrast, the two structures of the dithioated DNA complex show significantly different interactions between the DNA phosphorodithioate and the Lys57 side chain (Figure 3.9d). In one of the structures (Structure 1), the NH_3^+ group is in contact with phosphorodithioate, forming a CIP (N...S distance, 3.0 Å). In the other structure (Structure 2), the Lys57 side chain exhibited a solvent-separated ion pair with phosphorodithioate (N...S distance, 5.5 Å). These results suggest that the Lys57 side chain in the dithioated DNA complex is more dynamic than in the unmodified DNA complex. The phosphorodithioate itself also appears to be more dynamic than the corresponding phosphate. The DNA backbone torsion angles ϵ , ζ , α , and β for the phosphorodithioate differ by 10-20° between the two structures of the dithioated DNA complex, whereas the corresponding differences are less than 2° for the unmodified DNA complex (Figure 3.9e). Although these crystallographic data show two distinct modes, our NMR spectra show a single cross peak from the Lys57 NH_3^+ group. ¹⁵N relaxation dispersion data for the NH_3^+ groups also show no evidence that Lys57 undergoes

fast exchange between multiple states on a μs – ms timescale. The internal motions indicated by low order parameters from the ^{15}N relaxation data are in ps – ns timescale. The discrete states observed in the crystal structures might be due to the difference in temperature ($25\text{ }^{\circ}\text{C}$ in the NMR relaxation experiments vs. $\sim -185\text{ }^{\circ}\text{C}$ in the X-ray experiments), water activity, or crystal packing force. Nonetheless, it is important to note that the NMR and crystallographic data are consistent in that the Lys side chain becomes more dynamic upon the oxygen-to-sulfur substitution in DNA phosphate.

3.3.11 Crystal structures of the R_p and S_p monothioated DNA complexes

For each diastereomer of the monothioated DNA, we determined the crystal structure of the DNA complex with the Antp homeodomain. The crystal structure of the R_p -monothioated DNA complex was solved at a 2.09 \AA resolution (Figure 3.10A), whereas that of the S_p -monothioated DNA complex was solved at a 2.75 \AA resolution (Figure 3.10B). Table 3.5 provides a summary of the crystallographic data and structure refinement. Due to the different numbers of electrons in oxygen (8 electrons) and sulfur (16 electrons), the R_p and S_p diastereomers of the phosphoromonothioate group at the Lys57 site were readily distinguishable in the electron density maps, especially at a 2.09 \AA resolution. The crystallographic results directly confirmed the stereochemical assignment of the two major peaks in the Mono-Q chromatogram (Figure 3.2C).

Not surprisingly, the monothioation of a single phosphate group did not cause any drastic changes in the overall structures. Compared with our previous crystal structures of the unmodified and dithioated DNA complexes at a 2.7 \AA resolution (Table 3.4), the root-mean-square differences (r.m.s.d.) for the protein and DNA backbone atoms were less than 0.4 \AA for both monothioated DNA complexes. Interestingly, the tip (i.e., the $\text{N}\zeta$, $\text{C}\epsilon$, and

C δ groups) of the Lys57 side chain, which should interact with the R_P phosphoromonothioate, was disordered in the crystallographic electron density map. This was consistent with the finding that the Lys57 NH $_3^+$ group of the R_P -monothioated complex showed the smallest NMR order parameter among the four complexes, as shown below.

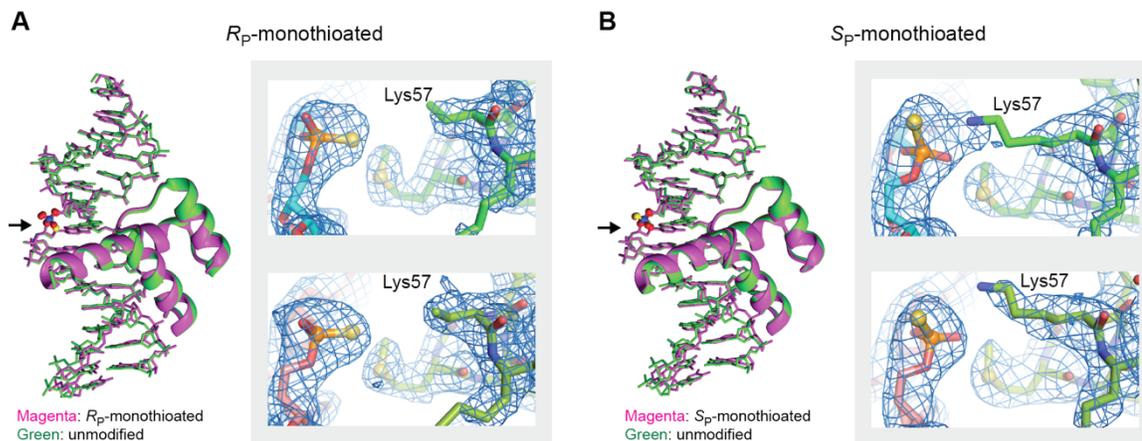


Figure 3.10: Crystal structures of the Antp homeodomain-DNA complexes containing R_P phosphoromonothioate (Panel A) or S_P phosphoromonothioate (Panel B) at the Lys57 interaction site. The crystal structures of the R_P and S_P -monothioated DNA complexes were determined at 2.09 and 2.75 Å resolutions, respectively. For each panel, superposition of the crystal structures of the modified (green) and unmodified (purple) complexes is shown. The modification sites are shown in ball-stick representation. (indicated by arrows). For each modified complex, the electron density map, of the phosphoromonothioate at Lys57 interaction site, for two structures in the asymmetric unit is shown. For the R_P -monothioated DNA complex, the tip (i.e., N ζ , C ϵ , and C δ groups) of the Lys57 side chain was not resolved in the crystallographic electron density map.

Table 3.5: Crystallographic data collection and refinement statistics for R_p and S_p -monothioated DNA complexes.

	R_p -monothioated DNA complex (PDB ID: 5JLW)	S_p -monothioated DNA complex (PDB ID: 5JLX)
Crystallographic data collection		
X-ray source	Bruker TXS	APS 21-ID-F
Wavelength (Å)	1.54	0.97921
Space group	P222 ₁	P222 ₁
Unit cell parameters (Å, °)	$a = 61.27, b = 75.89, c = 93.70,$ $\alpha = 90, \beta = 90, \gamma = 90$	$a = 60.86, b = 75.58, c = 91.88,$ $\alpha = 90, \beta = 90, \gamma = 90$
Sample temperature (K)	100	100
Resolution range (Å)	19.30 – 2.09	36.67 - 2.75
Total reflections	48110	9098
Non-anomalous reflections	25940	9098
Completeness (%)	97.5 (92.1) ^{a)}	99.4 (99.1) ^{a)}
Multiplicity	5.7 (4.2) ^{a)}	6.4 (6.2) ^{a)}
R_{merge}	0.0791 (0.50) ^{a)}	0.152 (1.2) ^{a)}
R_{pim}	0.0499 (0.35) ^{a)}	0.064 (0.513) ^{a)}
Refinement		
R_{work} (%)	21.81	19.26
R_{free} (%)	24.24	24.64
Bond RMSD from ideal values (Å)	0.009	0.007
Angle RMSD from ideal values (°)	1.1	0.8
Ramachandran plot: Favored (%)	114	115
Allowed (%)	0	0
Outliers (%)	0	0
No. non-H atoms [avg B-factor]:		
Protein	1067 [39]	1048 [53]
DNA	1218 [52]	1218 [67]
Water	212 [48]	25 [48]

^{a)} Numbers in parentheses are for the last shell (2.19 – 2.09 Å for 5JLW and 3.15 – 2.75 Å for 5JLX).

3.4 DISCUSSION

3.4.1 Mobilization of ion pair by oxygen-to-sulfur substitution in DNA phosphate

Our NMR and crystallographic data indicate that the sulfur substitution of the proximal non-bridging oxygen of the DNA phosphate mobilizes the interacting Lys side chain. The mobilization could be related to the relatively flat energy surface of H••S hydrogen bonds and a larger effective radius of sulfur (1.84 Å for sulfur vs. 1.40 Å for oxygen)¹¹⁶. Theoretical quantum chemical studies have shown that when compared with H••O hydrogen bonds, the enthalpy for H••S hydrogen bonds is slightly smaller and has a flatter energy surface.^{138, 139} Due to the flatter energy surface for sulfur, a slight deviation from an ideal hydrogen bond geometry causes only a marginal increase in enthalpy and may allow for a wider spatial distribution of a Lys side-chain NH₃⁺ group interacting with sulfur (Figure 3.11). This effect might make the Lys NH₃⁺ group more dynamic. Oxygen-to-sulfur substitutions might also perturb the dynamic equilibrium of the contact ion-pair (CIP) and solvent-separated ion-pair (SIP) states.¹⁷ As we recently demonstrated, the intermolecular ion pairs of Lys side-chain and DNA phosphate groups undergoes dynamic transitions between the CIP and SIP states on a sub-nanosecond timescale.²⁴ Oxygen-to-sulfur substitution in DNA phosphate might shift the CIP-SIP equilibrium of the intermolecular ion pair toward the SIP state, mobilizing the Lys NH₃⁺ group. Weaker interactions of the sulfur atoms with water molecules¹⁰⁶ might also contribute to mobilization of the intermolecular ion pair.

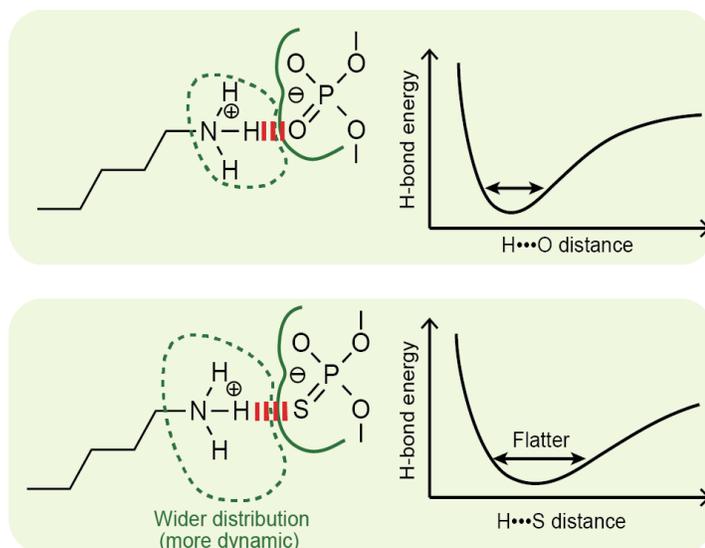


Figure 3.11: Possible reason for the mobilization of Lys57 side chain due to oxygen-to-sulfur substitution in DNA phosphate forming an intermolecular ion pair.

3.4.2 Entropic gain due to mobilization of the intermolecular ion pair

By using the experimental order parameters S^2_{axis} together with a particular motional model, we could roughly estimate the entropic gain due to mobilization of the Lys57 NH_3^+ group upon the oxygen-to-sulfur substitution in DNA phosphate. Assuming the diffusion-in-a-cone model⁹⁵, which gives entropic difference equal to $k_B \ln[\{3 - (1 + 8S^2_{axis,a})^{1/2}\} / \{3 - (1 + 8S^2_{axis,b})^{1/2}\}]$ (k_B , the Boltzmann constant), the increase in entropy for the symmetry axis of the Lys57 NH_3^+ group by the oxygen-to-sulfur substitution is estimated to be $1.0 \text{ cal K}^{-1} \text{ mol}^{-1}$. Additionally, difference in rotational entropy (S_{rot}) of an NH_3^+ group can contribute to the entropic gain. As considered previously for CH_3 groups¹⁴⁰, the rotational entropy (S_{rot}) of an NH_3^+ group is indirectly related to the bond-rotation kinetics because the probability distribution function for the bond torsion angle depends on the energy barrier for rotation. Provided that the Eyring

equation is applicable to NH_3^+ rotation⁵⁵, experimental τ_f data along with the analytical expression¹⁴⁰ of S_{rot} suggest that NH_3^+ rotational entropy could increase by $\sim 0.3 \text{ cal K}^{-1} \text{ mol}^{-1}$ upon the oxygen-to-sulfur substitution in the interacting DNA phosphate group. The overall entropic gain arising from mobilization of the NH_3^+ group (i.e., reorientational + rotational) is thus estimated to be $\sim 1.3 \text{ cal K}^{-1} \text{ mol}^{-1}$. Relatively high mobility of the phosphorodithioate group, which is implicated by our crystallographic data (see Figure 3.9e,f), can further increase the entropic gain for the ion pair. Although the mobility of the intermolecular ion pairs may depend on ionic strength, a previous NMR study showed only weak ionic-strength dependence for the Arg side-chain dynamics in peptide-RNA complex³³. If the oxygen-to-sulfur substitution in DNA phosphate increased the ion-pair mobility to a similar degree at 20 and 150 mM NaCl, this entropic effect would be comparable to the observed increase in binding entropy ($\Delta\Delta S = 2.5 \pm 0.5 \text{ cal K}^{-1} \text{ mol}^{-1}$) upon the dithioation.

3.4.3 Role of ion-pair dynamics in protein-DNA interactions

Although three-dimensional structures of macromolecular complexes provide insights into the enthalpic terms of the binding free energy for each complex, the entropic aspects of binding free energy are not immediately clear from the structures alone unless the dynamic aspects are studied. Our work presents such examples. The structural data alone would not predict higher affinities for the *R_P*-monothioated and dithioated DNA complexes. Due to the absence of static interactions between Lys57 and DNA, one may even predict weaker affinities for these complexes. However, their affinities are actually stronger due to the entropic gain. The enhanced dynamics of the intermolecular ion pair

can at least partly account for the affinity enhancement. Our study suggests that the ion-pair dynamics plays a role in protein-DNA association.

3.4.4 Entropic gain due to mobilization of the intermolecular ion pair

Our thermodynamic data clearly indicate that the affinity enhancement by the oxygen-to-sulfur substitution in DNA phosphate is entropy-driven. As described above, mobilization of the intermolecular ion pair seems to make a significant contribution to this entropic gain. Here, we consider other entropic effects that can contribute to the affinity enhancement.

