

LOCAL NUCLEIC ACID BASE CONFORMATION STUDY BY GUANINE  
FLUORESCENT ANALOGUE 6-METHYL ISOXANTHOPTERIN (6-MI) LABELED  
DNA

by

HUIYING JI

A DISSERTATION

Presented to the Department of Chemistry and Biochemistry  
and the Graduate School of the University of Oregon  
in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy

March 2019

DISSERTATION APPROVAL PAGE

Student: Huiying Ji

Title: Local Nucleic Acid Base Conformation Study by Guanine Fluorescent Analogue 6-Methyl Isoxanthopterin (6-MI) Labeled DNA

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Chemistry and Biochemistry by:

Michael Kellman	Chairperson
Andrew Marcus	Advisor
Peter von Hippel	Advisor
Marina Guenza	Core Member
Alice Barkan	Institutional Representative

and

Janet Woodruff-Borden      Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded March 2019

© 2019 Huiying Ji

## DISSERTATION ABSTRACT

Huiying Ji

Doctor of Philosophy

Department of Chemistry and Biochemistry

March 2019

Title: Local Nucleic Acid Base Conformation Study by Guanine Fluorescent Analogue 6-Methyl Isoxanthopterin (6-MI) Labeled DNA

Understanding the local conformations of DNA at the level of individual nucleic acid bases is important for the study of the mechanism of DNA sequence-dependent behavior. Here we apply linear absorption, circular dichroism (CD), and fluorescence spectroscopy to study the DNA local base conformation using 6-methyl Isoxanthopterin (6-MI) labeled DNA. We interpret excitation–emission peak shift (EES) measurements of the 6-MI, both as a ribonucleotide monophosphate in solution and as a site-specific substituent for guanine in various DNA constructs, by implementing a simple two-state model. We show that the spectroscopic properties of the 6-MI probe in DNA can be used to obtain detailed information about local base conformations and conformational heterogeneity and fluctuations. Based on these findings, we apply a simple theoretical model to calculate CD of 6-MI substituted DNA constructs. We find that our model can be used to extract base-sequence-dependent information about the local conformation of the 6-MI probe as modulated by the local base or base-pair environment. We next apply 6-MI to probe the ligand insertion of small molecules to duplex DNA, further extending the potential of 6-MI as a useful reporter of local nucleic acid base conformation. These studies served to establish a new level of sophistication in qualitatively analyzing 6-MI structural behavior in terms of local base stacking and unstacking conformations.



This dissertation contains previously published and unpublished co-authored material.

## CURRICULUM VITAE

NAME OF AUTHOR: Huiying Ji

### GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene  
Dalian University of Technology, Dalian, China

### DEGREES AWARDED:

Doctor of Philosophy, Chemistry, 2018, University of Oregon  
Bachelor of Science, Chemistry, 2012, Dalian University of Technology

### AREAS OF SPECIAL INTEREST:

Fluorescence spectroscopy  
Theoretical circular dichroism calculation  
Local base conformation of DNA

### PROFESSIONAL EXPERIENCE:

Graduate Research Assistant, Department of Chemistry and Biochemistry,  
University of Oregon, Eugene, 2012-2018

Graduate Teaching Assistant, Department of Chemistry and Biochemistry,  
University of Oregon, Eugene, 2012-2018

### PUBLICATION:

Neil P. Johnson, Huiying Ji, Thomas H. Steinberg, Peter H. von Hippel, and Andrew H. Marcus, "Sequence-Dependent Conformational Heterogeneity and Proton-Transfer Reactivity of the Fluorescent Guanine Analogue 6-Methyl Isoxanthopterin (6-MI) in DNA," *J. Phys. Chem. B*, **2015**, *119* (40), pp 12798–12807

## ACKNOWLEDGMENTS

I am extremely grateful for having the opportunity to complete my PhD program at the University of Oregon. During grad school, I have grown tremendously both as a scientist, and as a person. I want to thank everyone who has reached out and helped me in your own way.

First, I would like to thank my advisor, Professor Andy Marcus for giving me the opportunity to pursue a career in science in the Marcus group. Your guidance has been essential to my work. You have influenced me significantly as a researcher. I would like to thank my co-advisor, Professor Pete von Hippel. Your wisdom and positivity has been inspiring to us all. I would like to thank my committee members, Professor Mike Kellman, Marina Guenza and Alice Barkan. Thank you for asking important questions and giving me valuable advice.

I would like to thank my incredible mentor and collaborator Dr. Neil Johnson, who have influenced and helped me in many ways. Thank you for always making me feel excited about our research. I want to thank my collaborator Dr. Pablo Romano for encouraging me to learn programming and providing useful resources.

I want to thank all lab members in Marcus group and von Hippel group, who have created a welcoming work environment and showed enormous support, especially Dr. Loni Kringle. I might be the positive one that keeps our office balanced, but it is your perseverance that inspires me to keep trying.

I am grateful to all my friends, especially the ones with whom I went through grad school together. I feel so lucky to have met you. I will always cherish our shared memory in grad school. I will never make this far without you. Thank you for making me feel welcomed to the grad school community.

I want to thank my loving parents for supporting me unconditionally, although you do not understand what I am doing. I want to thank Walter Sorenson for helping me going through the last stage of grad school.

Finally, this work would not be possible without the financial support from the National Science Foundation and the National Institute of Health.

To Fellow Grad Students

## TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
Outline and Acknowledgement of the Contributions by Others to this Dissertation .....	4
II. SEQUENCE-DEPENDENT CONFORMATIONAL HETEROGENEITY AND PROTON- TRANSFER REACTIVITY OF THE FLUORESCENT GUANINE ANALOGUE 6-METHYL ISOXANTHOPTERIN (6-MI) IN DNA.....	7
Overview.....	7
Introduction.....	7
Materials and Methods.....	11
Results.....	13
Excitation-Emission Shift (EES) .....	13
Theoretical Considerations of the EES .....	14
EES of 6-MI Ribonucleotide Monophosphate (NMP) .....	19
EES of 6-MI Substituted DNA Constructs .....	22
Discussion and Conclusions .....	28
Summary and Bridge to Chapter III.....	34
III. THEORETICAL ASPECTS OF CIRCULAR DICHROISM CALCULATIONS .....	36
Overview.....	36
Introduction.....	36
Approximations and Principles of Theoretical Model.....	38
Electronic Coupling and Rotational Strength .....	39

Chapter	Page
Determination of Electric Dipole Transition Moments (EDTMs).....	44
EDTMs of natural bases.....	45
EDTMs of 6-MI .....	47
EDTMs of ACMA .....	50
Circular Dichroism Calculation of the Dinucleotides.....	52
MD simulation of the Dinucleotides .....	52
CD Calculation of the Dinucleotides .....	53
Bridge to Chapter IV.....	56
<b>IV. LOCAL NUCLEIC ACID BASE CONFORMATION REVEALED BY CIRCULAR DICHROISM STUDIES OF 6-METHYL ISOXANTHOPTERIN SUBSTITUTED DNA CONSTRUCTS .....</b>	<b>57</b>
Overview.....	57
Introduction.....	57
Materials and Methods.....	59
Experimental procedures .....	59
Preparation of DNA constructs labeled with 6-MI.....	59
Spectroscopy.....	60
Results.....	61
Discussion.....	70
Conclusions and Future Direction .....	73
Summary and Bridge to Chapter V.....	74
<b>V. STUDIES OF FREE AND TETHERED INTERCALATOR MOLECULES INTERACTING WITH 6-MI LABELED DNA .....</b>	<b>75</b>
Overview.....	75

Chapter	Page
Introduction.....	76
Materials and Methods.....	79
Chemicals.....	79
Spectroscopy.....	81
Results.....	81
The 6-MI/AO FRET Pair.....	81
Tethered ACMA and 6-MI Labeled DNA.....	86
Discussion.....	88
The AO-dsDNA Intercalation Complex.....	88
Tethered ACMA and 6-MI labeled DNA.....	95
Conclusions and Future Direction.....	98
VI. CONCLUDING REMARKS.....	100
Summary.....	100
Future Direction.....	101
APPENDICES.....	104
A. SUPPLEMENTARY INFORMATION FOR CHAPTER II: SEQUENCE-DEPENDENT CONFORMATIONAL HETEROGENEITY AND PROTON TRANSFER REACTIVITY OF THE FLUORESCENT GUANINE ANALOG 6-METHYL ISOXANTHOPTERIN (6-MI) IN DNA.....	104
B. SUPPLEMENTARY INFORMATION FOR CHAPTER III: THEORETICAL ASPECTS OF CIRCULAR DICHROISM CALCULATIONS.....	117
C. SUPPLEMENTARY INFORMATION FOR CHAPTER IV: LOCAL NUCLEIC ACID BASE CONFORMATION REVEALED BY CIRCULAR DICHROISM STUDIES OF 6-METHYL ISOXANTHOPTERIN SUBSTITUTED DNA CONSTRUCTS.....	122
D. SUPPLEMENTARY INFORMATION FOR CHAPTER V:	



Chapter	Page
STUDIES OF FREE AND TETHERED INTERCALATOR MOLECULES INTERACTING WITH 6-MI LABELED DNA .....	126
REFERENCES CITED.....	148

## LIST OF FIGURES

Figure	Page
1.1. The structure of guanine/cytosine base pair and 6-MI/cytosine base pair.....	2
2.1. Excitation–emission shift (EES). Absorbance and emission spectra of ssGXG and dsGXG constructs. ....	14
2.2. Homogeneous one-state model and heterogeneous two-state model for excitation–emission shift (EES) experiments. ....	15
2.3. EES data and model fits of 6-MI NMP in aqueous solutions for different pH conditions.....	20
2.4. EES data for 6-MI substituted DNA constructs.....	22
2.5. Standard free energy of reaction, free energy of activation, and equilibrium constant for proton transfer in 6-MI substituted DNA constructs.....	25
2.6. Reaction coordinates for proton exchange dsGXG and dsTXT .....	33
3.1. Energy level diagram of degenerate coupling between two energetically identical electronic transitions.....	41
3.2. Definitions for EDTM angles of natural orientation DNA bases .....	45
3.3. Definitions for EDTM angles of 6-MI Guanine fluorescent analogue .....	48
3.4. Excitation spectra and Gaussian decomposition analysis of 6-MI substituted ssDNA constructs. ....	49
3.5. Definitions for EDTM angles of ACMA. ....	51
3.6. Absorption spectrum and Gaussian decomposition analysis of ACMA.....	51
3.7. Circular dichroism of AA, AC, AG, AT.....	54
3.8. Circular dichroism of TA, TC, TG, TT. ....	54
3.9. Circular dichroism of GA, GC, GG, GT.....	55
3.10. Circular dichroism of CA, CC, CG, CT.....	55
4.1. Emission spectra of the YGG, GYG and GGY ssDNA constructs. ....	61
4.2. Experimental and simulated CD of ssDNA constructs YGG, GYG and GGY. ....	63

Figure	Page
4.3. Molecular structural model of the ss oligonucleotide YGG. ....	67
4.4. Comparison between experimental and simulated CD spectra of the 6-MI substituted ss DNA constructs YGG, GYG, and GGY. ....	68
4.5. Comparison between experimental and simulated CD spectra of the 6-MI substituted ds DNA constructs CCC/YGG, CCC/GYG, and CCC/GGY.....	69
4.6. Temperature dependence of CD signal for the indicated single stranded construct YGG, GYG, and GGY. ....	71
4.7. CD intensity at 350 nm as a function of temperature from Figure 4.6. ....	72
4.8. Calculations of the CD spectra of the GYG 6-MI substituted ss and ds DNA constructs. ....	72
5.1. Structure of free intercalator Acridine Orange (AO) and DNA backbone-tethered intercalator 9-amino-6-chloromethoxyacridine (ACMA). ....	77
5.2. Spectral overlap of 6-MI NMP and AO monomer in aqueous (buffered) solution.....	82
5.3. Fluorescence titration of AO into a solution of 6-MI NMP.....	82
5.4. Fluorescence titration experiments of AO, dsTGT, dsTXT, dsGGG, and dsGXG in buffer.....	84
5.5. Absorbance spectra of ACMA and 6-MI probe-labeled DNA constructs CZC, CZC/GGG, CZC/YGG, CZC/GYG, CZC/GGY. ....	87
5.6 Fluorescence titration experiments of AO with dsTXT, dsGXG and 6-MI NMP.....	90
5.7. AO absorbance overlaps with the emission spectra of dsGXG, dsTXT and 6-MI NMP.....	90
5.8. Fluorescence excited at 500 nm of AO in buffer solution containing dsTGT, dsGGG, dsTXT and dsGXG constructs. ....	92
5.9. Fluorescence excited at 335 nm of AO in buffer containing dsTGT, dsGGG, dsTXT and dsGXG constructs.....	93
5.10. The absorbance spectra of AO in buffer solution at concentrations ranging from 0.65 $\mu$ M to 21.41 $\mu$ M. ....	94

Figure	Page
5.11. $R_0$ calculated based on spectral overlap of the 6-MI NMP emission spectrum.....	95
5.12. Emission spectra of CZC/GGG, CZC/YGG, CZC/GYG and CZC/GGY. Excitation spectra of CZC, CZC/GGG, CZC/YGG, CZC/GYG and CZC/GGY. ....	96
5.13. A model of the CZC/GYG construct with the 6-MI base ‘flipped out’.....	96
A.1. EES data of 6-MI nucleotide monophosphate (NMP) in aqueous solution, as a function of rescaled excitation energy and pH. ....	112
A.2. EES data on pyrrolocytosine (PC) and 2-aminopurine (2-AP) substituted DNA, as a function of rescaled excitation energy. ....	113
A.3. EES versus excitation energy for 6-MI substituted ss and dsDNA constructs. ...	114
A.4. EES data versus rescaled excitation energy for 6-MI substituted ssDNA, as a function of pH = 6.2 (crosses), 7.6 (circles). ....	115
A.5. CD spectra of 6-MI residues in various dsDNA constructs. dsTXXT, dsGXXG, ssTXXT, and ssGXXG. ....	116
C.1. Comparison between experimental CD spectra and model calculations of DNA constructs in Table 4.1, in both wavelength and wavenumber units .....	122
C.2. Experimental and simulated CD (in wavenumber) of ssDNA constructs, which assume that the nucleotides adopt the B-form conformation.....	123
C.3. Comparison between experimental and simulated CD (in wavelength) spectra of the 6-MI substituted ss and ds DNA constructs as a function of the 6-MI probe rotation angle. ....	124
C.4. Comparison between experimental and simulated CD (in wavenumber) spectra of the 6-MI substituted ss and ds DNA constructs as a function of the 6-MI probe rotation angle. ....	125
D.1. Gaussian decomposition analysis of Acridine Orange (AO) absorbance. ....	127
D.2. Examples of Förster distance ( $R_0$ ) measurement fitting results.....	130
D.3. The initial placement of ACMA on the CCC/GGG constructs. ....	131
D.4. Angles corresponds to three conformational freedoms.....	131

Figure	Page
D.5. Comparison between experimental and simulated CD spectra of the ACMA tethered ss CZC construct as a function of ACMA rotation angles.....	132
D.6. The ss CZC conformation that produces optimal CD result. ....	133
D.7. Comparison between experimental and simulated CD spectra of the ACMA tethered CZC/GGG construct as a function of ACMA rotation angles. ...	134
D.8. The CZC/GGG conformation that produces optimal CD result. ....	135
D.9. Comparison between experimental and simulated CD spectra of the ACMA tethered CZC/YGG construct as a function of ACMA rotation angles. ....	136
D.10. The CZC/YGG conformation that produces optimal CD result. ....	136
D.11. Comparison between experimental and simulated CD spectra of the ACMA tethered CZC/GYG construct as a function of ACMA rotation angles. .	137
D.12. The CZC/GYG conformation that produces optimal CD result. ....	137
D.13. Comparison between experimental and simulated CD spectra of the ACMA tethered CZC/GGY construct as a function of ACMA rotation angles. .	138
D.14. The CZC/GGY conformation that produces optimal CD result. ....	139
D.15. Buffer-AO titration experiments. ....	139
D.16. dsGG(X)G-AO titration experiments.....	140
D.17. dsTG(X)T-AO titration experiments. ....	140
D.18. X-ray structure of 452D, d (CG (5-BrU) ACG) <sub>2</sub> crystalized with 4 DACA intercalating molecules. ....	142
D.19. AO-DNA CD calculation results based on the structure of 452D. ....	143
D.20. Sequence-dependent ACMA CD. ....	145
D.21. Sequence dependent ACMA fluorescence intensity. ....	147

## LIST OF TABLES

Table	Page
2.1. Base sequences and nomenclature for the 6-MI-containing ssDNA constructs used in Chapter II.....	12
3.1. Electric dipole transition moment (EDTM) parameters used for Thymine (T)....	46
3.2. EDTM parameters used for Adenine (A).....	46
3.3. EDTM parameters used for Cytosine (C). .....	47
3.4. EDTM parameters used for Guanine (G).....	47
3.5. Peak transition energies and standard deviations from Gaussian decomposition of excitation-emission spectra for the 6-MI substituted DNA constructs.....	50
3.6. EDTM parameters used for 6-methyl Isoxanthopterin (6-MI). .....	50
3.7. EDTM parameters used for 9-amino-6-chloro methoxyacridine (ACMA). .....	52
4.1. Base sequences and nomenclature for the 6-MI-containing ssDNA constructs used in Chapter IV. ....	59
5.1. Base sequences and nomenclature for the 6-MI-containing dsDNA constructs used in Chapter V. ....	80
5.2. Base sequences and nomenclature for the 6-MI- and ACMA-containing ssDNA constructs used in Chapter V.....	80
5.3. Concentration of Acridine Orange, dsDNA and the ratio of their concentrations in fluorescence titration experiment shown in Figure 5.3. ....	83
A.1. Parameters for fitting EES plot of 6-MI NMP .....	110
A.2. Parameters for fitting EES plot of ssDNA .....	111
A.3. Parameters for fitting EES plot of dsDNA.....	111
B.1. Comparison of experimental electronic transition moment in Thymine and Derivatives. ....	117
B.2. Comparison of experimental electronic transition moment in Adenine and Derivatives. ....	118

Table	Page
B.3. Comparison of experimental electronic transition moment in Cytosine and Derivatives.....	119
B.4. Comparison of experimental electronic transition moment in Guanine and Derivatives.....	120
B.5. Comparison of experimental electronic transition moment in Acridine Orange.....	121
D.1. AO absorbance fitting parameters.....	128
D.2. Concentration of Acridine Orange, dsDNA and the ratio of their concentrations in absorbance and CD titration experiments shown in Figure D.15, Figure D.16 and Figure D.17 .....	142
D.3. Base sequences and nomenclature for the ACMA-containing ssDNA constructs used in Appendix D. ....	144

# CHAPTER I

## INTRODUCTION

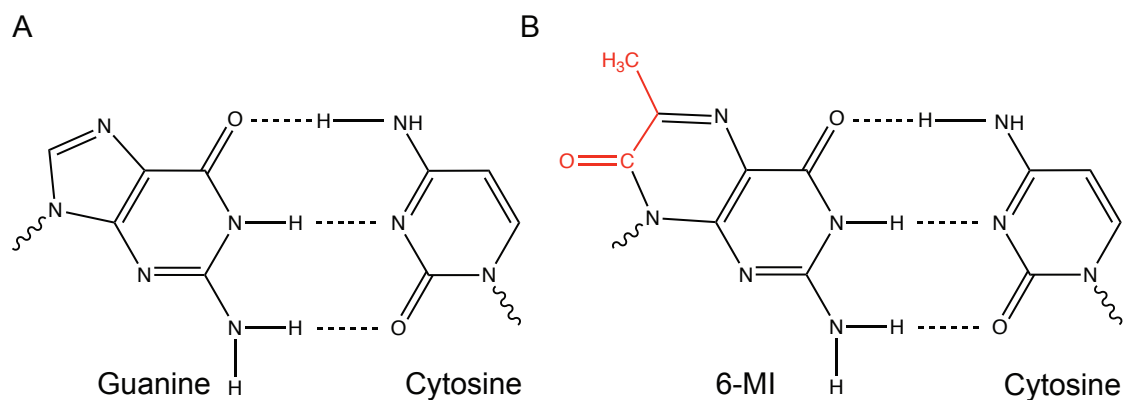
Deoxyribonucleic acid (DNA) carries genetic information that instructs and regulates genome expression. Fundamental gene regulational processes such as DNA replication, repair, and transcription all require precise reading and transferring of genetic information from DNA. Such biological information is encoded in the sequence of nucleotides and can be accessed by DNA-binding proteins. Since genetically coded base sequences are located 'inside' the DNA duplex, proteins that read-out base sequences must somehow access the DNA "interior". The accessibility of information from the DNA "interior" is facilitated by DNA "breathing", i.e. thermally driven DNA structural fluctuations at physiological temperatures that temporarily open up duplex DNA.<sup>1-4</sup>

How do we understand the sequence-dependent behavior of DNA dynamics and conformational heterogeneity? How are DNA-protein interactions achieved with such precision? What is the mechanism of recognition of a specific sequence? To begin to answer these questions requires knowledge of the local conformations of DNA at the level of individual nucleic acid bases.

Spectroscopy provides powerful tools to study conformations of DNA with instantaneous response and non-intrusive measurement during biological processes. However, overlapping spectral features across nucleic acid bases often hinder spectroscopic measurements that attempt to sensitively examine specific sites in DNA. To overcome this issue, fluorescent nucleic acid base analogues, which can be site-specifically substituted for natural bases, are often utilized to allow direct observation of regions of interests.<sup>1-3,5,6</sup>



This work focuses on the study of DNA local base conformation using 6-methyl Isoxanthopterin (6-MI) labeled DNA. 6-MI is a fluorescent guanine analogue that can serve as a sensitive reporter of DNA conformational changes at its labeled site. 6-MI structure is shown in Figure 1.1. 6-MI absorbs at a wavelength region that is transparent for canonical DNA bases and amino acids. 6-MI maintains hydrogen bonds with complementary cytosine. Depending on the flanking bases, its spectroscopic properties such as fluorescence, absorption and circular dichroism (CD) can be extremely sensitive to its local conformation and other environmental factors<sup>5,7,8</sup>.



**Figure 1.1:** The structure of (A) guanine/cytosine base pair and (B) 6-MI/cytosine base pair. Structural differences between guanine and 6-MI are emphasized in red.

In this work, I expanded upon the application possibilities of 6-MI as a fluorescent base analogue, and set important foundations for future research. By site-specifically positioning the 6-MI probe within DNA, my co-workers and I have demonstrated how one might extract base-sequence-dependent information about the local conformation of the 6-MI probe as modulated by the local base or base-pair environment.

We examined the conformational heterogeneity of 6-MI in various sequence contexts of duplex and ssDNA constructs. We gained more comprehensive understanding

of 6-MI electronic structure by studying the excitation–emission shift (EES) of the 6-MI substituted oligonucleotide constructs. Applying standard methods that were originally developed for nucleic acids by Schellman<sup>9–11</sup> Tinoco<sup>12,13</sup> and further extended by others<sup>14</sup>, I constructed protocols and computational modules to simulate circular dichroism in MATLAB and Python. I calculated CD of 6-MI substituted DNA constructs and achieved good agreement with experimental results. These studies served to establish a new level of sophistication in qualitatively analyzing 6-MI structural behavior in terms of local base stacking and unstacking conformations.

To further extend the potential of 6-MI as a useful reporter of local nucleic acid base conformation, I applied 6-MI to probe the ligand insertion of small molecules to duplex DNA, which is relevant to intercalation processes and protein-DNA interactions. I studied the interaction of duplex DNA with intercalator molecules Acridine Orange (AO) and 9-amino-6-chloro methoxyacridine (ACMA) that is covalently attached to the DNA backbone. I characterized 6-MI/AO Förster resonance energy transfer (FRET) pairs, thus establishing this as a meaningful system to study local DNA base conformation in future experiments involving single-molecule microscopy and multi-dimensional spectroscopy. We discovered that when 6-MI was substituted for guanine in DNA constructs, there was no significant preference or hindrance of AO intercalation at the 6-MI labeled site. Our results show that unlike free ACMA, which interacts with dsDNA via intercalation, tethered ACMA displaces the base on the complementary strand across from the ACMA attachment site. Duplex DNA constructs in which a tethered ACMA is placed directly opposite to a single 6-MI residue within the complementary strand may be an interesting model system to study base insertion and un-insertion processes. Such structures can be found in biologically important processes such as DNA methylation and DNA repair.

## **Dissertation Outline and Acknowledgements of the Contributions by Others**

This dissertation describes research done in the lab of Prof. Andrew H. Marcus and Prof. Peter H. von Hippel, with whom the experiments and calculations were co-developed, and the material was co-written. This dissertation also includes the contributions of following collaborators: Prof. Marina G. Guenza, Dr. Neil P. Johnson, Dr. Thomas H. Steinberg, Dr. Pablo G. Romano and Dr. Mohammadhasan Dinpajoo.

Chapter II presents the study of excitation–emission shift (**EES**) of the 6-MI substituted oligonucleotide constructs, and contains material co-authored with Neil P. Johnson, Thomas H. Steinberg, Peter H. von Hippel and Andrew H. Marcus. In these experiments, a variety of spectroscopic techniques were utilized to measure the EES of 6-MI. In this work, the fluorometry was performed by Neil P. Johnson and Thomas H. Steinberg. The two-state theoretical model of 6-MI electronic structure was co-developed with Andrew H. Marcus. In this chapter, we lay out the foundations of using the 6-MI probe as a tool to study sequence-dependent DNA local conformation. Manifested by the two-state model, we advanced our understanding of 6-MI spectroscopic properties that applied to studies described in later chapters.

Chapter III presents the theoretical treatments of circular dichroism (**CD**) calculations, which connect nucleic acid base conformation to spectroscopic observables. This chapter contains unpublished material co-authored with Neil P. Johnson, Mohammadhasan Dinpajoo, Marina G. Guenza, Peter H. von Hippel, and Andrew H. Marcus. We present the fundamentals of these CD calculations and the determination of critical spectroscopic parameters for this model. All of the CD calculations presented in my dissertation were based on the content described in this chapter. The description of CD calculations of the canonical base dinucleotides includes unpublished material co-

authored with Pablo G. Romano, Mohammadhasan Dinpajoooh and Marina G. Guenza. The results of molecular dynamics (MD) simulations used in the CD calculations were performed by Pablo Romano and Mohammadhasan Dinpajoooh in the Guenza group. The scheme to compare the results of MD simulation to CD using the matrix method was co-developed by Mohammadhasan Dinpajoooh, Pablo G. Romano, Marina G. Guenza, and Andrew H. Marcus.