The polyelectrolyte effect, which arises from release of condensed counterions upon DNA-association, could contribute to the entropic difference³⁰. The entropic term of this effect is given by $\Delta S_{PE} = -z\psi R \ln[M^+]$, where z is the number of DNA phosphates that interact with the protein; R is the gas constant; $[M^+]$ is the concentration of cations; and ψ is the number of the released cations per phosphate ($\psi \leq 1$), which is given as a function of the charge and axial phosphate distance along DNA^{30, 141}. Because the parameter z is identical for the unmodified and dithioated DNA complexes and only a single phosphate group is substituted to phosphorodithioate in the dithioated DNA, the entropic change in the polyelectrolyte effect upon the oxygen-to-sulfur substitution ($\Delta\Delta S_{PE}$) is given by $-(\psi_S - \psi_O)R \ln[M^+]$, in which ψ_O and ψ_S are for phosphate and phosphorodithioate, respectively. Because $\psi_O = 0.88$ for unmodified DNA¹⁴¹ and $\psi_S \leq 1$, the upper limit of $\Delta\Delta S_{PE}$ at 0.15 M NaCl is calculated to be $0.4 \text{ cal K}^{-1} \text{ mol}^{-1}$. Actual $\Delta\Delta S_{PE}$ could even be negative, because the weaker charge density of phosphorodithioate could cause $\psi_S < \psi_O$.

Thus, it is very unlikely that the polyelectrolyte effect contributes to the entropic enhancement of the affinity by the oxygen-to-sulfur substitution.

The hydrophobic effect could also contribute to the entropic enhancement of the affinity. Difference in the hydrophobic effect between the unmodified and dithioated DNA complexes may arise from different degree of solvent-exposure around the modification site. The entropic term in $\text{cal K}^{-1} \text{mol}^{-1}$ units for this effect can be roughly estimated by using an empirical equation of Spolar and Record³⁰, $\Delta S_{HE} = 0.32\Delta A_{np} \ln(T/386)$, where ΔA_{np} , the total change in accessible surface areas (in \AA^2) of nonpolar atoms due to binding; and T , temperature (in K). Because Lys57 of the dithioated complex is more exposed to solvent (i.e., smaller ΔA_{np}) in Structure 1 as shown in Figure 3.9, this entropic effect should be smaller for the dithioated complex, and therefore cannot account for our entropic data. Strictly speaking, however, more accurate assessment of the hydrophobic effect requires quantitative information on desolvation entropy for both phosphate and phosphorodithioate groups, though qualitatively, the lower charge density of phosphorodithioate group should weaken its interactions with water¹⁵.

3.4.5 Implications of transcription factor decoys for therapeutic applications

Since Morishita *et al.* used short DNA duplexes as decoys to inhibit the transcription factor E2F for therapeutic applications in 1995,¹⁴² the synthetic decoy DNA strategy has been examined for various transcription factors (reviewed in Refs.^{143, 144}). Typically, to increase stability against nucleases, all phosphate groups in the decoy DNA are replaced with phosphoromonothioate groups, each of which involves R_P and S_P diastereomers. Our data suggest that for monothioation of each phosphate that forms an intermolecular ion pair, only one of the two diastereomers of the DNA phosphoromonothioate can

significantly enhance binding affinity. For a system involving a total of 6 intermolecular ion pairs, only 1 in 64 DNA molecules exhibits the strongest affinity. The use of dithioate at these positions should resolve this problem. However, dithioation of all phosphates in DNA is known to significantly shift DNA conformation from the canonical B-form to the A-form-like conformation.^{145, 146} A practical strategy to design high-affinity synthetic decoy DNA could be the selective use of dithioate only at the ion-pair sites and monothioate at the other sites.

CHAPTER 4

Towards Elucidation of the Relationship Between Ion-Pair Dynamics and Molecular Kinetics

4.1 INTRODUCTION

The interactions of DNA-binding protein, such as transcription factors, are important for the initiation of many cellular functions. Such interactions require the protein to efficiently navigate the milieu of the cellular nucleus in order to find its target DNA sequence. Typically, these DNA-binding proteins must be able to recognize a 6-10 base pair sequence in a very dense concentration of DNA^{147, 148}. Moreover, the process of recognizing the target sequence by the DNA-binding proteins may be slowed by nonspecific binding, or even trapping by quasispecific, non-functional sites on a DNA sequence¹⁴⁹. Due to these phenomena, DNA-binding proteins must be able to efficiently dissociate from non-target DNA sites in order to effectively locate target sequences.

For protein-DNA interactions, the association of transcription factors to DNA is facilitated by ion-pair formation between basic side chains and the phosphate backbone. Ion pair interactions of protein-DNA complexes were observed from structural studies via X-ray crystallography². More importantly, ion pairs have been shown to be in dynamic equilibrium between two different states known as the contact ion-pair (CIP) and the solvent-separated ion-pair (SIP)¹⁸⁻²⁰. Through molecular dynamic simulations, our group determined the transition between the CIP-SIP states was on a ps-ns timescale²⁴. This rapid transition may facilitate dissociation of the protein from the DNA backbone through breaking of the CIPs and allow translocation along the DNA. Furthermore, studies on the

translocation of the HoxD9 homeodomain on nonspecific DNA revealed the timescale for translocation of the homeodomain was comparable to the timescale to break all CIPs of the homeodomain⁴³. Thus, ion-pair dynamics may play an important role in the kinetics of protein target search.

For DNA target search by DNA binding proteins such as transcription factors, there are four classical mechanisms for the efficient scanning of DNA^{150, 151}. In these processes the proteins must first encounter DNA through 3-dimensional diffusion. Following this, other processes such as sliding or 1-dimensional diffusion, microscopic dissociation/association (hopping), and intersegment transfer can occur to allow efficient target search. In order to determine the role of ion-pairs on the kinetics of protein-target search, these processes should be investigated. In this study, we establish an experimental approach to study the kinetic parameters of the Antennapedia homeodomain (Antp) using a macromolecular crowder (Ficoll PM70) and fluorescence-based stopped-flow experiments. Ficoll is typically used in molecular crowding studies^{44, 45}. With this approach, we determine the sliding length of the Antp homeodomain and the impact of Ficoll PM70 on Antp sliding length. Theoretical work suggests that protein dissociation may be slowed in the presence of macromolecular crowders¹⁵². Therefore, we also consider the impact of a macromolecular crowder (Ficoll) on 3-dimensional diffusion of the Antp homeodomain. While we use fluorescence-based stopped-flow experiments to investigate 1-D diffusion, we investigate the impact of Ficoll as a macromolecular crowder on 3-D diffusion of through NMR spectroscopy. Information on the diffusion parameters of the Antp homeodomain and DNA duplex are also presented in this work. Thus, the experimental

approach established in this chapter will facilitate further studies on the role of ion-pairs in protein target search.

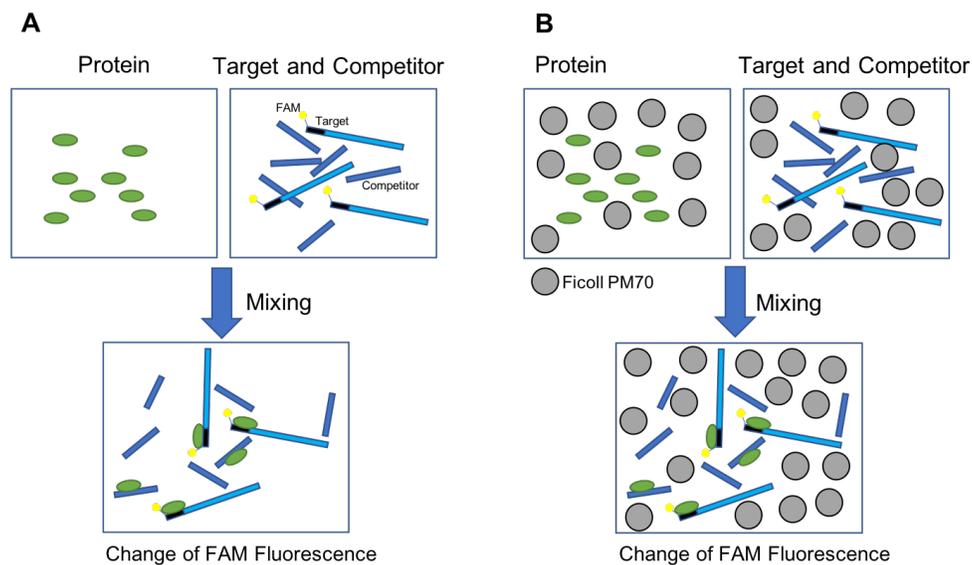


Figure 4.1: Experimental design of the using stopped-flow fluorescence kinetic assays. **(A)** Stopped-flow fluorescence assay for target search kinetics in the absence of Ficoll. **(B)** Stopped-flow fluorescence assay for target search kinetics in the presence of Ficoll (75 mg/ml).

4.2 MATERIALS AND METHODS

4.2.1 Preparation of the Antp Homeodomain protein

The 60 amino-acid construct of the fruit fly Antp homeodomain with a C39S mutation was expressed in the BL21 (DE3) strain of *E. coli* as previously described⁵⁰. Following expression, the protein was purified using SP fast-flow cation exchange, S-100 size-exclusion, and Resource-S cation exchange columns (GE Healthcare)^{24, 70, 74, 153}. The protein was quantified using UV light absorbance at 280 nm and an extinction coefficient of 15,470 M⁻¹ cm⁻¹ (<http://web.expasy.org/protparam/>). The protein was then buffer exchanged to a buffer containing 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 mM NaF (as a preservative) using an Amicon Ultra-4 centrifugal device.

4.2.2 Preparation of fluorescence-labeled probe DNA

A 33-mer single-stranded DNA with a FAM label attached to the 5'-terminus was purchased (Integrated DNA Technologies, Inc.) to prepare the DNA duplexes shown in Figure 4.2. The first 10 bases contain the 6-bp target sequence of the Antp homeodomain. The reverse primers and 100-mer template, named AntPSlides, were also purchased through Integrated DNA Technologies, Inc. The 33-bp duplex was prepared via annealing equimolar amounts of complimentary single-stranded DNA. Through PCR amplification using Vent DNA polymerase (New England), the 33-mer primer with a fluorescein amidite label (FAM), a reverse primer, and AntPSlides as the template, the 64-, 88-, 116-, and 145-bp DNA duplexes were prepared. The 46-bp duplex was prepared via hybridization of the FAM labeled 33-mer and single-stranded DNA, followed by base filling with DNA polymerase. Except for the 33-bp duplex, all duplexes were purified first, by a Resource Q anion-exchange column (GE Healthcare) and eluted on a gradient of 0-1.5 M NaCl in a

buffer containing 50 mM Tris-HCl (pH 7.5) and 1 mM ethylenediaminetetraacetic acid. The fractions containing the desired reaction product were purified via polyacrylamide gel electrophoresis (PAGE) using 4-20% gradient polyacrylamide/TBE gels (Invitrogen). The 33-bp duplex was also purified through PAGE purification. The bands of the desired products were then excised, crushed, and shaken in ultrapure water for ~16-18 hours to extract the DNA from the gel. All extracted DNA were purified using a PCR purification kit (Qiagen). All probes were buffer exchanged to a buffer containing 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 mM NaF (as a preservative) using an Amicon Ultra-4 centrifugal device.

4.2.3 Competitor DNA

Individual strands of a 15-bp nonspecific competitor DNA duplex used for the stopped-flow kinetic assays and NMR-based diffusion experiments were purchased through Integrated DNA Technologies, Inc. The individual sequences were purified using a Mono Q anion-exchange column (GE Healthcare) and GE ÄKTA purifier system, and the purified complimentary strands were annealed to form the duplex. The duplex was again isolated using the Mono Q anion-exchange chromatography. The sequences of the duplex can be seen in Figure 4.2A.

4.2.4 Stopped-flow fluorescence-based target association kinetics experiments

Target search kinetics of the Antp homeodomain were measured with an Applied Photophysics SX20-LED stopped-flow spectrofluorometer at 20°C. For these experiments, two solutions were rapidly mixed in a 1:1 volume (~0.5 ml) ratio by the stopped-flow device. The first solution contained 50 nM of the Antp homeodomain, while the second

solution consisted of 2.5 nM FAM-labeled probe DNA and 1000 nM competitor DNA. Both solutions were made with buffer containing 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 mM NaF (as a preservative). Immediately after flow for mixing stopped, time course data for fluorescence intensity were collected over a period over 15-40 s with time intervals ranging from 0.02 to 0.05 s. The FAM labeled probe was excited at 470 nm with a light-emitting diode (LED), and emission light was recorded using a long pass filter with a cut-off at 515 nm. All experiments were replicated 8-10 times.

For target search kinetics experiments of the Antp homeodomain in the presence of a macromolecular crowder, a stock solution of 100 mg/ml of Ficoll PM70 (Sigma-Aldrich) in a buffer containing 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 mM NaF was used. The protein and DNA solutions described above were then made with the Ficoll solution and buffer to a final concentration of 75 mg/ml Ficoll in 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 NaF mM. Experiments were performed as described above and replicated 8-10 times.

The concentrations described for the probe DNA (D_{tot}), competitor (C_{tot}), and protein (P_{tot}) were chosen to satisfy $D_{tot} \ll P_{tot} \ll C_{tot}$. This condition was used to create a pseudo-first-order reaction to simplify kinetic analyses. Apparent pseudo-first-order rate constants k_{obs} were determined from experimental data through non-linear least-squares fitting using MATLAB software (MathWorks) with $I(t) = (I_0 - I_\infty)\exp(-k_{obs}t) + I_\infty$, where I_0 and I_∞ represent intensities at time zero and infinite time, respectively.