Chapter IV presents the CD calculation conducted for 6-MI labeled DNA and DNA canonical dinucleotides. This chapter contains unpublished material co-authored with Neil P. Johnson, Peter H. von Hippel, and Andrew H. Marcus. Experimental CD and fluorometry were performed by Neil P. Johnson. We have found that single 6-MI substituted DNA constructs can be used to sensitively probe local nucleic acid base conformation. Qualitative agreement between calculations and experimental results reinforces the idea that the CD of 6-MI could be a useful probe of DNA secondary structure.

Chapter V presents the study of the interaction of the intercalator molecules Acridine Orange (AO) and 9-amino-6chloro methoxyacridine (ACMA) with 6-Methyl Isoxanthopterin (6-MI) substituted DNA constructs by steady state fluorometry and circular dichroism (CD) spectroscopy. This chapter contains unpublished material co-authored with Neil P. Johnson, Peter H. von Hippel and Andrew H. Marcus. Substantial work presented in this chapter that involves the determination and interpretation of experimental spectra of probe labeled DNA was carried out by Neil P. Johnson. We discovered that, in comparison to guanine, the 6-MI probe base does not significantly perturb the DNA lattice such that there is a preference or hindrance of intercalation at the 6-MI substitution site. Thus, 6-MI can be used as an accurate reporter of the intercalation

process. We subsequently explored the potential of using single site substituted 6-MI DNA constructs to measure ACMA, by labeling 6-MI on the complementary strand of ACMA within duplex DNA. Our results indicate that unlike free ACMA, which interacts with dsDNA via intercalation, tethered ACMA displaces the base on the complementary strand across from the ACMA attachment site. Thus, duplex DNA constructs in which a tethered ACMA is placed directly opposite to a single 6-MI residue within the complementary strand may be an interesting model system to study flipped out bases in dsDNA.

CHAPTER II

SEQUENCE-DEPENDENT CONFORMATIONAL HETEROGENEITY AND  
PROTON- TRANSFER REACTIVITY OF THE FLUORESCENT GUANINE  
ANALOGUE 6-METHYL ISOXANTHOPTERIN (6-MI) IN DNA

**Overview**

Chapter II presents the study of excitation–emission shift (EES) of the 6-MI substituted oligonucleotide constructs. This section is reprinted with permission from Neil P. Johnson, Huiying Ji, Thomas H. Steinberg, Peter H. von Hippel, and Andrew H. Marcus, “Sequence-Dependent Conformational Heterogeneity and Proton-Transfer Reactivity of the Fluorescent Guanine Analogue 6-Methyl Isoxanthopterin (6-MI) in DNA,” *J. Phys. Chem. B*, **2015**, *119* (40), pp 12798–12807. Copyright © 2015 American Chemical Society. In these experiments, a variety of spectroscopic techniques were utilized to measure the EES of 6-MI. In this work, the fluorometry was performed by Neil P. Johnson and Thomas H. Steinberg. The two-state theoretical model of 6-MI electronic structure was co-developed with Andrew H. Marcus. The additional co-authors provided editorial assistance. In this chapter, we lay out the foundations of using the 6-MI probe as a tool to study sequence-dependent DNA local conformation. Manifested by the two-state model, we advanced our understanding of 6-MI spectroscopic properties that applied to studies described in later chapters.

**Introduction**

A number of recent biophysical studies of DNA replication and transcription have relied on spectroscopic measurements that sensitively probe the local conformations and

dynamics of site-specific positions within protein–nucleic acid complexes. Some of these experiments use fluorescent chromophores that are incorporated into the nucleic acid backbone, or that are attached to a base using a flexible linker.<sup>1–3</sup> Another important class of fluorescently labeled nucleic acid constructs involves the site-specific substitution of one (or two) native base(s) within the nucleotide sequence by a fluorescent base “analogue”.<sup>4–9</sup> Such base-substituted DNA constructs are useful probes of local conformation, since the chromophores themselves can experience Watson–Crick base-pairing and base-stacking interactions similar to those of native bases. Moreover, the base analogues absorb light at significantly lower energies than do the canonical bases and proteins, so that they can be selectively excited and detected within large macro-molecular complexes.<sup>10–12</sup>

Previous experiments performed on fluorescent-base-analogue-substituted DNA constructs have used UV peak absorbance, fluorescence, and circular dichroism (CD) to study local base conformation. However, little attention has been devoted to understanding the presence of inhomogeneous broadening of the absorption line-shape and its relationship to local base conformational heterogeneity. When a base analogue chromophore is dissolved in a low-viscosity polarizable liquid such as water, changes to its charge distribution upon electronic excitation induce rapid structural reorganization of the local solvent. These excited-state relaxations occur on picosecond time scales, which lead to emission from the lowest energy level of the excited electronic state.<sup>13</sup> For a fluorescent base analogue that is substituted within a DNA strand, flanking bases can shield the chromophore from the aqueous solvent, so that local structural rearrangements might occur very slowly, possibly exceeding the nanosecond time scale of fluorescence. It is also possible that the fluorescent base analogue substituted within a DNA construct can

experience multiple local base conformations that are separated by relatively high activation barriers. If such conformational substates absorb at different energies, the resulting absorption spectrum is a weighted sum of substate contributions—a situation that leads to inhomogeneous broadening of the line shape.<sup>14</sup> Because the intrinsic lineshapes of individual substates are themselves broad, the cumulative absorption spectrum is often featureless and does not directly reveal the presence of multiple distinct species.

Information about the dispersion of local chromophore conformations and their interconversions can be determined using “site-selective” spectroscopic methods.<sup>13–16</sup> For example, when a spectrally narrow optical source is used to excite the system, the peak fluorescence energy can vary as the excitation energy is scanned across the absorption line.<sup>14</sup> It is sometimes found that the narrow distribution of substates selected using either red-edge or blue-edge excitation do not relax to a single globally minimized excited electronic state on the time scale of fluorescence. In such systems, a photoselected substate undergoes relaxation to its locally minimized conformation, which has its own characteristic emission energy. Thus, the presence of multiple substates comprising a heterogeneous mixture can affect the peak position of the fluorescence, depending on whether the system is excited on the blue-edge or the red-edge of the absorption line.<sup>14</sup>

The heterogeneity of chromophores in solution has been studied in high-viscosity solvents using various site-selective spectroscopies.<sup>14,17–19</sup> Such heterogeneity is most readily observed in polar solvents at low temperatures, for which the electronic transition dipole moment of the chromophore is strongly coupled to the slowly rearranging solvent shell. At elevated temperatures structural relaxations occur faster than the fluorescence lifetime, so that the system appears homogeneous on this time scale. When the environment relaxation is slower than fluorescence, site-selective measurements can provide



information about the presence of different probe microenvironments and their dynamical behavior. In this way, site-selective experiments have provided rough estimates of the mobility of chromophores in membranes<sup>20,21</sup> and in proteins.<sup>21,22</sup>

In this work, we apply steady-state excitation–emission spectroscopy to study conformational heterogeneity of the fluorescent guanine (G) analogue 6-methyl Isoxanthopterin (6-MI), which is site-specifically substituted for G in various DNA constructs. 6-MI was introduced as a fluorescent probe to study the local environments of G residues in DNA<sup>7,23–25</sup> (see Figure 1.1). The 6-MI:C base pair (bp) maintains the three Watson–Crick hydrogen bonds with C that are characteristic of G:C base pairs in duplex B-form DNA (dsDNA). The imino hydrogen at the N3 position of the 6-MI free base (6,8-dimethyl Isoxanthopterin) is reported to be slightly acidic, with  $pK_a = 8.4$ .<sup>26</sup> The behavior of 6-MI in DNA is similar to that of G with respect to its effect on the thermal stability of dsDNA, ligand access to the major groove, and template-dependent nucleotide addition by DNA polymerase (although subsequent elongation beyond the newly formed bp 6-MI:C is inhibited).<sup>5</sup>

The lowest lying electronic transition ( $S1 \leftarrow S0$ ) of the 6-MI chromophore in aqueous buffer is centered at  $\sim 340$  nm, and its peak emission lies close to 425 nm.<sup>27</sup> The  $S1$  excited state is characterized by a 6 ns lifetime and a fluorescence quantum yield of  $\sim 0.7$ . Two faster population relaxation processes ( $\sim 3$  ns,  $\leq 1$  ns) appear when 6-MI is substituted for G within single- stranded (ss) or double-stranded (ds) DNA.<sup>23</sup> The fluorescence intensity of 6-MI becomes progressively quenched in ss- and dsDNA constructs, which is thought to reflect effects of base stacking.<sup>5,23</sup>

Here we examine the conformational heterogeneity of 6-MI in various sequence contexts of duplex and ssDNA constructs. Accessibility of the weakly acidic N3 hydrogen

to potential hydrogen bonding acceptors depends on the local 6-MI environment, thus providing a spectroscopic signature to probe local base conformational heterogeneity. We find that conformational heterogeneities of 6-MI that are stable on time scales exceeding the fluorescence lifetime (nanoseconds or greater) can be observed in substituted DNA constructs in aqueous solution at room temperature. This local conformational heterogeneity of the 6-MI probe base depends significantly on DNA sequence context and strandedness.

### **Materials and Methods**

6-MI-substituted DNA constructs were purchased from Integrated DNA Technologies (Coralville, IA) and from Fidelity Systems (Gaithersburg, MD). 6-MI ribonucleotide monophosphate (NMP) was obtained from Fidelity Systems, and used without further purification. The sequences and nomenclatures of the DNA constructs are shown in Table 2.1. Oligonucleotide concentrations were determined using extinction coefficients provided by the manufacturer. Duplex DNA constructs were annealed by combining equimolar solutions of complementary strands, followed by heating the mixture to 90 °C and allowing it to gradually cool. Solutions were buffered using the following conditions: pH 7.6 (50 mM Tris, 150 mM NaCl); pH 6.2 (50 mM phosphate buffer, 150 mM NaCl). For titrations of 6-MI NMP, samples were prepared in 10 mM sodium pyrophosphate, pH 6–10.3, buffer conditions previously used to determine the pKa of 6,8-dimethylisoxanthopterin.<sup>26</sup> Unless otherwise stated, experiments were carried out at 20 °C.

Fluorescence spectra and steady-state fluorescence anisotropy were measured using a Jobin–Yvon FluoroMax-3 spectrophotometer. The samples were excited over the wavelength range 300–400 nm, and the emission spectra were measured over the 300–375

nm wavelength range, or in some cases over the 380–500 nm wavelength range. All spectra were corrected for background contributions from the buffer solution itself. The peak emission wavelengths were determined by fitting the emission spectra to a polynomial or a Gaussian function. The error associated with the peak emission wavelength determined by repetition of experiments was  $\pm 1$  nm ( $\sim \pm 100$ – $200$   $\text{cm}^{-1}$ ).

CD spectra of 4  $\mu\text{M}$  oligonucleotide solutions were measured using a Jasco model J-720 CD spectrophotometer with a temperature-controlled cell holder, as described previously.<sup>5</sup> The CD spectra were measured over the 300–450 nm wavelength range. Each measurement was tested for reproducibility by repeating 8–20 times, and the resulting data were averaged together. CD spectra are defined in terms of the difference in extinction of the sample excited using left- and right-circularly polarized light:  $\epsilon_L - \epsilon_R$  (in  $\text{M}^{-1} \text{cm}^{-1}$ ) per mole of 6-MI residue.

For DNA melting experiments, the UV absorbance was monitored at 260 nm using a Cary UV spectrophotometer equipped with a Peltier temperature controller. Samples were heated from 20–85  $^{\circ}\text{C}$  at a rate of 1  $^{\circ}\text{C} \text{min}^{-1}$ , and the absorbance was measured at 1  $^{\circ}\text{C}$  intervals.

**Table 2.1.** Base sequences and nomenclature for the 6-MI-containing ssDNA constructs used in in Chapter II.. The letter **X** indicate the 6-MI probe. The regions that contain the 6-MI analogue and flanking bases are highlighted in red. Duplex DNAs are the same molecules annealed with a fully complementary strand.

DNA construct	Nucleotide base sequence
ssGXG	5'-CTA ATC ATT <b>GXG</b> TTC GGT CCT TGC-3'
ssGXXG	5'-CTA ATC ATT <b>GXX</b> GTT CGG TCC TTG C-3'
ssTXT	5'- CTA ATC ATT <b>GTX</b> TTC GGT CCT TGC -3'
ssTXXT	5'- CTA ATC ATT <b>GTX</b> <b>XTT</b> CGG TCC TTG C -3'
ssTXG	5'- CTA ATC ATT <b>GTX</b> <b>GTT</b> CGG TCC TTG C -3'
ssTXXG	5'- CTA ATC ATT <b>XTX</b> <b>GTT</b> CGG TCC TTG C -3'

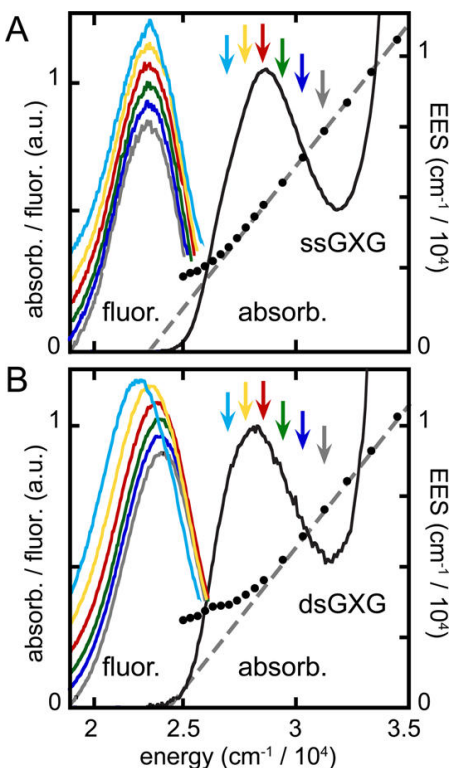
## Results

### Excitation-Emission Shift (EES)

We performed excitation–emission studies with the 6-MI substituted oligonucleotide constructs with sequences shown in Table 2.1. In some cases, the peak emission wavelength was relatively insensitive to the excitation energy. For example, when we excited the ssGXG sample over the 320–375 nm ( $26\,666$ – $31\,250\text{ cm}^{-1}$ ) wavelength range, the peak emission occurred at 425 nm ( $23\,529\text{ cm}^{-1}$ ), irrespective of the excitation wavelength (Figure 2.1 A). We define the excitation–emission shift (EES) as the difference between the energy of the exciting optical source and that of the peak emission, and we plot this function versus excitation energy in Figure 2.1 A (black filled circles). For this sample, the EES appeared to be a linear function of the excitation energy, with slope equal to approximately unity (0.97) over nearly the full range of excitation energies investigated. Deviation of the EES from linearity occurred only at the extreme red-edge of the absorption band.

In other cases, the plot of EES versus excitation energy exhibited a pronounced demarcation of behaviors. We illustrate this using the dsGXG construct (see Figure 2.1 B). At excitation energies much greater than that of the peak absorbance ( $\sim 350\text{ nm}$ ,  $28\,571\text{ cm}^{-1}$ ), the EES scaled linearly, similar to the behavior we observed for the ssGXG sample. However, as the excitation energy approached the absorption maximum, the emission peak began to shift so that the EES became a weaker function of excitation energy. Thus, for the dsGXG construct, it was possible to identify two distinct regimes of behavior. At high excitation energy the EES scaled linearly, while at intermediate energies a transition to nonlinear scaling was observed. These results suggest the presence of more than

one conformational substate, each with distinct electronic properties that persist on time scales longer than the fluorescence lifetime.

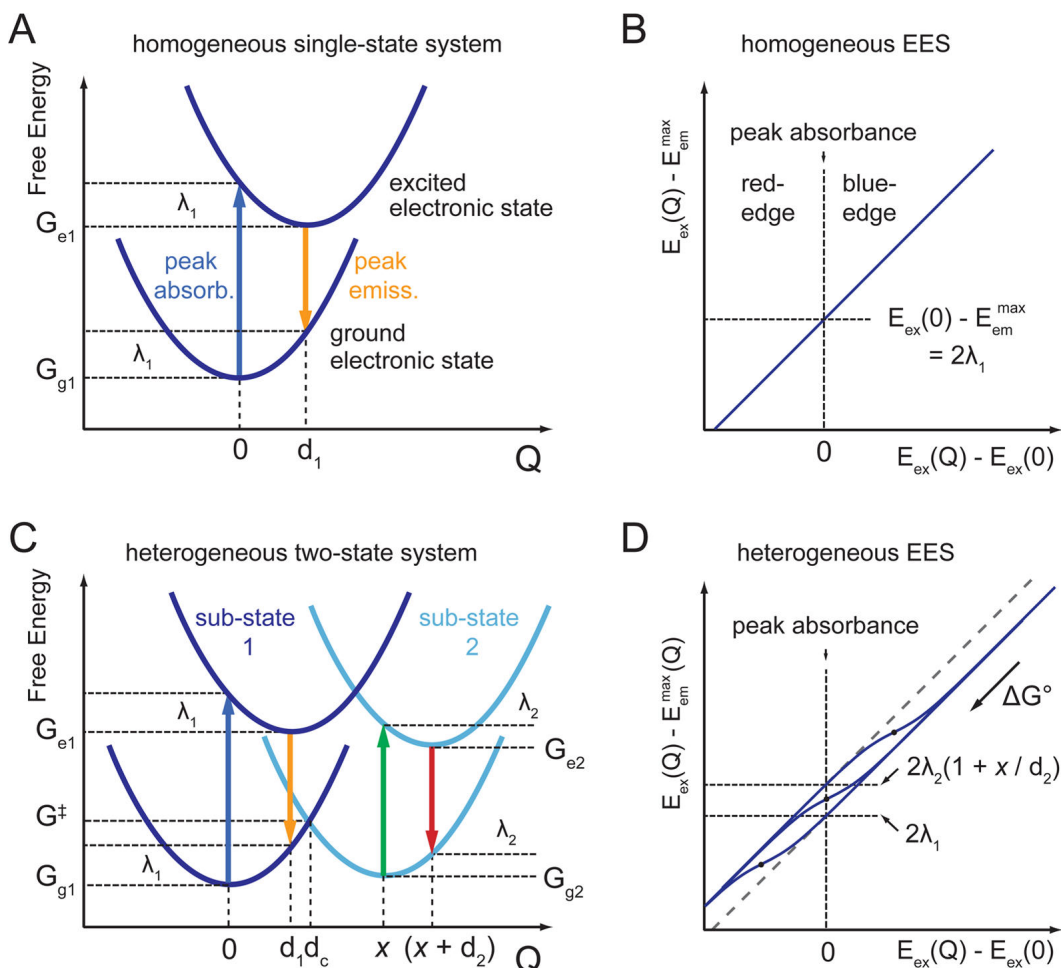


**Figure 2.1:** Excitation–emission shift (EES). Absorbance and emission spectra of **(A)** ssGXXG and **(B)** dsGXXG constructs, taken at 20 °C and pH = 7.6. Emission spectra were excited over the wavelength range 320–370 nm in 10 nm increments at the energies indicated by vertical arrows. Normalized emission spectra are shown in the corresponding color, and vertically displaced for clarity. Black circles are the “excitation–emission shift” (EES), i.e., the difference between the excitation and peak emission energies. Gray dashed lines are linear fits to the blue-edge of the EES data.

### Theoretical Considerations of the EES

To describe the relationship between the EES of 6-MI substituted DNA constructs and the presence of multiple conformational substates, we applied standard principles of molecular photochemistry.<sup>13, 28</sup> The mathematical details of the model are presented in the Appendix A.

We envision a multidimensional free energy hypersurface to characterize the ground and excited electronic states, which are coupled to nuclear motions (see Figure 2.2).



**Figure 2.2:** Homogeneous one-state model and heterogeneous two-state model for excitation–emission shift (EES) experiments. Panels (A) and (C) show hypothetical free energy surfaces associated with ground and excited electronic states, both as a function of the generalized conformation parameter  $Q$ . Panels (B) and (D) show the expected EES,  $Y \equiv E_{\text{ex}}(Q) - E_{\text{em}}^{\text{max}}(Q)$  (solid blue lines) as a function of the rescaled excitation energy [ $X \equiv E_{\text{ex}}(Q) - E_{\text{ex}}(0)$ ], where  $E_{\text{ex}}(0)$  is the energy of the peak absorbance. Panels A and B depict the situation for an optically homogeneous sample. Panels C and D depict an optically heterogeneous two-state sample with an activation barrier  $\Delta G^\ddagger = G^\ddagger - G_{g1}$ . For the heterogeneous two-state system, the EES undergoes an excitation energy-dependent transition between the two substates, as depicted in panel D for three different energetic conditions.

It is useful to think about a one-dimensional projection of this surface onto a single generalized nuclear coordinate  $Q$ , which serves as a reaction coordinate for state-to-state interconversion. The shape of the ground electronic surface, including the position of its minimum (defined at  $Q = 0$ ), is determined by the equilibrium distribution of local

environment configurations. It follows from thermodynamic principles that these surfaces can be well-approximated as parabolic functions near their minima.<sup>13</sup> Upon optical excitation, the induced polarization leads to relaxation of the local environment, which occurs much faster than the fluorescence lifetime. The equilibrium distribution of local configurations in the excited state is different from that in the ground state, so that the minimum of the excited electronic surface is displaced at  $Q = d_1$ . The extent to which the excited state is stabilized by environment relaxation is characterized by the reorganization free energy  $\lambda_1$ , which is the average optical energy that is absorbed and converted into heat before emission of fluorescence. The Stokes shift is the energy difference between the peaks of the absorption and fluorescence spectra. It can be understood from simple geometrical considerations that the Stokes shift is equal to twice the reorganization free energy  $2\lambda_1$  (see Appendix A).

For the ssGXG sample discussed above (see Figure 2.2 A), local conformations experienced by the 6-MI probe must rapidly interconvert, so that the ground and excited free energy surfaces may be described as single parabolic wells. This model is representative of an optically homogeneous system, for which the energy gap between ground and excited electronic surfaces is parametrized by the coordinate  $Q$ . The peak emission energy  $E_{em}^{max}$  is the energy gap at  $Q = d_1$ , which for a homogeneous system is independent of excitation energy.

We define the EES as the difference between the excitation and peak emission energies,  $Y \equiv E_{ex}(Q) - E_{em}^{max}$ . The energy of the peak absorbance of the 6-MI substituted DNA construct can vary depending on base sequence context and secondary structure<sup>27</sup>. We therefore define the rescaled excitation energy,  $X \equiv E_{ex}(Q) - E_{ex}(0)$ , where  $E_{ex}(0)$  is the peak absorption energy. It is straightforward to show that the EES plot (i.e. a graph

of  $Y$  versus  $X$ ) for an optically homogeneous system is a line with slope equal to unity, and  $y$ -intercept equal to the Stokes' shift  $2\lambda_1$  (see Appendix A and Figure 2.2 B).

Our EES experiments performed on dsGXG indicate that this system is optically heterogeneous (see Figure 2.1 B). Molecules excited at the red-edge of the absorption band represent energetically selectable substates, which do not readily interconvert with substates that are excited using higher energy light. The observed shift in peak emission energy with changing excitation is due to the presence of distinct conformations or chemical species and local environments of the 6-MI probe, each with spectroscopically resolvable optical transition energies.

To explain these observations, we consider a two-state model, which captures the essential elements of the problem (see Figure 2.2 C).<sup>29</sup> We assume that the 6-MI probe may adopt one of two possible conformations (labeled 1 and 2) in duplex DNA, each with ground and excited electronic free energy surfaces that depend on the generalized coordinate  $Q$ . As we shall see, these two conformations may be assigned to protonated and deprotonated forms of 6-MI. The ground-state surface on the left (shown in dark blue) represents substate 1 at  $Q = 0$ , while the surface on the right (shown in light blue) represents substate 2 at  $Q = x$ . Optical excitation of the system on the blue-edge of the absorption band favors substate 1, while excitation on the red-edge favors substate 2. The free energy of activation,  $\Delta G^\ddagger$ , between the two substates occurs at the coordinate  $Q = d_c$ , where the ground-state surfaces cross. The effect of tuning the excitation energy across the absorption band is to adjust the weights of substates 1 and 2 that are electronically excited, and thus to tune the peak energy of the fluorescence. This model can explain the observed fluorescence peak shift as a function of excitation energy.