Because two visible exponential processes were observed in the kinetic experiments for the samples tested in the presence of Ficoll, a biexponential fitting was used to calculate the observed rate constant k_{obs} using MATLAB software (MathWorks)

$A = A_1 e^{-k_1 t} + A_2 e^{-k_2 t} + C$, where the total change in fluorescence intensity was represented by two exponential terms to determine k_1 and k_2 , the rate constants for the first and second change in intensity. In this fitting, two rate constants are determined. The initial fluorescence intensity change was chosen to calculate the initial observed rate constant. The sliding length of the Antp homeodomain was estimated from the rate constants k_{obs} determined for each length of probe DNA using a global fitting adapted from Esadze, et al¹⁵⁴. Control experiments on the probe DNA in the presence of Ficoll revealed that the 2nd rate constant observed in our target association experiments were caused by photobleaching effects.

4.2.5 NMR sample preparation

To determine the translational and rotational diffusion parameters of the protein, NMR samples of free ¹⁵N-labeled Antp homeodomain were prepared in both the presence and absence of Ficoll PM70. For the protein sample tested in the absence of Ficoll, ¹⁵N-labeled Antp homeodomain was exchanged to a buffer containing 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 mM NaF. The 370 μ l solution containing a final concentration of 0.4 mM free ¹⁵N-labeled Antp homeodomain protein, enriched with 2.5% (v/v) D₂O, was sealed in an NMR tube. For the protein sample in the presence of Ficoll, the free ¹⁵N-labeled Antp homeodomain was exchanged to a buffer containing 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 mM NaF. The concentrated protein sample was then diluted with a buffer containing 100 mg/ml Ficoll PM 70, 20 mM potassium succinate (pH 5.8), 100 mM KCl, and 0.4 mM NaF to a final concentration of 0.4 mM ¹⁵N-labeled Antp homeodomain and 75 mg/ml Ficoll PM70. The translational and rotational diffusion parameters were also determined for free 15-bp DNA. Using the 15-bp nonspecific

competitor DNA, NMR samples of the free 15-bp duplex were prepared in the same manner to the final concentration of 0.4 mM 15-bp DNA. For the DNA sample prepared in the presence of Ficoll PM70, the final concentration of Ficoll was 75 mg/ml.

4.2.6 NMR Spectroscopy

To determine the rotational diffusion coefficient D_r of free ^{15}N -labeled Antp homeodomain and free 15-bp DNA in the presence and absence of Ficoll PM70 (75 mg/ml), backbone ^{15}N longitudinal and transverse (R_1 and R_2) relaxation rates were measured on Bruker Avance III spectrometer at the ^1H frequency of 750 MHz. Using the rotational diffusion coefficient D_r , the rotational correlation time τ_r was calculated from the following equation $\tau_r = (6D_r)^{-1}$. The measurements for the free protein and free DNA samples were performed at pH 5.8 and 25°C using a cryogenic probe.

For the translational diffusion parameters, the Pulsed-field-gradient (PFG)-based NMR Diffusion Ordered Spectroscopy (DOSY)¹⁵⁵ experiment was recorded for the samples of free Antp homeodomain protein in the absence and presence of Ficoll PM70 (75 mg/ml). Specifically the bipolar pulse longitudinal eddy current delay (BPP-LED)^{155, 156} pulse sequence was applied to the samples on the Bruker Avance III spectrometer at the ^1H frequency of 750 MHz. Two-parameter fitting of the signal intensities as a function of gradient strength were used to determine the translational diffusion coefficient D_t ¹³ using $Intensity = p_2 e^{-p_2 \gamma^2 s^2 d^2 (\Delta - d/3 - \tau/2)}$ where γ is the nuclear gyromagnetic ratio, d is the duration of the pulsed field gradient, Δ is the time between the 1st and 3rd pulses, and τ is the gradient recovery delay.

4.3 RESULTS

In this study, we investigated the sliding length of the Antennapedia homeodomain (Antp) and the impact of a macromolecular crowder on the Antp homeodomain's sliding length. For this, we performed target search kinetics experiments by monitoring the change in fluorescence over time as a function of Antp binding, using various DNA probes in the presence and absence of Ficoll PM70. Through NMR spectroscopy experiments, we observed the impact of Ficoll PM70 on mobility of the Antp homeodomain and DNA duplex.

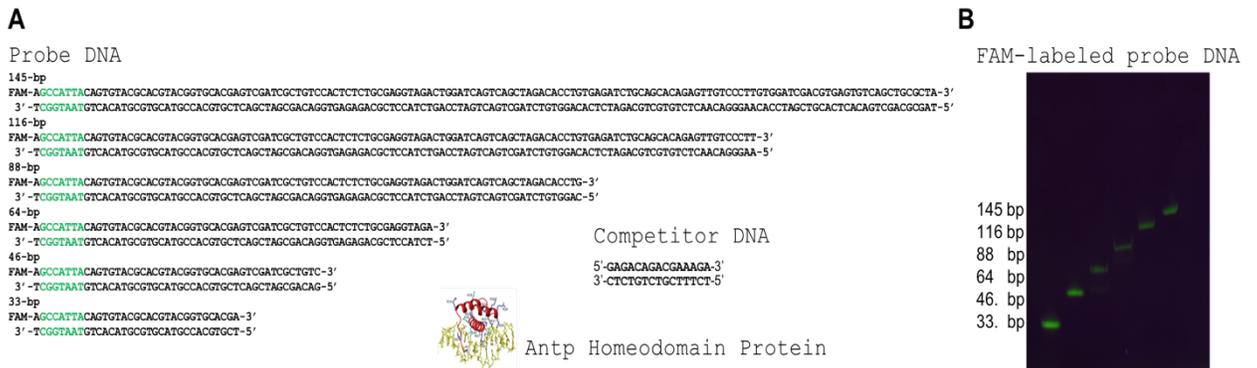


Figure 4.2: DNA duplexes used for kinetic studies on the Antp homeodomain. (A) Probe and competitor DNA duplexes used for the current study. The sequence of nucleotides in green represent the target sequence of the Antp Homeodomain. (B) PAGE of the purified duplexes, which were visualized by fluorescence arising from the covalently attached FAM.

4.3.1 Target association kinetics and determination of sliding length

As shown in Figure 4.1A, we mixed a solution containing protein with a solution containing probe and competitor DNA and recorded the time courses of FAM fluorescence intensity, immediately post-mixing with a stopped-flow device. For excitation of the FAM label, an LED with a maximum intensity at 470 nm was used. Emission light passing through a long pass filter with a cut-off at 515 nm was recorded. In these experiments, the conditions for the DNA duplexes and protein were $D_{tot} = 2.5$ nM, $C_{tot} = 1000$ nM, and $P_{tot} = 50$ nM in agreement with $D_{tot} \ll P_{tot} \ll C_{tot}$. The experiments were first run in the absence of a macromolecular crowder (Ficoll). With these conditions, we observed a mono-exponential change in fluorescence intensity, and attained pseudo-first-order rate constants k_{obs} , determined through mono-exponential fitting.

Sliding kinetics of the Antp homeodomain were analyzed using dependence of rate constant k_{obs} on various lengths of DNA. To determine rate constants, we prepared 33-, 46-, 64-, 88-, 116-, and 145-bp DNA duplexes each containing a covalently attached FAM label at the 5' terminal (Figure 4.2). This approach has been previously established by our group¹⁵⁴. As seen in Figure 4.3A, experimental k_{obs} data showed weak dependence on rate constants as a function of probe length. Qualitatively, this data suggests that the sliding length λ of the Antp homeodomain was very short. However, the weak dependence rendered sliding length λ quite difficult to quantitatively analyze.

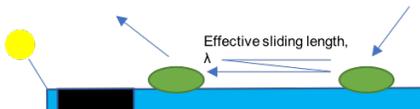
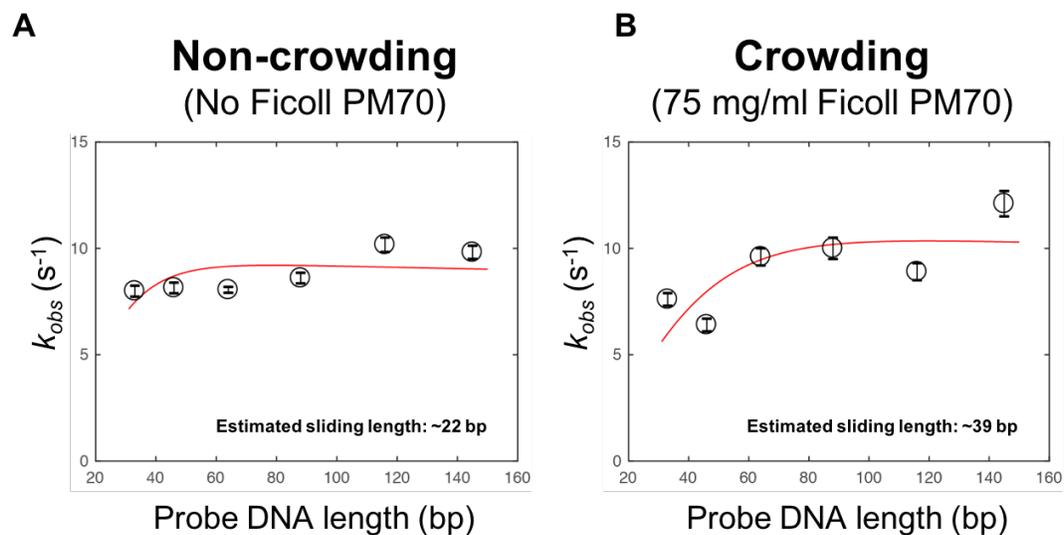


Figure 4.3: DNA length-dependence of target association kinetics measured for the Antp homeodomain protein in the presence and absence of 75 mg/ml Ficoll ($D_{tot} = 2.5$ nM, $P_{tot} = 50$ nM, $C_{tot} = 1000$ nM). The rate constant k_{obs} is plotted as a function of DNA length. The red line represents the best fit curve. The error bars represent the standard deviations for 8-10 replicates.

4.3.2 Impact of Ficoll on the Antp Homeodomain sliding length

In order to access more quantitatively analyzable DNA length dependence data, we employed a synthetic macromolecular crowder Ficoll PM70 (Figure 4.1B). Ficoll PM70 is often used as a macromolecular crowding agent due to its relatively globular shape and inert properties¹⁵⁷. In other words, Ficoll should not directly interfere with protein-DNA binding. These properties made Ficoll PM70 a good choice for our experiments. The same experiments were repeated as described in section 4.3.1 with the addition of 75 mg/ml Ficoll PM70. Again, the rate constants k_{obs} were determined for each duplex. In contrast to the length-dependent experiments run in the absence of Ficoll, we observed two visible exponential processes in the kinetic experiments. Therefore, bi-exponential fitting was used to appropriately determine rate constants k_{obs} . The equation used for bi-exponential fitting can be found in the Materials and Methods section. Figure 4.3B shows the length-dependence data for target association experiments of the Antp homeodomain in the presence of Ficoll. Using the same approach established by our group to determine sliding length, we estimated the sliding length of the Antp homeodomain in the presence of Ficoll to be ~39 bp, a two-fold increase from sliding length of the Antp homeodomain in the absence of Ficoll (Figure 4.3).

4.3.3 Impact of Ficoll on translational diffusion

To further understand how macromolecular crowders impact Antp sliding length, we evaluated the translational diffusion of Antp in the presence of Ficoll PM70. Translational diffusion represents the translocation of a molecule in solution and gives insight on the rate at which a molecule traverses from one point to another. To determine the translational diffusion coefficient D_t , we ran Pulsed-Field Gradient (PFG)-NMR Diffusion Ordered

Spectroscopy (DOSY) experiments on free Antp homeodomain and free 15-bp DNA duplex in both the absence and presence of Ficoll (75 mg/ml). Two samples of free Antp were made with a final concentration of 0.4 mM (3.0 mg/ml) in 75 mg/ml Ficoll and no Ficoll. Two samples of free 15-bp competitor DNA were also made in the same manner. In these experiments, the translational diffusion coefficient D_t is derived from a best-fit line for the change in normalized signal intensity as a function of a pulsed magnetic field, which is applied over a specific duration of time. We first compared the translational diffusion coefficient of the free Antp homeodomain protein in the presence ($D_t = (1.75 \pm 0.02) \times 10^{-6} \text{ cm}^2\text{s}^{-1}$) and absence of Ficoll ($D_t = (1.15 \pm 0.02) \times 10^{-6} \text{ cm}^2\text{s}^{-1}$) (Figure 4.4A and B). The difference in translational diffusion rates suggests a slower translational diffusion for the free Antp homeodomain in the presence of Ficoll. Similarly, slower diffusion was also observed for the free 15-bp DNA in the presence of Ficoll ($D_t = (1.00 \pm 0.01) \times 10^{-6} \text{ cm}^2\text{s}^{-1}$) compared to the absence of Ficoll ($D_t = (1.60 \pm 0.02) \times 10^{-6} \text{ cm}^2\text{s}^{-1}$) (Figure 4.4C and D). This data indicates that Ficoll PM70 causes slower translational diffusion for both the Antp homeodomain protein as well as the 15-bp DNA duplex.