We simplify the model, and thus reduce the number of descriptive parameters, by making the approximation that each substate is characterized by the same reorganization free energy ( $\lambda_1 = \lambda_2 = \lambda$ ), and by the same excited electronic state displacement ( $d_1 = d_2 = d$ ). The EES can then be described by the equation

$$Y = X + 2\lambda \left[ 1 + P_2 \left( \frac{x}{d} \right) \right] \quad (2.1)$$

where we have used the same definitions for  $X$  and  $Y$  as above for the homogeneous one-state system (see Appendix A for derivation). In Eq. (2.1),  $P_2 = 1 - P_1$ , is the Boltzmann population associated with sub-states 2. In the limit that only sub-state 1 is thermally populated ( $P_2 = 0$ ), we see that Eq. (2.1) recovers the behavior of the homogeneous system. In the opposite limit ( $P_2 = 1$ ), Eq. (1) reduces to a line with slope equal to unity and  $y$ -intercept  $2\lambda(1 + x/d)$ . These two limiting behaviors are depicted in Fig. 3D as vertically displaced dashed gray lines. When both sub-states are thermally populated, there exists a narrow range of excitation energies over which the EES makes a transition from one linear regime to the other, as shown in Fig. 3D. An inflection point occurs at the transition, which is specified by  $Q = d_c$ , and for which each sub-state is equally populated ( $P_1 = P_2 = 1/2$ ). The coordinates of the inflection point are given by  $X_c = -2(\lambda/d)d_c$  and  $Y_c = X_c + \lambda[2 + (x/d)]$ . The free energy of transition and the free energy of activation are given by the Marcus formulas<sup>13, 28</sup>

$$\Delta G^\circ = G_{g2} - G_{g1} = -\frac{1}{2}\alpha x^2 + \alpha x d_c, \quad (2.2)$$

and

$$\Delta G^\ddagger = \frac{1}{2} \alpha d_c^2, \quad (2.3)$$

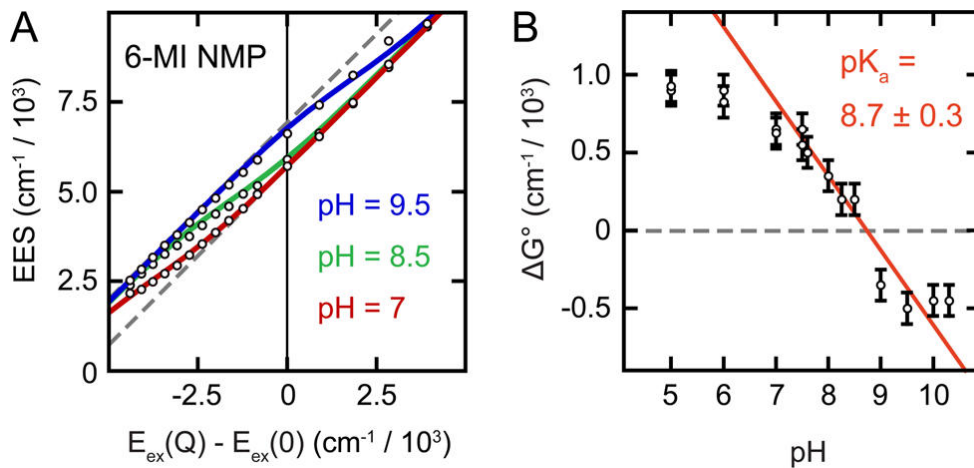
where  $\alpha = 2\lambda/d^2$  is a “polarizability” that characterizes the curvature of the free energy surfaces. According to Eq. (2.2), the free energy of reaction increases linearly with respect to  $d_c$ . Thus, an equilibrium system that favors substate 1 ( $\Delta G^\circ > 0$ ) will tend to exhibit an EES transition on the low-energy side (i.e., the red-edge) of the absorption band, while a system that favors substate 2 ( $\Delta G^\circ < 0$ ) will tend to exhibit a transition on the blue-edge of the absorption band. An isoenergetic system ( $\Delta G^\circ = 0$ ) will exhibit a transition on the red-edge of the absorption peak, at the coordinates  $X_c = -(\lambda/d)x$  and  $Y_c = 2\lambda$ .

A quantitative fit of EES data to Eq. (2.1), or to more general versions of the model (e.g. with  $\lambda_1 \neq \lambda_2$  and/or  $d_1 \neq d_2$ , see SI section), can be used to obtain optimized values for the parameters  $d_c$ ,  $x$ ,  $d_{1,2}$ ,  $\lambda_{1,2}$ , and through Eq. (2.2) and (2.3), values for  $\Delta G^\circ$  and  $\Delta G^\ddagger$ . In the remainder of this work, we have specialized to the case  $d_1 = d_2 = 1$ , and specific values of the nuclear coordinate  $Q$  (i.e.,  $d_c$ , and  $x$ ) are determined relative to the excited-state displacement. As mentioned previously, this simplification serves to reduce the number of parameters needed to faithfully represent our experimental data, while retaining the essential elements of the two-state model.

### **EES of 6-MI Ribonucleotide Monophosphate (NMP)**

As mentioned previously, the 6-MI chromophore is also a weak acid with ionizable hydrogen at the N3 (Watson–Crick) position (see Figure 1.1), which is characterized by a  $pK_a$  value that is considerably lower than that of the equivalent position (N1, see Figure 1.1) in guanine, and thus falls into the physiologically relevant pH range.<sup>26</sup> 6-MI NMP is soluble in aqueous solution, and the equilibrium concentrations of protonated and deprotonated forms depend on pH. To test whether the optical heterogeneity we observed

in 6-MI substituted DNA systems is related to the acid–base properties of the 6-MI probe, we performed pH-dependent EES experiments on aqueous solutions of 6-MI NMP.



**Figure 2.3:** EES data and model fits of 6-MI NMP in aqueous solutions for different pH conditions. **(A)** The EES [ $Y \equiv E_{\text{ex}}(Q) - E_{\text{em}}^{\text{max}}(Q)$ ] for 6-MI NMP is plotted versus the rescaled excitation energy  $X \equiv E_{\text{ex}}(Q) - E_{\text{ex}}(0)$ , where  $E_{\text{ex}}(0)$  is the energy of the peak absorbance. Solid curves are fits of Eq. (2.1) to the data. **(B)** The standard Gibbs free energy  $\Delta G^\circ$  was determined from an analysis of similar data over the pH range 5–10.3 (see Figure A2.1). A fit of these data to the acid–base equilibrium equation  $\Delta G^\circ = 2.3 RT (pK_a - \text{pH})$  (shown in red) yields the value  $pK_a = 8.7 \pm 0.3$ .

In Figure 2.3 A, we present examples of our pH-dependent EES measurements at pH = 7, 8.5, and 9.5. These data were fit to Eq. (2.1), and the model curves are shown superimposed on the experimental data points. Details of the fitting procedure are provided in the Appendix A. We see that, by increasing the pH through the vicinity of the  $pK_a$  (=8.4), the inflection points of the experimental EES curves move systematically toward increasing excitation energy, as predicted by the two-state model described above. We obtained excellent agreement between Eq. (2.1) and the experimental EES data over the full range of pH values we investigated (5–10.5). In Figure A.1, we show additional EES data and model fits for 6-MI NMP taken at other pH values. We obtained identical results from these experiments using pyrophosphate buffer, and independently using 50 mM Tris,

150 mM NaCl. In Table A.1, we list the optimized parameters that we determined from these experiments.

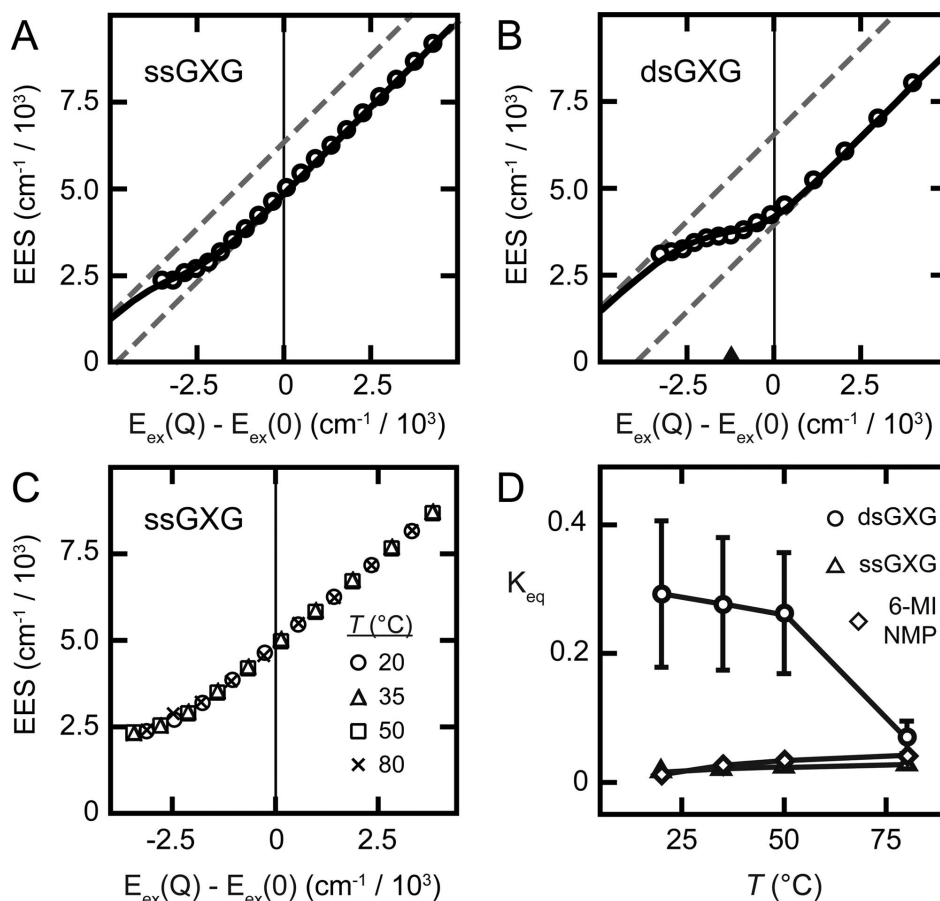
From the above analyses, we determined the reorganization free energy ( $\lambda = 2850 \pm 40 \text{ cm}^{-1}$ ) and the displacement of the ground electronic surface of the deprotonated state relative to the protonated one ( $x/d = 0.21$ ). We found that the optimized values we obtained for  $\lambda$  and  $x/d$  were nearly the same for all of the pH conditions that we examined, while the optimized values for  $\Delta G^\circ$  and  $\Delta G^\ddagger$  varied systematically. We compared these values to those predicted by the principles of acid–base equilibria. For the acid–base reaction  $\text{HA} \rightleftharpoons \text{H}^+ + \text{A}^-$ , we may write the equilibrium condition  $[\text{A}^-]/[\text{HA}] = 10^{(\text{pH} - \text{p}K_a)}$ . This ratio is equal to the equilibrium constant  $K_{\text{eq}}$  for the reaction 6-MI (protonated)  $\rightleftharpoons$  6-MI<sup>-</sup> (deprotonated). Thus, for a particular value of  $\text{p}K_a$ , we determined theoretical values for  $\Delta G^\circ = -RT \ln K_{\text{eq}}$ . In Figure 2.3 B, we plot the experimental values we obtained for  $\Delta G^\circ$  as a function of pH. The red line is the best fit of the function  $\Delta G^\circ = 2.3RT (\text{p}K_a - \text{pH})$  to these data, which yields the value  $\text{p}K_a = 8.7 \pm 0.3$ . We note that theory and experiment agree very well over the pH range 7–10.

The  $\text{p}K_a$  value we determined for 6-MI NMP ( $= 8.7 \pm 0.3$ ) is higher than that previously reported for the 6-MI base ( $= 8.4 \pm 0.1$ , in the form of 6,8DMI, 6,8-dimethylisoxanthopterin).<sup>26</sup> The larger  $\text{p}K_a$  value of the 6-MI NMP may be due to the presence of the phosphate group, which can stabilize the uncharged protonated form. Furthermore, the phosphate group of 6-MI NMP is itself ionizable with  $\text{p}K_a = 6\text{--}7$ .<sup>30</sup> The presence of this second ionizable site in 6-MI NMP is likely the reason for the departure of the experimentally observed values of  $\Delta G^\circ$  from theory at low pH ( $< 7$ , see Figure 2.3 B). This is because protonation of the phosphate group at low pH decreases its negative charge, which in turn stabilizes the deprotonated state of the 6-MI N3 site leading to a

decrease in the observed value of  $\Delta G^\circ$ . Our results suggest that the acid–base chemistry of 6-MI NMP in aqueous solution is responsible for the spectral heterogeneity apparent in the EES data for these samples.

### EES of 6-MI Substituted DNA Constructs

In Figure 2.4, we present experimental EES data from single-stranded (Figure 2.4 A) and duplex (Figure 2.4 B) DNA samples containing the GXG sequence. Superimposed on these are theoretical curves derived from our analysis using the heterogeneous two-state model (see previous and Appendix A).



**Figure 2.4:** EES data for 6-MI substituted DNA constructs. Experimental EES data and their corresponding fits to the heterogeneous two-state model are shown for (A) ssGXG and (B) dsGXG, as a function of the rescaled excitation energy at 20 °C. X = 6-MI. (C) EES data for ssGXG at 20 °C ( $\circ$ ), 35 °C ( $\Delta$ ), 50 °C ( $\square$ ), and 80 °C ( $\times$ ). (D) Temperature dependence of  $K_{\text{eq}} = [\text{6-MI}^- \text{ (deprotonated)}] / [\text{6-MI} \text{ (protonated)}]$  for the acid–base reaction of ssGXG, dsGXG, and 6-MI NMP. Oligonucleotide sequences are shown in Table 2.1.

The resulting values of  $\Delta G^\circ$  and  $\Delta G^\ddagger$  for the proton-transfer reaction show that the anionic form of 6-MI is more stable in dsGXG ( $K_{\text{eq}} = [\text{6-MI}^- \text{ (deprotonated)}]/[\text{6-MI (protonated)}] = 0.3 \pm 0.11$ ) than in ssGXG ( $K_{\text{eq}} = 0.03 \pm 0.01$ ). We found that EES data for ssDNA samples were insensitive to temperature variation, which is very similar behavior to that we found for the 6-MI NMP samples. In Figure 2.4 C, we show as an example EES data sets for the ssGXG sample taken over the 20–80 °C temperature range. On the other hand, our analyses of EES data for the duplex DNA samples did exhibit temperature-dependent changes in the values of  $K_{\text{eq}}$ . In Figure 2.4 D, we plot the values of  $K_{\text{eq}}$  for the dsGXG sample as a function of temperature, which decrease with increasing temperature. Upon complete thermal denaturation of the DNA duplex, the value of  $K_{\text{eq}}$  for the dsGXG sample reaches that obtained for ssGXG.

The dsGXG sample undergoes cooperative thermal melting with midpoint transition temperature  $T_m = 74$  °C. Because the value of  $K_{\text{eq}}$  also appears to undergo a cooperative change over this temperature range, we conclude that the enhanced stabilization of the deprotonated state we observed in room temperature duplex DNA is a consequence of secondary structure.