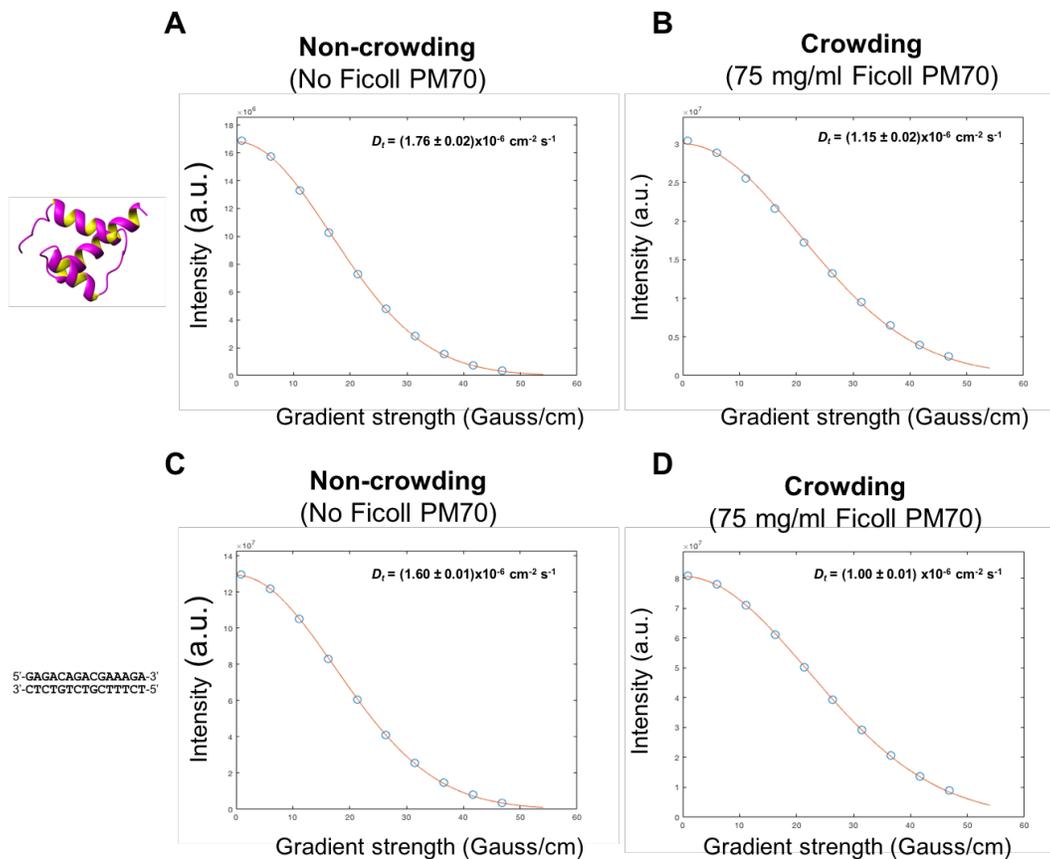


Figure 4.4: Translational diffusion measurements of free Antp homeodomain protein (0.4 mM; 3.0 mg/ml) and free 15-bp nonspecific DNA (0.4 mM; 4.0 mg/ml) in the presence and absence of Ficoll PM70 (75 mg/ml). Signal intensities were plotted as a function of gradient field strength for the Antp homeodomain in the absence (A) and presence (B) of Ficoll. Signal intensities were also plotted as a function of gradient field strength for 15-bp nonspecific DNA in the absence (C) and presence (D) of Ficoll. All experiments were run at pH 5.8 and 25°C at the ^1H frequency of 750 MHz.

4.3.4 Impact of Ficoll on rotational diffusion

While translational diffusion gives information on displacement of a molecule in solution, rotational diffusion gives information on the Brownian rotation of the molecule. Therefore, the impact of Ficoll on the rotational diffusion of Antp should also be considered using the rotational correlation time τ_r . The rotational correlation time τ_r can be related to the reorientation time of a molecule¹⁵⁸. In other words, the association of two molecules, such as protein and DNA, could be affected by the reorientation time for the molecules' interaction sites to properly align. To determine the rotational correlation time τ_r , backbone ¹⁵N longitudinal and transverse (R_1 and R_2) relaxation experiments were recorded for both samples of the free Antp homeodomain in the presence and absence of Ficoll PM70. For both R_1 and R_2 experiments, relaxation rates are shown for each residue (in Figure 4.5) of the Antp homeodomain in the presence and absence of Ficoll PM70. Data from these experiments gave a reorientation correlation time τ_r of 4.1 ns for the free Antp homeodomain in the absence of Ficoll, and a reorientation correlation time τ_r of 5.3 ns for the protein in the presence of Ficoll PM70 (75 mg/ml) (seen in Figure 4.5). These results suggest rotational diffusion is also slowed in the presence of Ficoll PM70.

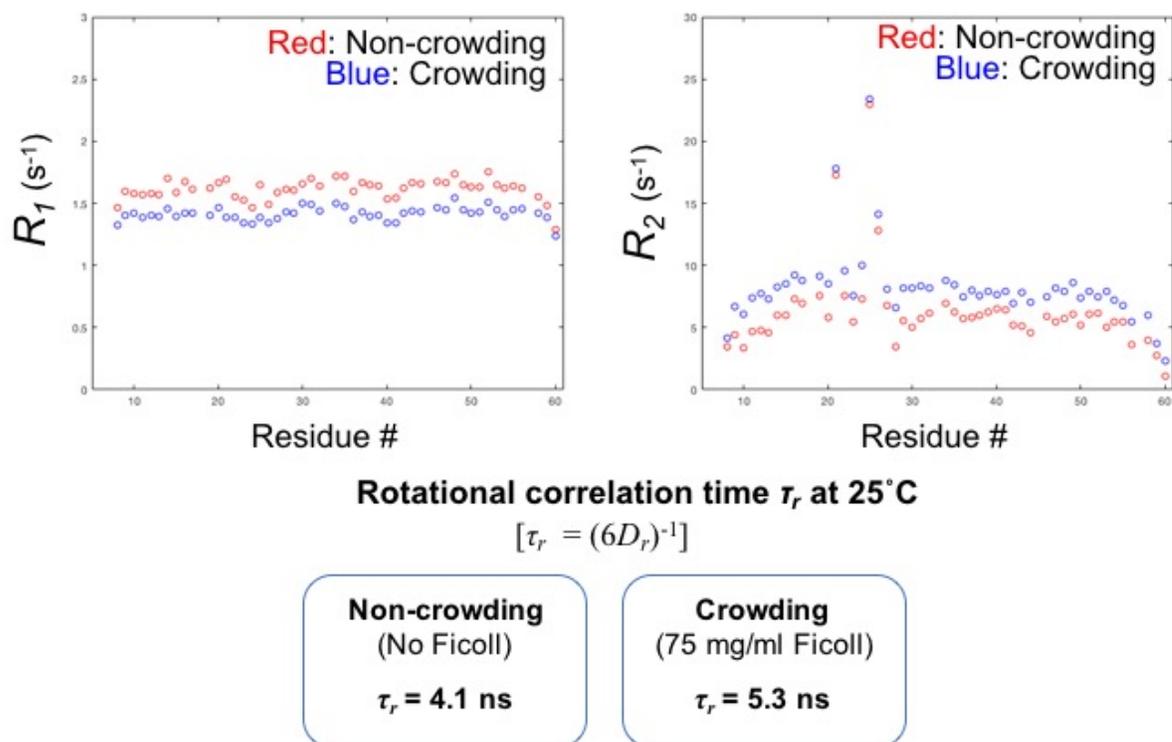


Figure 4.5: Determination of the rotational correlation coefficient for the Antp homeodomain protein in the presence and absence of Ficoll PM70 (75 mg/ml). ^{15}N R_1 and R_2 relaxation experiments used to determine the rotational correlation time τ_r . The relaxation rates are plotted as a function of individual residues. The blue circles represent relaxation rates for residues of the Antp homeodomain in the presence of Ficoll. The red circles represent relaxation rates for residues of the Antp homeodomain in the absence of Ficoll. All experiments were run at pH 5.8 and 25°C at a ^1H frequency of 750 MHz.

4.4 DISCUSSION

Overall, this study established the use of Ficoll as a macromolecular crowder to improve experiments to determine the kinetic parameters of the wild-type Antp homeodomain. While the previous studies in this dissertation work gave insight on the thermodynamic role of ion-pairs in protein-DNA interactions, we also wanted to determine the role of ion-pair interactions in regards to protein search. Because ion pair formations are important for facilitating protein-DNA binding, the breaking of these ion pairs would result in dissociation of the protein. The rate of protein dissociation from nonspecific DNA is an important parameter in the translocation kinetics of protein search¹⁵⁹. In fact, from NMR based line function analyses, our group found that the timescale of protein translocation on nonspecific DNA by the HoxD9 homeodomain protein was comparable to the timescale to break all CIPs simultaneously⁴³. Thus, the timescale of simultaneously breaking all ion pairs may play an important role in the kinetics of protein search.

Through quantitative analyses using fluorescence-based stopped-flow experiments, our objective was to determine the kinetic parameters of Antp homeodomain DNA search. This approach would aid in gaining more information on the sliding length and 1-D diffusion coefficient of the Antp homeodomain. We first studied the length dependence of the Antp homeodomain using various lengths of probe DNA labeled with a FAM fluorophore. However, due to the weak length-dependence data, the sliding length of the Antp homeodomain was difficult to analyze quantitatively. Theoretical work by other groups suggested that dissociation of protein from DNA would slow in the presence of a macromolecular crowder and potentially increase sliding length¹⁵². In the protein-DNA search process, a combination of 3-dimensional diffusion and 1-dimensional diffusion

allows for rapid search of the DNA target site. In the presence of a macromolecular crowder, protein dissociation may become less energetically favorable, allowing 1-D search processes such as sliding to become more efficient. Therefore, we introduced a synthetic macromolecular crowder (Ficoll PM70) into our system. Our results indicated a stronger length-dependence for the Antp homeodomain in the presence of Ficoll PM70, which gave more quantitatively analyzable data to determine the sliding length of Antp. The resultant length-dependence experimental data suggested that sliding length indeed increased by ~ 2 fold. It should be noted that the length-dependence data attained from target-association stopped-flow experiments in the presence of Ficoll required a bi-exponential fitting model. This was due to the presence of two observable exponential processes, one of which was caused by a photobleaching effect based on control experiments on probe DNA to determine the effect of photobleaching in the presence of Ficoll PM70. Furthermore, as seen in our fitting data for length-dependence in the presence of Ficoll, a bi-exponential fitting may cause a larger error and affect fitting. Therefore, it is pertinent to optimize the experimental conditions to further improve data quality.

As previously mentioned, the recognition of a target DNA site by DNA binding proteins occurs through a combination of 1-D and 3-D diffusion. The combination of these processes allows for more efficient scanning. We effectively increased the sliding length of Antp by ~ 2 -fold through the addition of Ficoll to our target-association kinetics experiments, which suggests that 1-D diffusion was enhanced in the presence of Ficoll. In order to determine the effect of Ficoll on 3-D diffusion, we investigated the translational and rotational diffusion parameters of free Antp as well as the free DNA duplex. To do this, we performed NMR experiments to determine the diffusion parameters of the Antp

homeodomain and DNA duplex in the presence and absence of Ficoll. For both the translational and rotational diffusion parameters, we observed overall slower diffusion for the free Antp homeodomain protein in the presence of Ficoll. Rotational diffusion was also slower for the free DNA duplex in the presence of Ficoll compared to free DNA in the absence of Ficoll. Collectively, these data indicated that Ficoll slowed 3-dimensional diffusion of the protein and DNA in solution.

More insight on the role of ion pairs can be attained through individual site mutations of basic residues of the Antp homeodomain that form ion pairs with the DNA phosphate backbone. In fact, our group has already developed and prepared two Antp homeodomain mutants (K57A and K46A) that result in a reduced number of total ion pairs in the Antp homeodomain-DNA complex. We hypothesize that the reduced number of ion pairs would result in faster translocation. As seen in this work, the wild-type Antp homeodomain was found to have a short sliding length. Thus, studies on mutants of the Antp homeodomain would be quite difficult to execute without the aid of Ficoll. Moreover, further examination on the different mechanisms of target search on DNA in the presence of a macromolecular crowder should also be considered because macromolecular crowding facilitates a more *in vitro*-like environment¹⁶⁰⁻¹⁶³. Of the four classical mechanisms of target search on DNA, we can surmise that 3-D diffusion is slowed and 1-D diffusion is enhanced in the presence of Ficoll. However, microscopic dissociation (hopping) and intersegment transfer should also be considered in order to grasp a fuller picture of DNA target search by Antp in the presence of a crowder. This work will not only improve our studies on the role of ion-pairs in protein target search by the Antp homeodomain, we also establish a good experimental approach to examine the effect of macromolecular crowding on protein target search.

4.5 FUTURE DIRECTIONS

4.5.1 Further delineating the role of ion pair dynamics on Protein-DNA search kinetics

In Chapter 4, we established the use of Ficoll as an approach to improve the study of kinetic parameters for the wild-type Antp homeodomain. However, the role of ion pairs in protein sliding still remains to be elucidated. In previous work from our group on the sliding of the HoxD-9 homeodomain, the timescale of sliding for HoxD-9 was determined through NMR line shape analysis⁴³. This data suggested that the timescale to simultaneously break all CIP's of the homeodomain was comparable to the timescale of sliding along the DNA. To further examine this relationship, ion pairs should be systematically examined for the Antp homeodomain. For example, mutations of basic side chains to prevent ion-pair formation would reduce the total number of ion pairs and may result in faster sliding. Other ion pair mutations to perturb the CIP-SIP equilibrium should also be used to observe a change in sliding kinetics (Figure 4.6). This strategy will result in a deeper understanding of the role of ion pairs on kinetic parameters of the Antp Homeodomain.

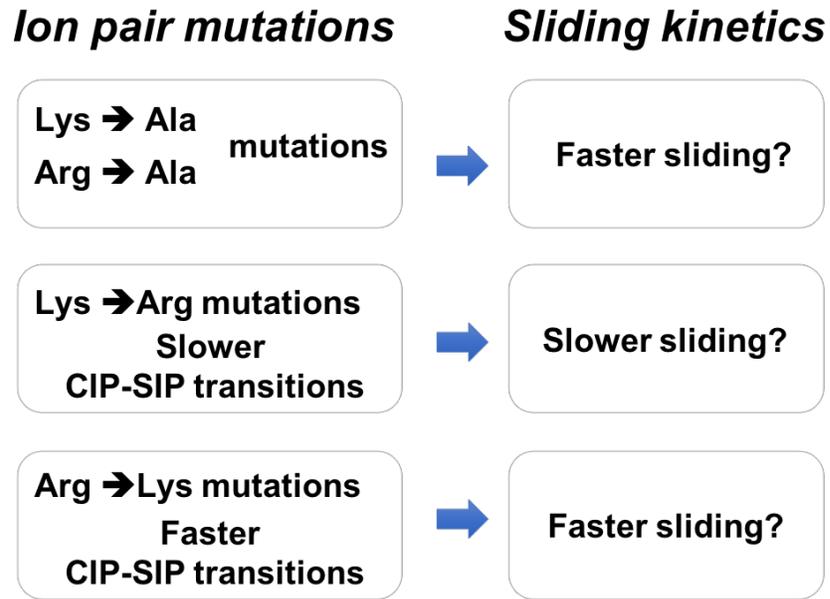


Figure 4.6: A schematic of ion pair mutations to study the role of ion pairs in protein sliding kinetics.