We note that the ssGXG sample also unstacks over this temperature range, albeit noncooperatively and to a lesser extent than does dsDNA. If base stacking contributed to the temperature effects observed for the dsDNA samples, we might expect to observe a similar decreasing trend in the values of  $K_{\text{eq}}$  for the ssGXG sample over this temperature range. On the contrary, the value of  $K_{\text{eq}}$  for this sample appears to increase only very slightly over this range, and its behavior is indistinguishable from that of the 6-MI NMP.

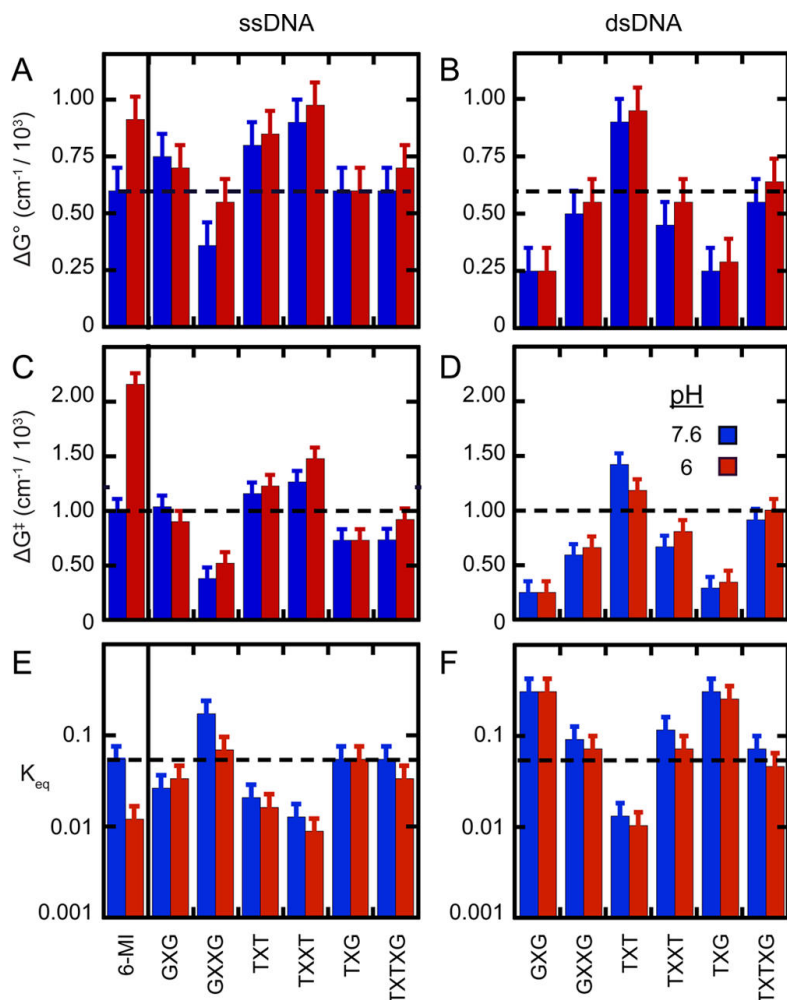
To better understand the proton-transfer reactivity of 6-MI-substituted DNA constructs, we examined the behavior of the 6-MI probe in various sequence contexts (see

Table 2.1). In Figure A.3, we show comparisons between experimental data for all DNA constructs and model curves, which are based on our two-state analysis.

For the case of ssDNA constructs, we obtained best fits to our data when we set  $\lambda_1 = \lambda_2$ , and when the value of  $x$  was nearly the same for all samples (see Table A.2). We thus found that the optimization conditions required to model 6-MI substituted ssDNA constructs were very similar to those for the 6-MI NMP samples previously discussed, although the displacement of the ground-state surface of the deprotonated state for ssDNA constructs ( $x \approx 0.31$ ) was somewhat larger than for 6-MI NMP ( $x \approx 0.21$ ). This result indicates that 6-MI in ssDNA requires a greater displacement of its proton along the reaction coordinate to achieve its deprotonated form, in comparison to 6-MI NMP in aqueous solution. The excellent agreement between theory and experiment for all of the ssDNA constructs indicates that the optical heterogeneity we observed in these systems can be explained using the two-state model. In contrast, in order to obtain agreement between theory and experiment for dsDNA, it was necessary to allow  $\lambda_1$  and  $\lambda_2$  to vary independently, and for the value of  $x$  to vary within the range 0.2–0.4 (see Table A.3). The results of this analysis yielded values for  $\Delta G^\circ$ ,  $\Delta G^\ddagger$ , and  $K_{eq}$  characteristic of the equilibrium between protonated and deprotonated states of 6-MI substituted DNA with different sequence contexts. These values are compared graphically in Figure 2.5, and are reported in Tables A.2 and A.3.

To further test whether the effects we observed are specific to the 6-MI probe, we examined the EES for single-stranded and duplex DNA constructs that contained the fluorescent cytosine analogue pyrrolocytosine (PC), and the fluorescent adenine analogue 2-aminopurine (2-AP). In Figure S2, we present EES data for single-stranded and dsDNA

constructs containing the sequences C(PC)A or G(2-AP)G (see Table 2.1 for complete sequences).



**Figure 2.5:** Standard free energy of reaction, free energy of activation, and equilibrium constant for proton transfer in 6-MI substituted DNA constructs. Standard free energies of reaction  $\Delta G^\circ$  (A, B), free energies of activation  $\Delta G^\ddagger$  (C, D), and deprotonation reaction equilibrium constant  $K_{eq}$  (E, F) of ssDNA (left) and dsDNA (right) at pH = 7.6 (blue) and pH = 6 (red). Horizontal dashed lines indicate the values for 6-MI NMP at pH = 7.6. Representative fits of EES curves for all single-stranded and duplex DNA constructs are shown in Figure A2.2. The optimized values of the fitting parameters are presented in Table A.2 and Table A.3.

In each case, the EES data appear to behave as an optically homogeneous one-state system. The pKa values of PC and 2-AP are less than 4,<sup>6,31</sup> so that both of these analogues are expected to exist entirely in their deprotonated forms near neutral pH. Thus, our



observations support the conjecture that a proton-transfer reaction plays a role in the optical heterogeneity we have observed in 6-MI-substituted DNA constructs.

We note that the local environment experienced by 6-MI in DNA is likely very different from that of 6-MI NMP in aqueous solution. The effects of stacking with neighboring bases in both single-stranded and duplex DNA constructs, in addition to the further screening from the solvent due to Watson–Crick base-pairing complementarity in the duplex constructs, are expected to limit accessibility of the aqueous solvent to the 6-MI probe. To test this conjecture, we performed EES experiments on 6-MI substituted single-stranded and duplex DNA at two different pH values (6 and 7.6). We found for nearly all of the DNA constructs we investigated, our results are experimentally indistinguishable at the two different pH values (see Figure 2.5, and Figures A.2 and A.3). In contrast, this same variation in pH shifts the equilibrium concentrations of 6-MI NMP between its protonated and deprotonated forms by nearly an order of magnitude (Figure 2.5 and Figure 2.1). This apparent insensitivity to pH of the EES data of 6-MI-substituted ss- and dsDNA constructs could be interpreted in multiple ways. For the case of duplex DNA constructs, variations in pH over this narrow range may not influence the local environment of the 6-MI probe since the cooperatively stabilized structure of the dsDNA in which the 6-MI probe is contained is also insensitive to pH over this range.<sup>32</sup> On the other hand, it is possible that these results indicate that the participation of solvent water does not dominate the proton-transfer reaction for either ss or duplex DNA, which instead depends on the local proximity of adjacent or opposing DNA bases.

A notable exception to the general behavior described above occurs for the ssGXXG construct (Figure 2.5 E). In this case, the change in pH shifts the equilibrium concentrations between protonated and deprotonated forms by a factor of 2.6. It is possible

that the ssGXXG sequence somewhat destabilizes the local ssDNA stacking, thus increasing the exposure of the 6-MI dimer probe to the solvent environment, although much less so than 6-MI NMP in aqueous solution. To obtain some information about local base conformation within the single-stranded and duplex GXXG and TXXT constructs, we performed circular dichroism (CD) experiments at pH = 7.6 (see Figure A.5). CD spectra for these constructs are sensitive to the exciton interactions between nearest neighbor 6-MI probes, which depend on the degree of base-stacking interactions. The CD spectra of the ssGXXG and ssTXXT constructs are nearly identical, and the signs and magnitudes of the CD indicate that the two 6-MI residues are stacked in a right-handed B-form conformation (Figure A.5). We speculate that some previously unconsidered interaction between the two 6-MI probes (e.g., a minor pH-dependent conformation change) may render the ssGXXG sequence sensitive to pH.

Our findings allow us to make some general statements pertaining to the local environment experienced by the 6-MI probe in DNA. In 6-MI-substituted ssDNA constructs, the protonated form of the 6-MI probe can be stabilized relative to that of the 6-MI NMP in solution (see Figure 2.5 E). The effect appears to be most pronounced when 6-MI is positioned adjacent to a T, and much less so when 6-MI is next to a G. Exceptions are the sequences GXXG (discussed above), TXG and TXXG. The latter represent combinations of 6-MI next to flanking T and G, which may lead to an overall cancellation of effects. When 6-MI is positioned adjacent to a G (or to another 6-MI), the effect of hybridization in duplex DNA is to stabilize the deprotonated form of the probe relative to the equilibrium in its corresponding single-stranded construct (and in the NMP). This enhanced stability of the deprotonated 6-MI in dsDNA is absent for TXT and TXXG sequences with probe residue flanked on both sides by a T (see Figure 2.5 F). Sequences

containing G or (6-MI):C base pairs are expected to be more stable than sequences containing A:T base pairs. These results indicate that  $K_{\text{eq}} = [6\text{-MI}^- \text{ (deprotonated)}] / [6\text{-MI (protonated)}]$  increases in stable duplex structures.

## Discussion and Conclusions

We have studied the optical heterogeneity of the fluorescent guanine analogue 6-MI, which we have shown to be sensitive to pH changes near neutrality when it is dissolved in aqueous solution, or to its local macromolecular environment when it is substituted for G in single-stranded and duplex DNA. We have quantified the degree of optical heterogeneity of 6-MI using steady-state excitation–emission shift (EES) measurements, in which the difference between the peak emission and excitation energies is recorded as a function of excitation energy. We demonstrated that the EES for 6-MI NMP in solution, and 6-MI-substituted DNA constructs, are well-described using a heterogeneous two-state model. On the basis of this approach, we determined the values of  $\Delta G^\circ$  and  $\Delta G^\ddagger$ , the standard free energies of reaction and activation, to characterize the equilibrium distribution of the two substates. This approach can potentially be generalized to provide information about fluorescent probes of local macromolecular structure other than 6-MI, which may exist in multiple substates that are stable on the time scale of the fluorescence lifetime.

By performing EES experiments on 6-MI NMP in aqueous solution, we determined that the distribution of the two substates is pH-dependent. These data can be understood in terms of simple acid–base chemistry, in which the first substate was identified as the protonated species, and the second substate as the deprotonated species. As expected for two-state behavior, EES data obtained from the 6-MI NMP system could be self-

consistently modeled for all values of pH using the same value of the parameter  $x$ , which represents the relative displacement of the protonated and deprotonated ground-state free energy surfaces. We obtained excellent agreement between theory and experiment under the constraint  $\lambda_1 = \lambda_2$ . We observed similar behavior for 6-MI-substituted ssDNA constructs (see below). We calculated the equilibrium constant  $K_{\text{eq}} = [\text{6-MI}^- \text{ (deprotonated)}] / [\text{6-MI (protonated)}]$  from the value of  $\Delta G^\circ$ , which we determined by fitting the EES data to the two-state model. The resulting value for the  $\text{p}K_{\text{a}}$  of 6-MI ( $8.7 \pm 0.3$ , see Figure 2.3 B) agrees very well with that previously reported in the literature, where it was established that the ionizable proton resides at the N3 position of 6-MI<sup>26</sup>. We note that the  $\text{p}K_{\text{a}}$  values determined from fluorescence measurements are the same as those determined by UV absorbance,<sup>26</sup> suggesting that the pH-dependent optical heterogeneity of 6-MI NMP reflects a ground-state equilibrium phenomenon.

Our EES experiments on 6-MI-substituted DNA also revealed the presence of optical heterogeneity in these systems. Here too, the heterogeneity could be understood using the same two-state model as was used to explain the 6-MI NMP samples. We interpret these observations to mean that the optical heterogeneity of 6-MI in DNA is due to an equilibrium distribution between protonated and deprotonated species. This assumption is supported by control experiments on single-stranded and duplex DNA constructs with nucleic acid base substitutions using the fluorescent probes pyrrolocytosine (PC) and 2-aminopurine (2-AP). Neither PC or 2-AP contain an ionizable proton, and we found that single-stranded and duplex DNA constructs containing these probes as substituents behave as homogeneous one-state systems, as expected (see **Figure A2.2**). The above findings lead us to conclude that the value for the equilibrium constant  $K_{\text{eq}}$ , which we determined from our two-state analysis of EES data, is a measure of the lability

of the N3 proton of 6-MI. Our approach can thus be used to quantify the “acidity” of 6-MI within different macromolecular environments, such as DNA.