CHAPTER 5

Conclusions

5.1 MAJOR FINDINGS IN THIS WORK

The interactions of small ion pairs have been studied for some time¹, yet the importance of ion pairs for biological macromolecular interactions still required further investigation (especially for the dynamics of ion pairs). In this project, we focused on improving our understanding of the role of ion-pair dynamics in protein-DNA interactions. By use of the Antp homeodomain transcription factor, we extensively studied ion pairs formed by the lysine and arginine basic side chains of the Antp homeodomain with the phosphate backbone of DNA.

In Chapter 2, we compared the side chain dynamics of lysine and arginine side chains of the Antp homeodomain in the free and DNA-bound states. In previous work performed on the Egr-1 zinc-finger protein, the mobility of lysine and arginine side chains was largely retained for the basic side chains that form an ion pair with the phosphate of the DNA backbone³⁷. Using NMR relaxation experiments to determine the S^2 order parameters of the basic side chains, our studies demonstrated the same substantial retention of side chain mobility for the Antp homeodomain upon ion-pair formation with the DNA backbone. The retention of side chain mobility suggested a reduction in the overall loss of conformational entropy as a result. However, that this high mobility was retained upon ion pair formation was quite surprising due to hydrogen bond formation and strong short-ranged electrostatic interactions. Based on previous computational work and theoretical consideration, it is likely that the high mobility of basic side chains could be a result of the rapid dynamic

transitions between the CIP and SIP states of the ion pair²⁴. Furthermore, our studies in Chapter 2 suggested that the retained mobility could mitigate the entropic costs for protein-DNA binding, and this hypothesis was tested in the following chapter.

In Chapter 3, we tested whether an increase in ion-pair mobility could reduce the entropic cost for protein-DNA binding. To do this, studies on the impact of the oxygen-to-sulfur substitution (*a.k.a* phosphorodi- or monothioate) on Antp homeodomain binding were performed in order to facilitate a deeper understanding of the role of ion pair dynamics in protein-DNA interactions. For these studies, a single phosphate at the interaction site of Lys57 was modified with a sulfur substitution of one- or both of the non-bridging oxygen atoms. Through fluorescence-based binding assays, an enhanced binding affinity was observed for the Antp homeodomain binding to dithioated DNA. Isothermal titration calorimetry data indicated that this enhancement was not due to enthalpic effects, but rather an entropic contribution. By applying NMR relaxation experiments previously used in Chapter 2, we also studied the impact of the oxygen-to-sulfur substitution on side chain mobility. We observed Lys57 interacting with the dithioated DNA phosphate to be even more mobile than the unmodified DNA. Furthermore, our structural studies comparing the Antp homeodomain-DNA complexes with the dithioated and unmodified DNA phosphate suggested the presence of two distinctly different interactions for Lys57 with the dithioated phosphate, as opposed to the unmodified DNA, based on N...S distances. The different N...S distances observed in the dithioated DNA complex indicated that the Lys57 ion pair was found in both the CIP and SIP state, supporting the high mobility of Lys57 observed by NMR spectroscopy. Similar mobility was also observed for the R_p diastereomer of monothioate, which is reasonable considering the R_p diastereomer contained the sulfur

substitution that is more proximal to Lys57. Collectively, these data suggest that the enhancement in binding affinity was a result of entropic contributions arising from conformational dynamics.

Along with the thermodynamic role of ion pairs in protein-DNA interactions, the role of ion pairs on the kinetics of protein-DNA search should also be considered. In Chapter 4, we studied the impact of a macromolecular crowder on the kinetic parameters of the Antp homeodomain using Ficoll PM70. Using stopped-flow kinetic experiments, we tested several lengths of DNA containing the target sequence of the Antp homeodomain through length-dependence target association experiments to determine the sliding length and 1-D diffusion coefficient of the Antp homeodomain. This approach was previously established using the Egr-1 zinc finger protein^{149, 154}. However, unlike Egr-1, the Antp homeodomain showed very weak length dependence, rendering quantitative analysis of sliding for the Antp homeodomain quite difficult. In fact, the weak length dependence data suggested the sliding length of the Antp homeodomain was fairly short. In order to improve the quality of the length dependence data, we tested Antp homeodomain sliding in the presence of a macromolecular crowder, Ficoll PM70. In the presence of Ficoll, often used for macromolecular crowding studies^{44, 45, 164-166}, the Antp homeodomain sliding length significantly increased by about two-fold. Data from NMR-based diffusion experiments to determine the translational and rotational diffusion coefficients of the Antp homeodomain and nonspecific DNA, in the presence and absence of Ficoll, indicated that both the Antp homeodomain and nonspecific DNA showed slower diffusion in the presence of Ficoll. The use of Ficoll showed to be a successful and practical approach to achieve analyzable length dependence data.

5.2 FUTURE PERSPECTIVES

5.2.1 Implications on transcription factor decoys for therapeutic applications⁵

Since Morishita *et al.* used short DNA duplexes as decoys to inhibit the transcription factor E2F for therapeutic applications in 1995,¹⁴² the synthetic decoy DNA strategy has been examined for various transcription factors (reviewed in Refs.^{143, 144}). Typically, to increase stability against nucleases, all phosphate groups in the decoy DNA are replaced with phosphoromonothioate groups, each of which involves R_P and S_P diastereomers. Our data suggest that for monothioation of each phosphate that forms an intermolecular ion pair, only one of the two diastereomers of the DNA phosphoromonothioate can significantly enhance binding affinity. For a system involving a total of 6 intermolecular ion pairs, only 1 in 64 DNA molecules exhibits the strongest affinity. The use of dithioate at these positions should resolve this problem. However, dithioation of all phosphates in DNA is known to significantly shift DNA conformation from the canonical B-form to the A-form-like conformation.^{145, 146} A practical strategy to design high-affinity synthetic decoy DNA could be the selective use of dithioate only at the ion-pair sites and monothioate at the other sites.

5.3 FINAL REMARKS: POTENTIAL IMPACT OF THIS WORK

In this work, we asked several questions to gain a deeper insight on the role of ion-pair dynamics in protein-DNA interactions. We addressed the thermodynamic significance of ion-pair dynamics in Chapters 2 and 3; first through studying the dynamics the basic side chains of the Antp homeodomain in the free and DNA-bound states, then through

⁵ Paragraph adapted with permission from John Wiley and Sons (see Appendix). Nguyen, D., et al. (2016) Stereospecific effects of oxygen-to-sulfur substitution in DNA phosphate on ion-pair dynamics and protein-DNA affinity. *ChemBioChem* 17, 1636-42.

mechanistically delineating the enhancement of the Antp homeodomain binding affinity by the oxygen-to-sulfur modification. We found that basic side chains of the Antp homeodomain forming an ion pair with the DNA backbone retained their conformational mobility when bound to DNA. Despite the strong short-ranged interactions and hydrogen bonds of ion-pairs, computational and theoretical work suggested the mobility was a result of rapid transitions between the CIP-SIP states and could potentially reduce the entropic costs of binding. In order to test this hypothesis, we used the oxygen-to-sulfur substitution, which served as a great tool to study ion-pair dynamics. The lysine side chain that formed an ion pair with the di- and monothioated DNA phosphate showed even higher mobility when compared to the unmodified DNA, suggesting the major entropic contribution came from conformational mobility. Unlike enthalpic changes, entropic effects of binding free energies are difficult to observe. However, our experimental approach facilitated a means to understanding entropic effects that occur as a result of ion-pair dynamics. Not only did we determine that the enhancement in affinity by the oxygen-to-sulfur substitution was due to an entropic effect, but our approach mechanistically deduced that the substitution resulted in an increase in conformational dynamics, rather than other entropic effects arising from the polyelectrolyte or hydrophobic effects. Also, while dithioation of DNA oligonucleotides have been studied for therapeutic applications^{40, 41}, our work suggests a more strategic approach to appropriately modifying DNA phosphates of the oligonucleotides that form ion-pairs in order to achieve greater affinities.

To help address the role of ion pairs on the kinetics of protein search, we analyzed the kinetic parameters of the Antp homeodomain sliding in the absence and presence of a macromolecular crowder. With this approach, we were able to attain quantitatively

analyzable length dependence data to determine the sliding length of the wild-type form of the Antp homeodomain. Further studies on the mutants of the Antp homeodomain can now be achievable. Although the practical intention of introducing a macromolecular crowder to our system was to improve quantitative data analysis, the use of a synthetic macromolecular crowder also facilitates a more *in vitro*-like environment. There are very few experimental studies on the effects of macromolecular crowding. However, to grasp a biologically relevant understanding of protein interactions, the crowded environment of the cell should be deeply considered.

Although some progress has been made toward understanding the ion pair, this work deepens our knowledge on how ion pair dynamics affect protein-DNA interactions in this work. Our previous work confirmed that the ion pair can dynamically transition between the CIP and SIP states, and the dynamics of lysine and arginine side chains could be observed by NMR spectroscopy^{24,36,37,42,58}. The current work in this dissertation will serve to explain the mechanism behind the retained (or enhanced) mobility upon protein-DNA association. We also demonstrate how side chain mobility contributes to conformational entropy, and how this entropic contribution impacts the binding free energy for the Antp homeodomain-DNA complex.

Our work will aid in advancing the drug design process to benefit human health. We demonstrate the importance of short-ranged electrostatic interactions for protein-DNA complexes by using the Antp homeodomain-DNA complex. While prior knowledge was limited to static x-ray crystallography studies, NMR spectroscopy allows us to improve our understanding of the atomic-level mechanisms that govern protein function through experiments that provide insight on ion pair dynamics. This experimental data can be used

to validate force field parameters, which will improve *in silico* drug design methods and will increase the efficiency of pharmaceutical research and development. In fact, NMR data on Lys side chains were used for validation of CHARMM36 force field¹⁶⁷.

Oligonucleotides have been used as potential human therapeutics, but they often lack a strong enough affinity to show efficacy¹⁶⁸⁻¹⁷⁰. One approach to overcome this issue would be to increase the concentration of oligonucleotides that will bind to the targeted proteins. However, this strategy is impractical and costly, and other means for improving oligonucleotide binding through modifications can be used. We demonstrate the advantage of oligonucleotide modifications for increasing protein-DNA binding affinities using the oxygen-to-sulfur substitution. Through understanding the mechanism behind the increased affinity for protein-DNA binding, better strategies can be developed to improve nucleic acid therapeutics and develop better drugs.

Appendix

4/23/2018

Rightslink® by Copyright Clearance Center



RightsLink®

[Home](#)[Create Account](#)[Help](#)

ACS Publications
Most Trusted. Most Cited. Most Read.

Title: Internal Motions of Basic Side Chains of the Antennapedia Homeodomain in the Free and DNA-Bound States

Author: Dan Nguyen, Zoe A. Hoffpauir, Junji Iwahara

Publication: Biochemistry

Publisher: American Chemical Society

Date: Nov 1, 2017

Copyright © 2017, American Chemical Society

LOGIN

If you're a **copyright.com user**, you can login to RightsLink using your copyright.com credentials. Already a **RightsLink user** or want to [learn more?](#)

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:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)[CLOSE WINDOW](#)

Copyright © 2018 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement.](#) [Terms and Conditions.](#) Comments? We would like to hear from you. E-mail us at customercare@copyright.com



RightsLink®

Home

Account
Info

Help



Title: Entropic Enhancement of Protein-DNA Affinity by Oxygen-to-Sulfur Substitution in DNA Phosphate

Author: Levani Zandarashvili, Dan Nguyen, Kurtis M. Anderson, Mark A. White, David G. Gorenstein, Junji Iwahara

Publication: Biophysical Journal

Publisher: Elsevier

Date: 1 September 2015

Copyright © 2015 Biophysical Society. Published by Elsevier Inc. All rights reserved.

Logged in as:

Dan Nguyen

LOGOUT

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW

Copyright © 2018 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#). Comments? We would like to hear from you. E-mail us at customercare@copyright.com

**JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS**

May 25, 2018

This Agreement between Dan Nguyen ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4345510105242
License date	May 10, 2018
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	ChemBioChem
Licensed Content Title	Stereospecific Effects of Oxygen-to-Sulfur Substitution in DNA Phosphate on Ion Pair Dynamics and Protein-DNA Affinity
Licensed Content Author	Dan Nguyen, Levani Zandarashvili, Mark A. White, Junji Iwahara
Licensed Content Date	Jul 13, 2016
Licensed Content Pages	7
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Biophysical Analyses of Intermolecular Ion-Pairs in Protein-DNA Complexes
Expected completion date	Jun 2018
Expected size (number of pages)	100
Requestor Location	Dan Nguyen 15803 Hermitage Oaks Dr TOMBALL, TX 77377 United States Attn: Dan Nguyen
Publisher Tax ID	EU826007151
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at <http://myaccount.copyright.com>).