The existence of the deprotonated form of 6-MI was detected in all of the DNA constructs we investigated (see Figure 2.5, Figure A.3, Table A.2 and Table A.3). The value of the equilibrium constant of 6-MI substituted ssDNA constructs was generally less than or equal to that of the free probe in aqueous solvent at physiological pH, and varied between  $K_{\text{eq}} = 0.01$  and  $0.06$ , depending on the probe sequence context. An exception was the sequence ssGXXG, which has two guanine bases flanking the 6-MI probes. Our EES data for ssDNA constructs were relatively insensitive to variation in pH ( $=6, 7.6$ , see Figure A.4). In contrast, the value of  $K_{\text{eq}}$  for 6-MI NMP in aqueous solution was sensitive to pH ( $=0.06 \pm 0.02$  at pH 7.6,  $=0.02 \pm 0.01$  at pH 6, see Figure 2.4, Figure A.4, Table A.1 and Table A.2). Our results suggest that while the acid–base equilibrium of 6-MI NMP depends strongly on solution pH, a 6-MI probe residue embedded in ssDNA experiences a local environment that stabilizes the protonated form, and is at least partially shielded from the aqueous solvent.

We next consider how the secondary structure of ssDNA might perturb the acidity of the 6-MI N3 hydrogen. It is possible that base-stacking interactions, which play a dominant role in the stability of both single-stranded and duplex DNA, might severely restrict the accessibility of hydrogen acceptors that drive proton exchange with 6-MI. If base stacking were the predominant factor affecting proton-transfer reactivity, we might expect the equilibrium to shift toward the deprotonated state if the temperature is elevated to induce thermal unstacking of bases in ssDNA. On the contrary, we found that the value of  $K_{\text{eq}}$  was nearly equal to that of 6-MI NMP at room temperature, and this value only

marginally increased over the temperature range (20–80 °C, see Figure 2.4 D). Hence, base-stacking interactions do not appear to influence the acidity of 6-MI in ssDNA.

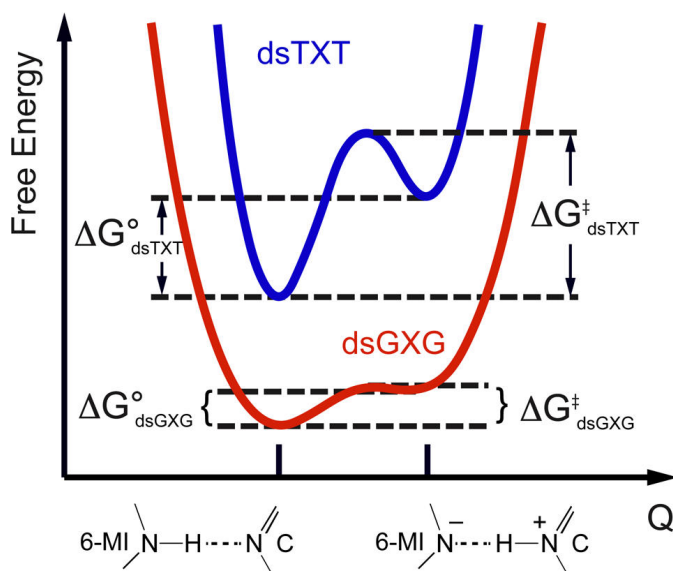
For DNA constructs in which the 6-MI probe residues were positioned adjacent to a G (or to another 6-MI), we found that the value of  $K_{\text{eq}}$  was generally larger in dsDNA than for the same sequences in ssDNA. For example, the value of  $K_{\text{eq}}$  is nearly 10-fold greater for dsGXG than for ssGXG. Exceptions to this behavior were observed for the sequences TXT and TXTXG, in which the 6-MI probe residues were positioned at sites adjacent on both sides to a T (see Figure 2.5, Tables A.2 and Table A.3). Stacking interactions between 6-MI and adjacent T residues are expected to be much weaker than the interactions between 6-MI and an adjacent G (or to another 6-MI) due to the considerable difference in surface areas between G and T. These results suggest that mechanically rigid sequences of dsDNA, in which the probe residue was positioned adjacent to a G, resulted in stabilization of the deprotonated form of 6-MI. In contrast, more flexible sequences of dsDNA, in which the probe residue was positioned adjacent to a T, did not experience a significant change in acid–base equilibrium relative to the same sequence in ssDNA.

As we previously discussed, thermal unstacking of ssGXG does not significantly alter the value of  $K_{\text{eq}}$  (see Figure 2.4 C). However, the enhanced value of  $K_{\text{eq}}$  in dsGXG does drop precipitously at elevated temperatures due to thermal denaturation, suggesting that the enhanced stability of the deprotonated form of 6-MI in duplex DNA is due to the cooperatively stabilized duplex structure. We note that the deprotonated form of 6-MI is generally more favored in duplex DNA than it is for the 6-MI NMP in solution, while the protonated form of 6-MI is more favored in ssDNA. Finally, the secondary structure associated with the duplex DNA environment is expected to decrease the accessibility of

the 6-MI probe to solvent water relative to the 6-MI NMP and the 6-MI-substituted ssDNA constructs. Taken together, these observations suggest that the enhanced stability of the deprotonated form of 6-MI in duplex DNA is likely associated with the Watson–Crick hydrogen bond between the 6-MI N3 hydrogen and its complementary C on the opposite DNA strand. Unlike the hydrogen bonds between the canonical DNA bases G and C, in which the proton remains associated with the donor base G (see Figure 1.1), the 6-MI:C hydrogen bond seems to involve significant proton transfer to the imino nitrogen of the acceptor base C. The latter situation is sometimes referred to as a low-barrier hydrogen bond (LBHB), in which the proton can be nearly equally shared between the hydrogen bond donor and acceptor sites. Such systems have been studied by determining hydrogen bond distances by X-ray crystallography, or through NMR measurements of  $^1\text{H}$ – $^{15}\text{N}$  dipolar coupling.<sup>33</sup> The possible existence of LBHBs in proteins and small molecules has been an area of active interest and debate.<sup>34–38</sup> The nuclear wave function of an LBHB in its ground vibrational level is expected to have its maximal amplitude coincident with the position of the energy barrier, so that the ionizable proton might undergo quantum mechanical tunneling between protonated and deprotonated states.<sup>34</sup> The existence of this system would represent an interesting system for future spectroscopic investigations.

The values of  $\Delta G^\circ$  and  $\Delta G^*$  that we have determined for the various DNA substrates can be used to construct projections of the free energy landscape onto the proton-transfer coordinate. Because these surfaces depend sensitively on base sequence context of the 6-MI probe, this information could be used to infer couplings between the proton-transfer coordinate and other nuclear degrees of freedom, such as base torsional motions and the disruption of Watson–Crick hydrogen bonds. As an example, we compared model free energy surfaces derived from our results for the duplex DNA constructs dsGXG and

dsTXT, which we illustrate in Figure 2.6. The forward proton-transfer reaction is thermodynamically and kinetically more favorable for the dsGXG construct ( $K_{\text{eq}} = 0.30$ ) than for the dsTXT construct ( $K_{\text{eq}} = 0.01$ , see Table A.3). In the dsGXG construct, the 6-MI probe is expected to experience a mechanically more stable local environment than in the dsTXT construct. This leads to a free energy surface for the proton-transfer reaction with nearly isoenergetic end-states bridged by a relatively low activation barrier.



**Figure 2.6:** Reaction coordinates for proton exchange dsGXG and dsTXT. The absolute positions of the curves on the  $Y$ -axis are arbitrary, and they have been offset for clarity.

We note that hydrogen bond strength is greatest when the  $\text{p}K_{\text{a}}$  values of the donor and acceptor groups are similar to each other.<sup>39</sup> The lower  $\text{p}K_{\text{a}}$  value of 6-MI NMP (=8.7, see Figure 2.6) in comparison to dGMP (=9.7<sup>40</sup>) likely contributes to the enhanced lability of the 6-MI N3 proton in duplex DNA. Furthermore, this Watson–Crick hydrogen bond involves the directional interaction between the 6-MI N3 proton and the lone pair electronic orbital of the imino nitrogen of the opposing C. The weaker acidity of 6-MI probes with adjacent T residues (in comparison to G residues, see Figure 2.5 and Figure 2.6) may be



due to the disruption of this orbital overlap resulting from the larger thermal fluctuations associated with T-rich sequences.

In this work, we have shown that the spectroscopic properties of the 6-MI probe in DNA can be used to obtain detailed information about local base conformations and conformational heterogeneity and fluctuations. Such information may be helpful in understanding detailed aspects of the mechanisms of specific and nonspecific protein–DNA interactions, including those involved in DNA replication, recombination, and repair. In work in progress, we are utilizing 6-MI substituted DNA constructs to understand the roles of thermal fluctuations of local base conformations in the assembly and functional pathways of protein–DNA complexes, and we expect that the improved spectroscopic understanding of the 6-MI probe obtained in this study will provide new insights into what “it sees” within DNA molecules.

### **Summary and Bridge to Chapter III**

In this chapter we have reported excitation–emission peak shift (EES) measurements of the fluorescent guanine (G) analogue 6-methyl Isoxanthopterin (6-MI), both as a ribonucleotide monophosphate (NMP) in solution and as a site-specific substituent for G in various DNA constructs. Changes in the peak positions of the fluorescence spectra as a function of excitation energy indicate that distinct subpopulations of conformational states exist in these samples on time scales longer than the fluorescence lifetime. Our pH-dependent measurements of the 6-MI NMP in solution show that these states can be identified as protonated and deprotonated forms of the 6-MI fluorescent probe. We implement a simple two-state model, which includes four vibrationally coupled electronic levels to estimate the free energy change, the free energy of activation, and the

equilibrium constant for the proton transfer reaction. These parameters vary in single-stranded and duplex DNA constructs, and also depend on the sequence context of flanking bases. Our results suggest that proton transfer in 6-MI-substituted DNA constructs is coupled to conformational heterogeneity of the probe base, and can be interpreted to suggest that Watson–Crick base pairing between 6-MI and its complementary cytosine in duplex DNA involves a “low-barrier-hydrogen-bond”.

We next utilize these findings in using the 6-MI probe to extract local base conformational information. In Chapter III, we have laid out the theoretical treatments of circular dichroism (CD) calculations, which connect nucleic acid base conformation to spectroscopic observables.

## CHAPTER III

### THEORETICAL ASPECTS OF CIRCULAR DICHROISM CALCULATIONS

#### **Overview**

Chapter III presents the theoretical treatments of circular dichroism (CD) calculations, which connect nucleic acid base conformation to spectroscopic observables. This chapter contains unpublished material co-authored with Neil P. Johnson, Mohammadhasan Dinpajoo, Marina G. Guenza, Peter H. von Hippel, and Andrew H. Marcus. We present the fundamentals of these CD calculations and the determination of critical spectroscopic parameters for this model. All the CD calculations presented in my dissertation were based on the content described in this chapter. The description of CD calculations of the canonical base dinucleotides includes unpublished material co-authored with Pablo G. Romano, Mohammadhasan Dinpajoo and Marina G. Guenza. The results of molecular dynamics (MD) simulations used in the CD calculations were performed by Pablo Romano in the Guenza group. The scheme to compare the results of MD simulation to CD using the matrix method was co-developed by Mohammadhasan Dinpajoo, Pablo G. Romano, Marina G. Guenza, and Andrew H. Marcus.

#### **Introduction**

Circular dichroism (CD) spectroscopy is routinely utilized to qualitatively examine the interactions between exciton-coupled molecular chromophores. This method is extremely sensitive to the presence of chiral chromophore conformations, and to protein- and ligand-induced changes of chromophore conformations. For example, the CD spectrum of base analogue dinucleotide substituted DNA constructs will change dramatically upon addition of various amounts of gene 32 protein, which is due to the

protein-induced unstacking of the analogue bases. Moreover, CD can be used to indicate the polarity preference of gene 32 binding upon binding to the ssDNA lattice.<sup>1</sup>

In order to quantitatively analyze CD spectra, we have constructed a simple theoretical model to aid in our understanding of local interactions between a fluorescent base analogue site-specifically positioned within a DNA construct and the native bases and ligands that can comprise its immediate environment.

In principle, quantitative structural information about single base-analogue-substituted DNA constructs can be obtained by comparing the results of spectroscopic measurements (e.g., CD and absorption) to the predictions of theoretical models that can be used to estimate the conformation-dependent resonant interactions between the Electric Dipole Transition Moments (or EDTMs) of nearby residues.

CD is particularly useful to understand our fluorescent Guanine analogue substituted DNA constructs due to the low-energy electronic transitions of 6-MI, which can be spectrally distinguished from the natural DNA bases. By using the fluorescent analogue bases as site-specific probes of local base conformation, it is possible to isolate the CD signals from these probes outside of the spectral window that is obscured by the natural DNA bases. We calculated CD of 6-MI substituted DNA constructs and achieved good agreement with experimental results (See Chapter IV). We then further extended our analysis to study the CD of 6-MI DNA constructs in the presence of the intercalator dye 9-amino-6-chloro methoxyacridine (ACMA) (see Chapter V).

To simulate the CD spectra of the probe-labeled DNA constructs, we applied ‘the Matrix’ methods that were originally developed for nucleic acids by Schellman<sup>2-4</sup>, Tinoco<sup>5,6</sup>, and further extended by others<sup>7</sup>. This approach, which is based on estimating the couplings between site EDTMs, has been shown to produce results that agree with

experiment for DNA sequences that favor the B form at wavelengths greater than 220 nm<sup>6,8</sup>. In this chapter, we outline the approach as we have implemented it.