Bibliography

1. Marcus, Y.; Hefter, G., Ion pairing. *Chemical reviews* **2006**, *106* (11), 4585-4621.
2. Nadassy, K.; Wodak, S. J.; Janin, J., Structural features of protein– nucleic acid recognition sites. *Biochemistry* **1999**, *38* (7), 1999-2017.
3. Jones, S.; Daley, D. T.; Luscombe, N. M.; Berman, H. M.; Thornton, J. M., Protein–RNA interactions: a structural analysis. *Nucleic acids research* **2001**, *29* (4), 943-954.
4. McCoy, A. J.; Epa, V. C.; Colman, P. M., Electrostatic complementarity at protein/protein interfaces1. *Journal of molecular biology* **1997**, *268* (2), 570-584.
5. Chien, E. Y.; Liu, W.; Zhao, Q.; Katritch, V.; Han, G. W.; Hanson, M. A.; Shi, L.; Newman, A. H.; Javitch, J. A.; Cherezov, V., Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **2010**, *330* (6007), 1091-1095.
6. Wu, H.; Wacker, D.; Mileni, M.; Katritch, V.; Han, G. W.; Vardy, E.; Liu, W.; Thompson, A. A.; Huang, X.-P.; Carroll, F. I., Structure of the human κ -opioid receptor in complex with JDTic. *Nature* **2012**, *485* (7398), 327.
7. Bax, B. D.; Chan, P. F.; Eggleston, D. S.; Fosberry, A.; Gentry, D. R.; Gorrec, F.; Giordano, I.; Hann, M. M.; Hennessy, A.; Hibbs, M., Type IIA topoisomerase inhibition by a new class of antibacterial agents. *Nature* **2010**, *466* (7309), 935.
8. Luscombe, N. M.; Austin, S. E.; Berman, H. M.; Thornton, J. M., An overview of the structures of protein-DNA complexes. *Genome Biol* **2000**, *1* (1), REVIEWS001.
9. Rohs, R.; Jin, X.; West, S. M.; Joshi, R.; Honig, B.; Mann, R. S., Origins of specificity in protein-DNA recognition. *Annual review of biochemistry* **2010**, *79*, 233-269.
10. Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M., Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res* **2001**, *29* (13), 2860-74.
11. Privalov, P. L.; Dragan, A. I.; Crane-Robinson, C., Interpreting protein/DNA interactions: distinguishing specific from non-specific and electrostatic from non-electrostatic components. *Nucleic Acids Res* **2011**, *39* (7), 2483-91.
12. Dragan, A. I.; Li, Z.; Makeyeva, E. N.; Milgotina, E. I.; Liu, Y.; Crane-Robinson, C.; Privalov, P. L., Forces driving the binding of homeodomains to DNA. *Biochemistry* **2006**, *45* (1), 141-51.

13. Dragan, A. I.; Read, C. M.; Makeyeva, E. N.; Milgotina, E. I.; Churchill, M. E.; Crane-Robinson, C.; Privalov, P. L., DNA binding and bending by HMG boxes: energetic determinants of specificity. *J Mol Biol* **2004**, *343* (2), 371-93.
14. Dragan, A. I.; Liggins, J. R.; Crane-Robinson, C.; Privalov, P. L., The energetics of specific binding of AT-hooks from HMGA1 to target DNA. *J Mol Biol* **2003**, *327* (2), 393-411.
15. Collins, K. D., Charge density-dependent strength of hydration and biological structure. *Biophys J* **1997**, *72* (1), 65-76.
16. Fennell, C. J.; Bizjak, A.; Vlachy, V.; Dill, K. A.; Sarupria, S.; Rajamani, S.; Garde, S., Ion pairing in molecular simulations of aqueous alkali halide solutions. *J Phys Chem B* **2009**, *113* (44), 14837-8.
17. Iwahara, J.; Esadze, A.; Zandarashvili, L., Physicochemical Properties of Ion Pairs of Biological Macromolecules. *Biomolecules* **2015**, *5* (4), 2435-63.
18. Simon, J. D.; Peters, K. S., Picosecond dynamics of ion pairs: the effect of hydrogen bonding on ion-pair intermediates. *Journal of the American Chemical Society* **1982**, *104* (24), 6542-6547.
19. Simon, J. D.; Peters, K. S., Direct observation of the special salt effect: picosecond dynamics of ion-pair exchange. *Journal of the American Chemical Society* **1982**, *104* (22), 6142-6144.
20. Masnovi, J.; Kochi, J., Direct observation of ion-pair dynamics. *Journal of the American Chemical Society* **1985**, *107* (26), 7880-7893.
21. Masunov, A.; Lazaridis, T., Potentials of mean force between ionizable amino acid side chains in water. *Journal of the American Chemical Society* **2003**, *125* (7), 1722-1730.
22. Friedman, R. A.; Mezei, M., The potentials of mean force of sodium chloride and sodium dimethylphosphate in water: An application of adaptive umbrella sampling. *The Journal of chemical physics* **1995**, *102* (1), 419-426.
23. Rozanska, X.; Chipot, C., Modeling ion-ion interaction in proteins: a molecular dynamics free energy calculation of the guanidinium-acetate association. *The Journal of Chemical Physics* **2000**, *112* (22), 9691-9694.
24. Chen, C.; Esadze, A.; Zandarashvili, L.; Nguyen, D.; Montgomery Pettitt, B.; Iwahara, J., Dynamic Equilibria of Short-Range Electrostatic Interactions at Molecular Interfaces of Protein-DNA Complexes. *The journal of physical chemistry letters* **2015**, *6* (14), 2733-7.
25. Marcus, Y., Ionic volumes in solution. *Biophysical chemistry* **2006**, *124* (3), 200-207.

26. Wand, A. J., The dark energy of proteins comes to light: conformational entropy and its role in protein function revealed by NMR relaxation. *Curr Opin Struct Biol* **2013**, *23* (1), 75-81.
27. Manning, G. S.; Ray, J., Counterion condensation revisited. *J Biomol Struct Dyn* **1998**, *16* (2), 461-76.
28. Manning, G. S., Limiting laws and counterion condensation in polyelectrolyte solutions. V. Further development of the chemical model. *Biophys Chem* **1978**, *9* (1), 65-70.
29. Manning, G. S., Limiting laws and counterion condensation in polyelectrolyte solutions. IV. The approach to the limit and the extraordinary stability of the charge fraction. *Biophys Chem* **1977**, *7* (2), 95-102.
30. Spolar, R. S.; Record, M. T., Coupling of local folding to site-specific binding of proteins to DNA. *Science* **1994**, *263* (5148), 777-84.
31. Yang, D.; Kay, L. E., Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding. *J Mol Biol* **1996**, *263* (2), 369-82.
32. Trbovic, N.; Cho, J. H.; Abel, R.; Friesner, R. A.; Rance, M.; Palmer, A. G., 3rd, Protein side-chain dynamics and residual conformational entropy. *J Am Chem Soc* **2009**, *131* (2), 615-22.
33. Wilkinson, T. A.; Botuyan, M. V.; Kaplan, B. E.; Rossi, J. J.; Chen, Y., Arginine side-chain dynamics in the HIV-1 rev-RRE complex. *J Mol Biol* **2000**, *303* (4), 515-29.
34. Esadze, A.; Li, D. W.; Wang, T.; Brüschweiler, R.; Iwahara, J., Dynamics of lysine side-chain amino groups in a protein studied by heteronuclear ¹H-¹⁵N NMR spectroscopy. *J Am Chem Soc* **2011**, *133* (4), 909-19.
35. Li, Z.; Raychaudhuri, S.; Wand, A. J., Insights into the local residual entropy of proteins provided by NMR relaxation. *Protein Sci* **1996**, *5* (12), 2647-50.
36. Anderson, K. M.; Esadze, A.; Manoharan, M.; Brüschweiler, R.; Gorenstein, D. G.; Iwahara, J., Direct observation of the ion-pair dynamics at a protein-DNA interface by NMR spectroscopy. *J Am Chem Soc* **2013**, *135* (9), 3613-9.
37. Esadze, A.; Chen, C.; Zandarashvili, L.; Roy, S.; Pettitt, B. M.; Iwahara, J., Changes in conformational dynamics of basic side chains upon protein-DNA association. *Nucleic acids research* **2016**, *44* (14), 6961-70.
38. Yang, X.; Bassett, S. E.; Li, X.; Luxon, B. A.; Herzog, N. K.; Shope, R. E.; Aronson, J.; Prow, T. W.; Leary, J. F.; Kirby, R., Construction and selection of bead-bound combinatorial oligonucleoside phosphorothioate and

- phosphorodithioate aptamer libraries designed for rapid PCR-based sequencing. *Nucleic acids research* **2002**, *30* (23), e132-e132.
39. Yang, X.; Fennewald, S.; Luxon, B. A.; Aronson, J.; Herzog, N. K.; Gorenstein, D. G., Aptamers containing thymidine 3'-O-phosphorodithioates: synthesis and binding to nuclear factor- κ B. *Bioorganic & medicinal chemistry letters* **1999**, *9* (23), 3357-3362.
 40. Marshall, W.; Caruthers, M., Phosphorodithioate DNA as a potential therapeutic drug. *Science* **1993**, *259* (5101), 1564-1570.
 41. Marshall, W. S.; Beaton, G.; Stein, C. A.; Matsukura, M.; Caruthers, M. H., Inhibition of human immunodeficiency virus activity by phosphorodithioate oligodeoxycytidine. *Proceedings of the National Academy of Sciences* **1992**, *89* (14), 6265-6269.
 42. Zandarashvili, L.; Esadze, A.; Iwahara, J., NMR studies on the dynamics of hydrogen bonds and ion pairs involving lysine side chains of proteins. *Adv Protein Chem Struct Biol* **2013**, *93*, 37-80.
 43. Sahu, D.; Iwahara, J., Discrete-State Kinetics Model for NMR-Based Analysis of Protein Translocation on DNA at Equilibrium. *J Phys Chem B* **2017**, *121* (41), 9548-9556.
 44. Kumar, S.; Sharma, D.; Kumar, R., Role of Macromolecular Crowding on Stability and Iron Release Kinetics of Serum Transferrin. *J Phys Chem B* **2017**, *121* (37), 8669-8683.
 45. Gtari, W.; Bey, H.; Aschi, A.; Bitri, L.; Othman, T., Impact of macromolecular crowding on structure and properties of pepsin and trypsin. *Mater Sci Eng C Mater Biol Appl* **2017**, *72*, 98-105.
 46. McGinnis, W.; Levine, M. S.; Hafen, E.; Kuroiwa, A.; Gehring, W. J., A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature* **1984**, *308* (5958), 428-33.
 47. Müller, M.; Affolter, M.; Leupin, W.; Otting, G.; Wüthrich, K.; Gehring, W. J., Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. *EMBO J* **1988**, *7* (13), 4299-304.
 48. Affolter, M.; Percival-Smith, A.; Müller, M.; Leupin, W.; Gehring, W. J., DNA binding properties of the purified Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **1990**, *87* (11), 4093-7.
 49. Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A., The third helix of the Antennapedia homeodomain translocates through biological membranes. *J Biol Chem* **1994**, *269* (14), 10444-50.

50. Fraenkel, E.; Pabo, C. O., Comparison of X-ray and NMR structures for the Antennapedia homeodomain-DNA complex. *Nat Struct Biol* **1998**, *5* (8), 692-7.
51. Cummins, L.; Graff, D.; Beaton, G.; Marshall, W. S.; Caruthers, M. H., Biochemical and physicochemical properties of phosphorodithioate DNA. *Biochemistry* **1996**, *35* (26), 8734-41.
52. Wang, S.; Lee, R. J.; Cauchon, G.; Gorenstein, D. G.; Low, P. S., Delivery of antisense oligodeoxyribonucleotides against the human epidermal growth factor receptor into cultured KB cells with liposomes conjugated to folate via polyethylene glycol. *Proceedings of the National Academy of Sciences* **1995**, *92* (8), 3318-3322.
53. Wagner, R. W.; Flanagan, W. M., Antisense technology and prospects for therapy of viral infections and cancer. *Molecular medicine today* **1997**, *3* (1), 31-38.
54. Yang, X.; Hodge, R. P.; Luxon, B. A.; Shope, R.; Gorenstein, D. G., Separation of synthetic oligonucleotide dithioates from monothiophosphate impurities by anion-exchange chromatography on a mono-q column. *Anal Biochem* **2002**, *306* (1), 92-9.
55. Zandarashvili, L.; Iwahara, J., Temperature Dependence of Internal Motions of Protein Side-Chain NH_3^+ Groups: Insight into Energy Barriers for Transient Breakage of Hydrogen Bonds. *Biochemistry* **2015**, *54* (2), 538-45.
56. Esadze, A.; Zandarashvili, L.; Iwahara, J., Effective strategy to assign ^1H - ^{15}N heteronuclear correlation NMR signals from lysine side-chain NH_3^+ groups of proteins at low temperature. *J Biomol NMR* **2014**, *60* (1), 23-7.
57. Lipari, G.; Szabo, A., Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *Journal of the American Chemical Society* **1982**, *104* (17), 4546-4559.
58. Zandarashvili, L.; Li, D. W.; Wang, T.; Brüsweiler, R.; Iwahara, J., Signature of mobile hydrogen bonding of lysine side chains from long-range ^{15}N - ^{13}C scalar J-couplings and computation. *J Am Chem Soc* **2011**, *133* (24), 9192-5.
59. Pérez, C.; Löhr, F.; Rüterjans, H.; Schmidt, J. M., Self-Consistent Karplus Parametrization of 3 J Couplings Depending on the Polypeptide Side-Chain Torsion χ_1 . *Journal of the American Chemical Society* **2001**, *123* (29), 7081-7093.
60. Chou, J. J.; Case, D. A.; Bax, A., Insights into the mobility of methyl-bearing side chains in proteins from $(3)\text{J}(\text{CC})$ and $(3)\text{J}(\text{CN})$ couplings. *J Am Chem Soc* **2003**, *125* (29), 8959-66.
61. Rohs, R.; Jin, X.; West, S. M.; Joshi, R.; Honig, B.; Mann, R. S., Origins of specificity in protein-DNA recognition. *Annual review of biochemistry* **2010**, *79*, 233-69.

62. Caro, J. A.; Harpole, K. W.; Kasinath, V.; Lim, J.; Granja, J.; Valentine, K. G.; Sharp, K. A.; Wand, A. J., Entropy in molecular recognition by proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114* (25), 6563-6568.
63. Berglund, H.; Baumann, H.; Knapp, S.; Ladenstein, R.; Härd, T., Flexibility of an arginine side chain at a DNA-protein interface. *J Am Chem Soc* **1995**, *117* (51), 12883-12884.
64. Iwahara, J.; Jung, Y. S.; Clore, G. M., Heteronuclear NMR spectroscopy for lysine NH₃ groups in proteins: unique effect of water exchange on (15)N transverse relaxation. *J Am Chem Soc* **2007**, *129* (10), 2971-80.
65. Esadze, A.; Li, D. W.; Wang, T.; Brüscheiler, R.; Iwahara, J., Dynamics of lysine side-chain amino groups in a protein studied by heteronuclear ¹H-¹⁵N NMR spectroscopy. *J Am Chem Soc* **2011**, *133*, 909-919.
66. Zandarashvili, L.; Li, D. W.; Wang, T.; Bruscheiler, R.; Iwahara, J., Signature of mobile hydrogen bonding of lysine side chains from long-range ¹⁵N-¹³C scalar J-couplings and computation. *J Am Chem Soc* **2011**, *133* (24), 9192-5.
67. Werbeck, N. D.; Kirkpatrick, J.; Hansen, D. F., Probing arginine side-chains and their dynamics with carbon-detected NMR spectroscopy: application to the 42 kDa human histone deacetylase 8 at high pH. *Angewandte Chemie* **2013**, *52* (11), 3145-7.
68. Anderson, K. M.; Esadze, A.; Manoharan, M.; Bruscheiler, R.; Gorenstein, D. G.; Iwahara, J., Direct Observation of the Ion-Pair Dynamics at a Protein-DNA Interface by NMR Spectroscopy. *J Am Chem Soc* **2013**, *135* (9), 3613-9.
69. Nguyen, D.; Lokesh, G. L. R.; Volk, D. E.; Iwahara, J., A Unique and Simple Approach to Improve Sensitivity in 15N-NMR Relaxation Measurements for NH(3)(+) Groups: Application to a Protein-DNA Complex. *Molecules* **2017**, *22* (8), 1355.
70. Zandarashvili, L.; Nguyen, D.; Anderson, K. M.; White, M. A.; Gorenstein, D. G.; Iwahara, J., Entropic Enhancement of Protein-DNA Affinity by Oxygen-to-Sulfur Substitution in DNA Phosphate. *Biophys J* **2015**, *109* (5), 1026-37.
71. Fraenkel, E.; Rould, M. A.; Chambers, K. A.; Pabo, C. O., Engrailed homeodomain-DNA complex at 2.2 Å resolution: a detailed view of the interface and comparison with other engrailed structures. *Journal of molecular biology* **1998**, *284* (2), 351-61.
72. Billeter, M.; Qian, Y. Q.; Otting, G.; Muller, M.; Gehring, W.; Wüthrich, K., Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. *Journal of molecular biology* **1993**, *234* (4), 1084-93.

73. Fernandez, C.; Szyperski, T.; Billeter, M.; Ono, A.; Iwai, H.; Kainosho, M.; Wüthrich, K., Conformational changes of the BS2 operator DNA upon complex formation with the Antennapedia homeodomain studied by NMR with $^{13}\text{C}/^{15}\text{N}$ -labeled DNA. *Journal of molecular biology* **1999**, *292* (3), 609-17.
74. Zandarashvili, L.; Esadze, A.; Kemme, C. A.; Chattopadhyay, A.; Nguyen, D.; Iwahara, J., Residence Times of Molecular Complexes in Solution from NMR Data of Intermolecular Hydrogen-Bond Scalar Coupling. *The journal of physical chemistry letters* **2016**, *7* (5), 820-4.
75. Chen, C. Y.; Esadze, A.; Zandarashvili, L.; Nguyen, D.; Pettitt, B. M.; Iwahara, J., Dynamic Equilibria of Short-Range Electrostatic Interactions at Molecular Interfaces of Protein-DNA Complexes. *J Phys Chem Lett* **2015**, *6* (14), 2733-2737.
76. Ernst, R. R.; Bodenhausen, G.; Wokaun, A., Heteronuclear polarization transfer. In *Principles of nuclear magnetic resonance in one and two dimensions*, Oxford University Press: New York, 1987.
77. van de Ven, F. J. M., Dephasing coherences. In *Multidimensional NMR in liquids: basic principles and experimental methods*, VCH Publishers: New York, 1995; pp 211-224.
78. Kay, L. E.; Bull, T. E.; Nicholson, L. K.; Griesinger, C.; Schwalbe, H.; Bax, A.; Torchia, D. A., The Measurement of Heteronuclear Transverse Relaxation-Times in Ax3 Spin Systems Via Polarization-Transfer Techniques. *Journal of magnetic resonance* **1992**, *100* (3), 538-558.
79. Palmer, A. G.; Wright, P. E.; Rance, M., Measurement of Relaxation-Time Constants for Methyl-Groups by Proton-Detected Heteronuclear Nmr-Spectroscopy. *Chem Phys Lett* **1991**, *185* (1-2), 41-46.
80. Kupče, E.; Boyd, J.; Campbell, I. D., Short selective pulses for biochemical applications. *J Magn Reson Ser B* **1995**, *106* (3), 300-303.
81. Hansen, D. F.; Vallurupalli, P.; Kay, L. E., An improved ^{15}N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J Phys Chem B* **2008**, *112* (19), 5898-904.
82. Korzhnev, D. M.; Tischenko, E. V.; Arseniev, A. S., Off-resonance effects in ^{15}N T2 CPMG measurements. *Journal of biomolecular NMR* **2000**, *17* (3), 231-7.
83. Iwahara, J.; Clore, G. M., Sensitivity improvement for correlations involving arginine side-chain $\text{N}\epsilon/\text{H}\epsilon$ resonances in multi-dimensional NMR experiments using broadband ^{15}N 180 degrees pulses. *Journal of biomolecular NMR* **2006**, *36* (4), 251-7.

84. Nguyen, D.; Iwahara, J., Impact of ^{15}N - ^{15}N scalar couplings on ^{15}N transverse relaxation measurements for arginine side chains of proteins. *Journal of biomolecular NMR* **2018**, *in press*.
85. Shaka, A. J.; Keeler, J.; Freeman, R., Evaluation of a New Broad-Band Decoupling Sequence - Waltz-16. *J Magn Reson* **1983**, *53* (2), 313-340.
86. Hansen, D. F.; Kay, L. E., Improved magnetization alignment schemes for spin-lock relaxation experiments. *J Biomol NMR* **2007**, *37* (4), 245-55.
87. Woessner, D. E., Nuclear spin relaxation in ellipsoids undergoing rotational Brownian motion. *J Chem Phys* **1962**, *37*, 647-654.
88. Iwahara, J.; Peterson, R. D.; Clubb, R. T., Compensating increases in protein backbone flexibility occur when the Dead ringer AT-rich interaction domain (ARID) binds DNA: a nitrogen-15 relaxation study. *Protein Sci* **2005**, *14* (5), 1140-50.
89. Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A., Rotational diffusion anisotropy of human ubiquitin from N-15 NMR relaxation. *J Am Chem Soc* **1995**, *117* (50), 12562-12566.
90. Lipari, G.; Szabo, A., Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .1. Theory and Range of Validity. *J Am Chem Soc* **1982**, *104* (17), 4546-4559.
91. Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M., Deviations from the simple 2-parameter model-free approach to the interpretation of N-15 nuclear magnetic-relaxation of proteins. *J Am Chem Soc* **1990**, *112* (12), 4989-4991.
92. d'Auvergne, E. J.; Gooley, P. R., The use of model selection in the model-free analysis of protein dynamics. *Journal of biomolecular NMR* **2003**, *25* (1), 25-39.
93. Abe, N.; Dror, I.; Yang, L.; Slattery, M.; Zhou, T.; Bussemaker, H. J.; Rohs, R.; Mann, R. S., Deconvolving the recognition of DNA shape from sequence. *Cell* **2015**, *161* (2), 307-18.
94. Dror, I.; Zhou, T.; Mandel-Gutfreund, Y.; Rohs, R., Covariation between homeodomain transcription factors and the shape of their DNA binding sites. *Nucleic acids research* **2014**, *42* (1), 430-41.
95. Yang, D. W.; Kay, L. E., Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: Application to protein folding. *Journal of molecular biology* **1996**, *263* (2), 369-382.
96. Masunov, A.; Lazaridis, T., Potentials of mean force between ionizable amino acid side chains in water. *J Am Chem Soc* **2003**, *125* (7), 1722-30.

97. Rohs, R.; West, S. M.; Sosinsky, A.; Liu, P.; Mann, R. S.; Honig, B., The role of DNA shape in protein-DNA recognition. *Nature* **2009**, *461* (7268), 1248-53.
98. Dias, N.; Stein, C. A., Antisense oligonucleotides: basic concepts and mechanisms. *Mol Cancer Ther* **2002**, *1* (5), 347-55.
99. Kole, R.; Krainer, A. R.; Altman, S., RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov* **2012**, *11* (2), 125-40.
100. Morishita, R.; Higaki, J.; Tomita, N.; Ogihara, T., Application of transcription factor "decoy" strategy as means of gene therapy and study of gene expression in cardiovascular disease. *Circ Res* **1998**, *82* (10), 1023-1028.
101. Eckstein, F., Phosphorothioates, essential components of therapeutic oligonucleotides. *Nucleic acid therapeutics* **2014**, *24* (6), 374-87.
102. King, D. J.; Bassett, S. E.; Li, X.; Fennewald, S. A.; Herzog, N. K.; Luxon, B. A.; Shope, R.; Gorenstein, D. G., Combinatorial selection and binding of phosphorothioate aptamers targeting human NF-kappa B RelA(p65) and p50. *Biochemistry* **2002**, *41* (30), 9696-706.
103. Marshall, W. S.; Beaton, G.; Stein, C. A.; Matsukura, M.; Caruthers, M. H., Inhibition of human immunodeficiency virus activity by phosphorodithioate oligodeoxycytidine. *Proc Natl Acad Sci USA* **1992**, *89* (14), 6265-9.
104. Marshall, W. S.; Caruthers, M. H., Phosphorodithioate DNA as a potential therapeutic drug. *Science* **1993**, *259* (5101), 1564-70.
105. Yang, X.; Bassett, S. E.; Li, X.; Luxon, B. A.; Herzog, N. K.; Shope, R. E.; Aronson, J.; Prow, T. W.; Leary, J. F.; Kirby, R.; Ellington, A. D.; Gorenstein, D. G., Construction and selection of bead-bound combinatorial oligonucleoside phosphorothioate and phosphorodithioate aptamer libraries designed for rapid PCR-based sequencing. *Nucleic acids research* **2002**, *30* (23), e132.
106. Pallan, P. S.; Yang, X. B.; Sierant, M.; Abeydeera, N. D.; Hassell, T.; Martinez, C.; Janicka, M.; Nawrot, B.; Egli, M., Crystal structure, stability and Ago2 affinity of phosphorodithioate-modified RNAs. *Rsc Advances* **2014**, *4* (110), 64901-64904.
107. Yang, X. B.; Mierzejewski, E., Synthesis of nucleoside and oligonucleoside dithiophosphates. *New J Chem* **2010**, *34* (5), 805-819.
108. Wagner, R. W.; Flanagan, W. M., Antisense technology and prospects for therapy of viral infections and cancer. *Mol Med Today* **1997**, *3* (1), 31-8.
109. Wang, S.; Lee, R. J.; Cauchon, G.; Gorenstein, D. G.; Low, P. S., Delivery of antisense oligodeoxyribonucleotides against the human epidermal growth factor receptor into cultured KB cells with liposomes conjugated to folate via polyethylene glycol. *Proc Natl Acad Sci USA* **1995**, *92* (8), 3318-22.

110. Wang, L.; Chen, S.; Vergin, K. L.; Giovannoni, S. J.; Chan, S. W.; DeMott, M. S.; Taghizadeh, K.; Cordero, O. X.; Cutler, M.; Timberlake, S.; Alm, E. J.; Polz, M. F.; Pinhassi, J.; Deng, Z.; Dedon, P. C., DNA phosphorothioation is widespread and quantized in bacterial genomes. *Proc Natl Acad Sci U S A* **2011**, *108* (7), 2963-8.
111. Wang, L.; Chen, S.; Xu, T.; Taghizadeh, K.; Wishnok, J. S.; Zhou, X.; You, D.; Deng, Z.; Dedon, P. C., Phosphorothioation of DNA in bacteria by *dnd* genes. *Nat Chem Biol* **2007**, *3* (11), 709-10.
112. Xie, X.; Liang, J.; Pu, T.; Xu, F.; Yao, F.; Yang, Y.; Zhao, Y. L.; You, D.; Zhou, X.; Deng, Z.; Wang, Z., Phosphorothioate DNA as an antioxidant in bacteria. *Nucleic acids research* **2012**, *40* (18), 9115-24.
113. Gambari, R., Recent patents on therapeutic applications of the transcription factor decoy approach. *Expert Opin Ther Pat* **2011**, *21* (11), 1755-71.
114. Morishita, R.; Sugimoto, T.; Aoki, M.; Kida, I.; Tomita, N.; Moriguchi, A.; Maeda, K.; Sawa, Y.; Kaneda, Y.; Higaki, J.; Ogihara, T., In vivo transfection of cis element "decoy" against nuclear factor-kappaB binding site prevents myocardial infarction. *Nat Med* **1997**, *3* (8), 894-9.
115. Ohtani, K.; Egashira, K.; Usui, M.; Ishibashi, M.; Hiasa, K. I.; Zhao, Q.; Aoki, M.; Kaneda, Y.; Morishita, R.; Takeshita, A., Inhibition of neointimal hyperplasia after balloon injury by cis-element 'decoy' of early growth response gene-1 in hypercholesterolemic rabbits. *Gene Ther* **2004**, *11* (2), 126-32.
116. Lide, D. R., *CRC handbook of chemistry and physics*. 84th ed.; CRC Press: Boca Raton, FL, 2003.
117. Affolter, M.; Percival-Smith, A.; Muller, M.; Leupin, W.; Gehring, W. J., DNA binding properties of the purified Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **1990**, *87* (11), 4093-7.
118. Otting, G.; Qian, Y. Q.; Billeter, M.; Muller, M.; Affolter, M.; Gehring, W. J.; Wuthrich, K., Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J* **1990**, *9* (10), 3085-92.
119. Qian, Y. Q.; Billeter, M.; Otting, G.; Muller, M.; Gehring, W. J.; Wuthrich, K., The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* **1989**, *59* (3), 573-80.
120. Dill, K. A.; Bromberg, S., *Molecular Driving Forces: statistical thermodynamics in biology, chemistry, physics, and nanoscience*. Garland Sciences: New York, 2011.
121. Clore, G. M.; Gronenborn, A. M., Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol* **1998**, *16* (1), 22-34.

122. Andre, I.; Linse, S.; Mulder, F. A., Residue-specific pKa determination of lysine and arginine side chains by indirect ¹⁵N and ¹³C NMR spectroscopy: application to apo calmodulin. *J Am Chem Soc* **2007**, *129* (51), 15805-13.
123. 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* **1995**, *6* (3), 277-293.
124. Johnson, B. A.; Blevins, R. A., Nmr View - a Computer-Program for the Visualization and Analysis of Nmr Data. *Journal of biomolecular NMR* **1994**, *4* (5), 603-614.
125. Headd, J. J.; Echols, N.; Afonine, P. V.; Grosse-Kunstleve, R. W.; Chen, V. B.; Moriarty, N. W.; Richardson, D. C.; Richardson, J. S.; Adams, P. D., Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. *Acta Crystallogr D Biol Crystallogr* **2012**, *68* (Pt 4), 381-90.
126. Painter, J.; Merritt, E. A., Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **2006**, *62* (Pt 4), 439-50.
127. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 4), 486-501.
128. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *Journal of applied crystallography* **2007**, *40* (Pt 4), 658-674.
129. Florián, J.; Strajbl, M.; Warshel, A., Conformational flexibility of phosphate, phosphonate, and phosphorothioate methyl esters in aqueous solution. *J Am Chem Soc* **1998**, *120* (31), 7959-7966.
130. Frederiksen, J. K.; Piccirilli, J. A., Separation of RNA phosphorothioate oligonucleotides by HPLC. *Methods in enzymology* **2009**, *468*, 289-309.
131. Freiburger, L.; Auclair, K.; Mittermaier, A., Global ITC fitting methods in studies of protein allostery. *Methods* **2015**, *76*, 149-61.
132. Ladbury, J. E.; Wright, J. G.; Sturtevant, J. M.; Sigler, P. B., A thermodynamic study of the trp repressor-operator interaction. *J Mol Biol* **1994**, *238* (5), 669-81.
133. Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Rance, M.; Skelton, N. J., *Protein NMR Spectroscopy: Principles and Practice*. 2 ed.; Elsevier Academic Press: Burlington, 2007.
134. Gorenstein, D. G., Conformation and dynamics of DNA and protein-DNA complexes by ³¹P NMR. *Chem Rev* **1994**, *94* (5), 1315-1338.

135. Gorenstein, D. G., 31P NMR of DNA. *Methods Enzymol* **1992**, *211*, 254-86.
136. Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C., Precision neutron diffraction structure determination of protein and nucleic acid components. I. The crystal and molecular structure of the amino Acid L-alanine. *J Am Chem Soc* **1971**, *94* (8), 2657-2660.
137. Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K., Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res* **2009**, *37* (17), 5917-29.
138. Howard, D. L.; Kjaergaard, H. G., Hydrogen bonding to divalent sulfur. *Phys Chem Chem Phys* **2008**, *10* (28), 4113-4118.
139. Wennmohs, F.; Staemmler, V.; Schindler, M., Theoretical investigation of weak hydrogen bonds to sulfur. *J Chem Phys* **2003**, *119* (6), 3208-3218.
140. Krishnan, M.; Smith, J. C., Response of small-scale, methyl rotors to protein-ligand association: a simulation analysis of calmodulin-peptide binding. *J Am Chem Soc* **2009**, *131* (29), 10083-91.
141. Record, M. T., Jr.; Lohman, M. L.; De Haseth, P., Ion effects on ligand-nucleic acid interactions. *J Mol Biol* **1976**, *107* (2), 145-58.
142. Morishita, R.; Gibbons, G. H.; Horiuchi, M.; Ellison, K. E.; Nakajima, M.; Zhang, L.; Kaneda, Y.; Ogihara, T.; Dzau, V. J., A Gene-Therapy Strategy Using a Transcription Factor Decoy of the E2f Binding-Site Inhibits Smooth-Muscle Proliferation in-Vivo. *Proc Natl Acad Sci USA* **1995**, *92* (13), 5855-5859.
143. Mann, M. J.; Dzau, V. J., Therapeutic applications of transcription factor decoy oligonucleotides. *The Journal of clinical investigation* **2000**, *106* (9), 1071-5.
144. Morishita, R.; Higaki, J.; Tomita, N.; Ogihara, T., Application of transcription factor "decoy" strategy as means of gene therapy and study of gene expression in cardiovascular disease. *Circ Res* **1998**, *82* (10), 1023-8.
145. Cho, Y.; Zhu, F. C.; Luxon, B. A.; Gorenstein, D. G., 2D 1H and 31P NMR spectra and distorted A-DNA-like duplex structure of a phosphorodithioate oligonucleotide. *Journal of biomolecular structure & dynamics* **1993**, *11* (3), 685-702.
146. Volk, D. E.; Yang, X.; Fennewald, S. M.; King, D. J.; Bassett, S. E.; Venkitachalam, S.; Herzog, N.; Luxon, B. A.; Gorenstein, D. G., Solution structure and design of dithiophosphate backbone aptamers targeting transcription factor NF-kappaB. *Bioorg Chem* **2002**, *30* (6), 396-419.
147. Badis, G.; Berger, M. F.; Philippakis, A. A.; Talukder, S.; Gehrke, A. R.; Jaeger, S. A.; Chan, E. T.; Metzler, G.; Vedenko, A.; Chen, X.; Kuznetsov, H.; Wang, C. F.; Coburn, D.; Newburger, D. E.; Morris, Q.; Hughes, T. R.; Bulyk, M. L.,

- Diversity and complexity in DNA recognition by transcription factors. *Science* **2009**, *324* (5935), 1720-3.
148. Wunderlich, Z.; Mirny, L. A., Different gene regulation strategies revealed by analysis of binding motifs. *Trends Genet* **2009**, *25* (10), 434-40.
 149. Esadze, A.; Kemme, C. A.; Kolomeisky, A. B.; Iwahara, J., Positive and negative impacts of nonspecific sites during target location by a sequence-specific DNA-binding protein: origin of the optimal search at physiological ionic strength. *Nucleic Acids Res* **2014**, *42* (11), 7039-46.
 150. Winter, R. B.; Berg, O. G.; von Hippel, P. H., Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor--operator interaction: kinetic measurements and conclusions. *Biochemistry* **1981**, *20* (24), 6961-77.
 151. Berg, O. G.; Winter, R. B.; von Hippel, P. H., Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **1981**, *20* (24), 6929-48.
 152. Krepel, D.; Gomez, D.; Klumpp, S.; Levy, Y., Mechanism of Facilitated Diffusion during a DNA Search in Crowded Environments. *J Phys Chem B* **2016**, *120* (43), 11113-11122.
 153. Nguyen, D.; Zandarashvili, L.; White, M. A.; Iwahara, J., Stereospecific Effects of Oxygen-to-Sulfur Substitution in DNA Phosphate on Ion Pair Dynamics and Protein-DNA Affinity. *Chembiochem : a European journal of chemical biology* **2016**, *17* (17), 1636-42.
 154. Esadze, A.; Iwahara, J., Stopped-flow fluorescence kinetic study of protein sliding and intersegment transfer in the target DNA search process. *J Mol Biol* **2014**, *426* (1), 230-44.
 155. Johnson Jr, C. S., Diffusion ordered nuclear magnetic resonance spectroscopy: principles and applications. *Progress in Nuclear Magnetic Resonance Spectroscopy* **1999**, *34* (3-4), 203-256.
 156. Wu, D.; Chen, A.; Johnson, C. S., An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. *Journal of magnetic resonance, Series A* **1995**, *115* (2), 260-264.
 157. Wenner, J. R.; Bloomfield, V. A., Crowding effects on EcoRV kinetics and binding. *Biophys J* **1999**, *77* (6), 3234-41.
 158. Yao, S.; Babon, J. J.; Norton, R. S., Protein effective rotational correlation times from translational self-diffusion coefficients measured by PFG-NMR. *Biophys Chem* **2008**, *136* (2-3), 145-51.

159. Esadze, A.; Kemme, C. A.; Kolomeisky, A. B.; Iwahara, J., Positive and negative impacts of nonspecific sites during target location by a sequence-specific DNA-binding protein: origin of the optimal search at physiological ionic strength. *Nucleic Acids Res* **2014**.
160. Bancaud, A.; Huet, S.; Daigle, N.; Mozziconacci, J.; Beaudouin, J.; Ellenberg, J., Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J* **2009**, *28* (24), 3785-98.
161. Brackley, C. A.; Cates, M. E.; Marenduzzo, D., Intracellular facilitated diffusion: searchers, crowders, and blockers. *Phys Rev Lett* **2013**, *111* (10), 108101.
162. Tabaka, M.; Kalwarczyk, T.; Hołyst, R., Quantitative influence of macromolecular crowding on gene regulation kinetics. *Nucleic Acids Res* **2014**, *42* (2), 727-38.
163. Liu, L.; Cherstvy, A. G.; Metzler, R., Facilitated Diffusion of Transcription Factor Proteins with Anomalous Bulk Diffusion. *J Phys Chem B* **2017**, *121* (6), 1284-1289.
164. van den Berg, B.; Wain, R.; Dobson, C. M.; Ellis, R. J., Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell. *EMBO J* **2000**, *19* (15), 3870-5.
165. Tokuriki, N.; Kinjo, M.; Negi, S.; Hoshino, M.; Goto, Y.; Urabe, I.; Yomo, T., Protein folding by the effects of macromolecular crowding. *Protein Sci* **2004**, *13* (1), 125-33.
166. Munishkina, L. A.; Cooper, E. M.; Uversky, V. N.; Fink, A. L., The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J Mol Recognit* **2004**, *17* (5), 456-64.
167. Huang, J.; MacKerell, A. D., CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* **2013**, *34* (25), 2135-45.
168. Rojanasakul, Y., Antisense oligonucleotide therapeutics: drug delivery and targeting. *Advanced Drug Delivery Reviews* **1996**, *18* (2), 115-131.
169. Crooke, S. T.; Bennett, C. F., Progress in antisense oligonucleotide therapeutics. *Annual Review of Pharmacology and Toxicology* **1996**, *36* (1), 107-129.
170. Crooke, S. T., Progress toward oligonucleotide therapeutics: pharmacodynamic properties. *The FASEB journal* **1993**, *7* (6), 533-539.

VITA

Dan Nguyen was born on October 10th, 1987 to Tai Nguyen and Hong-Thu Phan in Houston, TX. He received his Bachelor of Science in Biochemistry from the University of Texas at Austin in 2010. After receiving his undergraduate degree, Dan spent three and a half years at Baylor College of Medicine, where he worked under Dr. Rajagopal Sekhar to improve metabolic disease outcomes in patients using glutathione as an antioxidant. In 2014, Dan entered the Biochemistry and Molecular Biology graduate program through the Molecular Biophysics Education Track at the University of Texas Medical Branch. Under the guidance of Dr. Junji Iwahara, Dan worked on understanding the role of ion pair dynamics in protein-DNA interactions through biophysical analyses. During his dissertation work, Dan received several accolades including the Houston Area Molecular Biophysics fellowship, a highly competitive federal T32 grant. In addition to his research, Dan served as the Chair for the Biological Chemistry Student Organization and co-founded Innovation Labs @ UTMB.

PUBLICATIONS

Anderson, K.M., **Nguyen, D.**, Esadze, A., Zandarashvili, L., Gorenstein, D.G., Iwahara, J. (2015) A chemical approach for site-specific identification of NMR signals from protein side-chain NH_3^+ groups forming intermolecular ion pairs in protein-nucleic acid complexes. *J Biomol NMR* 62, 1-5.

Chen, C., Esadze, A., Zandarashvili, L., **Nguyen, D.**, Pettitt, B.M., Iwahara, J. (2015) Dynamic equilibria of short-range electrostatic interactions at molecular interfaces of protein-DNA complexes. *J Phys Chem Lett* 6, 2733-7.

Zandarashvili, L. and Nguyen, D., Anderson, K.M., White, M.A., Gorenstein, D.G., Iwahara, J. (2015) Entropic enhancement of protein-DNA affinity by oxygen-to-sulfur substitution in DNA phosphate. *Biophys J* 109, 1026-37.

Zandarashvili, L., Esadze, A., Kemme, C.A., Chattopadhyay, A., **Nguyen, D.**, Iwahara, J. (2016) Residence times of molecular complexes in solution from NMR data of intermolecular hydrogen-bond scalar coupling. *J Phys Chem Lett* 7, 820-4.

Kemme, C.A., **Nguyen, D.**, Chattopadhyay, A., Iwahara, J. (2016) Regulation of transcription factors via natural decoys in genomic DNA. *Transcription* 7, 115-20.

Nguyen, D., Zandarashvili, L., White, M.A., Iwahara, J. (2016) Stereospecific effects of oxygen-to-sulfur substitution in DNA phosphate on ion-pair dynamics and protein-DNA affinity. *ChemBioChem* 17, 1636-42.

Nguyen, D., Lokesh, G.L.R., Volk, D.E., Iwahara, J. (2017) A unique and simple approach to improve sensitivity in ^{15}N NMR relaxation measurements for NH_3^+ groups: Application to a protein-DNA complex. *Molecules* 22, E1355.

Nguyen, D., Hoffpauir, Z., Iwahara, J. (2017) Internal motions of basic side chains of the Antennapedia homeodomain in the free and DNA-bound states. *Biochemistry* 56, 5866-9.

Nguyen, D., Iwahara, J. (2018) Impact of two-bond ^{15}N - ^{15}N scalar couplings on ^{15}N transverse relaxation measurements for arginine side chains of proteins. *J Biomol NMR*, in press.

Permanent address: 15803 Hermitage Oaks Dr, Tomball, TX 77377

This dissertation was typed by Dan Nguyen.