### **Approximations and Principles of Theoretical Model**

In our calculations, chromophores including natural bases and fluorescent probes are treated as a group of spectroscopically well-defined individual sub-units. A DNA polymer can be regarded as an ensemble of purines and pyrimidines, which are treated as such sub-units. We assume that such sub-unit are only slightly perturbed by other sub-units and the environment. Note that these sub-units do not necessarily belong to the same molecule.

The linear rearrangement of electrons on a chromophore induced by the electric field is characterized by the electric dipole transition moments (EDTMs), denoted by  $\mu$ . The circular rearrangement of electrons of a chromophore induced by the magnetic field is characterized by the magnetic dipole transition moments (MDTMs), denoted by  $m$ . Electronic transitions occur when either (or both) electric or magnetic field radiation displaces electrons to a new state. We approximate the electronic transitions on each sub-unit as a set of point dipoles originating from the chromophore centers.

Is the point dipole approximation (PDA) adequate to describe the electronic couplings between nucleic acid base sites within a DNA molecule? The accuracy of the PDA is expected to improve when the inter-site separation  $\langle R_{ij} \rangle$  is greater than: (i) the ‘size’ of the molecule; and (ii) the transition dipole radius  $|\mu_i|/e$ , where the fundamental charge unit  $e = 1.6 \times 10^{-19}$  C. For a pair of stacked bases in dsDNA,  $\langle R_{ij} \rangle \sim 3.5$  Å. Although this separation is comparable to the average dimension of a nucleotide base ( $\sim 5$  Å), it is significantly greater than the dipole radii of nearly all the transitions relevant to

this work ( $< 1 \text{ \AA}$ , see Table 3.1-3.4, Table 3.6-3.7). Based on the above considerations, it is unclear how much error is introduced by using the Point Dipole Approximation. Nevertheless, the PDA represents a reasonable starting point to model resonant coupling within the 6-MI- and ACMA-labeled DNA constructs. We note that previous studies using the PDA to model oligonucleotides have been shown to be successful in a number of situations<sup>4,6,8,9</sup>.

A non-zero CD spectrum requires a helical redistribution of electronic charge density upon electric or magnetic (or both) excitation of the coupled transitions within the molecule. This effect is summarized by the Rosenfeld equation for the rotational strength

$$RS = Im\{\boldsymbol{\mu} \cdot \boldsymbol{m}\} \quad (3.1)$$

In the following, we assume that the intrinsic CD of individual chromophores is negligible, since individual chromophoric sub-unit are essentially achiral. Finite CD bands arise only because of pairwise chromophore-chromophore interactions. In an achiral molecule, the electron redistribution of a transition is always planar. Since most chromophores, we considered in the calculation are planar, we assume there's negligible intrinsic magnetic dipole moment ( $\boldsymbol{\mu} \neq \mathbf{0}$ ,  $\boldsymbol{m} = \mathbf{0}$ ). We consider solely the coupling between an electric dipole allowed transition and the magnetic dipoles induced by other electric dipole allowed transitions on neighboring chromophores. The cumulative CD is therefore the sum of all pairwise couplings.

### Electronic Coupling and Rotational Strength

We consider the probe substituted DNA construct as a collection of probe and nucleotide base sites. Each site can support a variable number of local EDTMs. The  $i$ th EDTM is associated with an optical transition from a ground electronic state  $|g_i\rangle$  to an

excited state  $|e_i\rangle$  with transition energy  $\varepsilon_i$  and frequency  $\nu_i$ . The EDTM is specified by a molecular frame orientation angle  $\theta_i$  and magnitude  $|\boldsymbol{\mu}_i|$ .

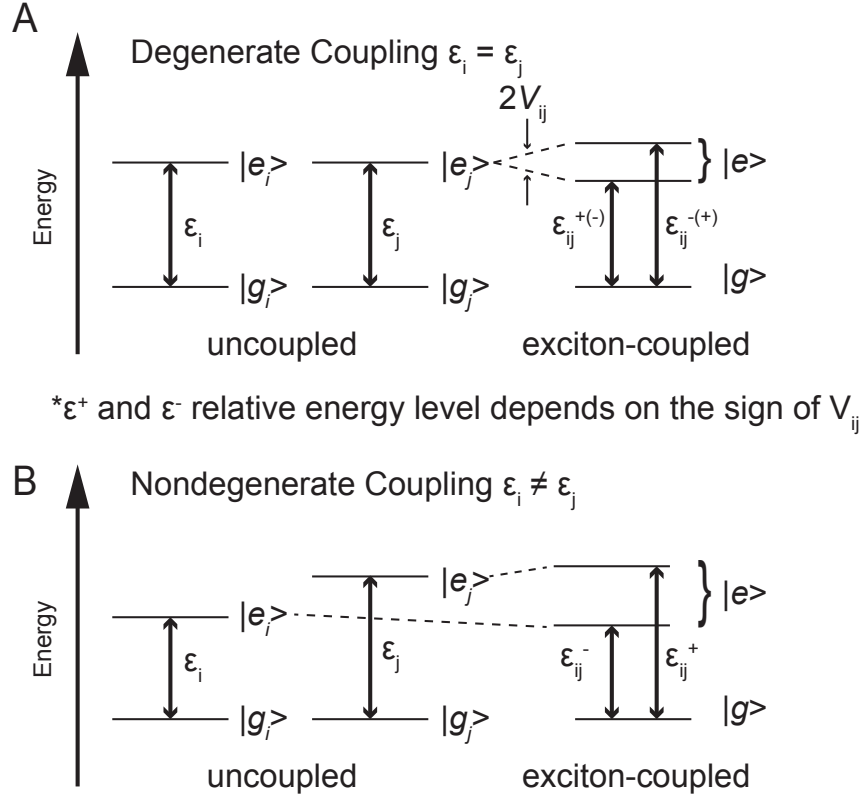
If a pair of electronic transitions are not too close to each other to deviate from the point dipole approximation, their electronic coupling is purely Coulombic given no electron exchange between them. We approximate the resonant electronic coupling between the  $i$ th and  $j$ th EDTM (assumed to be on different sites) using the point dipole-dipole expression.

$$V_{ij} = \frac{\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j - 3\hat{\mathbf{R}}_{ij} \cdot \boldsymbol{\mu}_i \boldsymbol{\mu}_j \cdot \hat{\mathbf{R}}_{ij}}{4\pi D \varepsilon_0 R_{ij}^3} \quad (3.2)$$

In Eq. (3.2),  $\mathbf{R}_{ij} = \mathbf{R}_j - \mathbf{R}_i$  is the inter-site separation vector, and the ‘hats’ indicate vectors of unit length.  $\varepsilon_0 = 8.854 \times 10^{-12} \text{ kg}^{-1} \text{ m}^{-3} \text{ s}^2 \text{ C}^2$  is the permittivity of a vacuum.  $D$  is the relative permittivity, or the dielectric constant. Note that  $V_{ij} = V_{ji}$ .

In general, the resonant coupling between the  $i$ th and  $j$ th EDTMs with energies  $\varepsilon_i$  and  $\varepsilon_j$ , respectively, gives rise to a pair of energetically shifted transitions, one with symmetric (+) and the other with anti-symmetric (−) molecular frame polarizations. Eq. (3.2) can be applied to both degenerate ( $\varepsilon_i = \varepsilon_j$ ) and non-degenerate ( $\varepsilon_i \neq \varepsilon_j$ ) types of pairwise coupling. We next consider separately the treatment of degenerate versus non-degenerate electronic coupling. The energy diagram of electronic coupling is shown in Figure 3.1.

For cases in which the  $i$ th and  $j$ th EDTMs are *degenerate* with energy  $\varepsilon_i = \varepsilon_j = \overline{\varepsilon_{ij}}$ , each exciton is assigned a rotational strength, which depends on the spatial arrangement of the individual sites according to the triple product



**Figure 3.1:** (A) Energy level diagram of degenerate coupling between two energetically identical electronic transitions. (B) Energy level diagram of nondegenerate coupling between two energetically different electronic transitions.

$$RS_{ij}^{\pm} = \pm \frac{\overline{\epsilon_{ij}}}{4\hbar} \{ \boldsymbol{\mu}_j \times \boldsymbol{\mu}_i \cdot \mathbf{R}_{ij} \} \quad (3.3)$$

These two CD bands of opposite sign occur at energies slightly above and below  $\overline{\epsilon_{ij}}$ .

$$\epsilon_{ij}^{\pm} = \overline{\epsilon_{ij}} \pm V_{ij} \quad (3.4)$$

$RS_{ij}^+$  corresponds to the transition that occurs at energy  $\epsilon_{ij}^+$  and  $RS_{ij}^-$  corresponds to the transition that occurs at energy  $\epsilon_{ij}^-$ .



Eq. (3.4) describes the special case in which the coupled two-level systems have degenerate transition energies. For cases in which the  $i$ th and  $j$ th EDTMs correspond to nondegenerate transitions with  $\varepsilon_i \neq \varepsilon_j$ , a more general solution<sup>10</sup> is

$$\varepsilon_{ij}^{\pm} = \overline{\varepsilon_{ij}} \pm \sqrt{\Delta_{ij}^2 + |V_{ij}|^2} \quad (3.5)$$

Here we have defined the average energy  $\overline{\varepsilon_{ij}} = \frac{1}{2}(\varepsilon_i + \varepsilon_j)$  and the non-degeneracy parameter  $\Delta_{ij} = \frac{1}{2}(\varepsilon_j - \varepsilon_i)$ . The preceding expressions accurately account for the effects of the resonant coupling, irrespective of the degree of degeneracy and the extent of exciton delocalization. In all cases, the effect of coupling between the  $i$ th and  $j$ th EDTMs is to increase the separation between their respective transition energies. For purposes of notation, we designate the energies of coupled pairs of degenerate transitions  $\varepsilon_{ij}^{\pm}$ , and for non-degenerate transitions  $\varepsilon_{i(j)}^{nd}$  and  $\varepsilon_{j(i)}^{nd}$ . If  $\varepsilon_i > \varepsilon_j$ ,  $\varepsilon_{i(j)}^{nd} = \varepsilon_{ij}^+$ ,  $\varepsilon_{j(i)}^{nd} = \varepsilon_{ij}^-$ . If  $\varepsilon_i < \varepsilon_j$ ,  $\varepsilon_{i(j)}^{nd} = \varepsilon_{ij}^-$ ,  $\varepsilon_{j(i)}^{nd} = \varepsilon_{ij}^+$ .

For cases in which the  $i$ th and  $j$ th EDTMs are *non-degenerate*, with respective energies  $\varepsilon_i \neq \varepsilon_j$ , the effect of the resonant coupling on the rotational strengths is described by Eq. (3.6)

$$RS_{i(j)}^{nd} = -\frac{\varepsilon_i \varepsilon_j V_{ij}}{\hbar(\varepsilon_j^2 - \varepsilon_i^2)} \{ \boldsymbol{\mu}_j \times \boldsymbol{\mu}_i \cdot \mathbf{R}_{ij} \} \quad (3.6)$$

Here  $RS_{i(j)}^{nd}$  is the rotational strength experienced by the  $i$ th EDTM due to its resonant coupling to the  $j$ th EDTM. We note that interchanging the indices  $i$  and  $j$  in Eq. (3.6) inverts the sign such that  $RS_{i(j)}^{nd} = -RS_{j(i)}^{nd}$ .  $RS_{i(j)}^{nd}$  corresponds to the transition that occurs at energy  $\varepsilon_{i(j)}^{nd}$  and  $RS_{j(i)}^{nd}$ , which corresponds to the transition that occurs at energy  $\varepsilon_{j(i)}^{nd}$ .

For each pairwise interaction, degenerate and non-degenerate coupled, there are two pairs of resulting rotational strengths with corresponding energy levels. The total CD is the result of all pairwise interactions.

The  $i$ th rotational strength is related to the  $i$ th differential absorption band with respect to left and right circularly polarized light

$$RS_i = \frac{3 \ln 10 c^2 D \epsilon_0 \hbar}{4\pi N_A} \int_{-\infty}^{\infty} \frac{[\alpha_l^i(\nu) - \alpha_r^i(\nu)]}{10\nu} d\nu = \frac{3 \ln 10 c^2 D \epsilon_0}{4\pi N_A \cdot 10} \cdot \frac{\Delta A_i}{\nu} \quad (3.7)$$

In Eq. (3.7),  $\alpha_l^i(\nu)$  and  $\alpha_r^i(\nu)$  are the absorption coefficients for left and right circularly polarized light, and  $\Delta A_i = \int_{-\infty}^{\infty} [\alpha_l^i(\nu) - \alpha_r^i(\nu)] d\nu$  is the CD band that corresponds to  $RS_i$  and  $\nu_i$ .  $N_A$  is Avogadro's number. We assume Gaussian lineshape for the line broadening

$$\alpha_l^i(\nu) - \alpha_r^i(\nu) = \frac{RS_i}{\sqrt{2\pi} \cdot \sigma_i} e^{-\frac{(\nu-\nu_i)^2}{2\sigma_i^2}} \quad (3.8)$$

where  $\sigma_i$  is the width of the band.

The total CD signal is a function of energy (frequency) resulting from the sum of individual CD bands, each with rotational strength  $R(\nu_i)$  centered at frequency  $\nu_i$

$$\alpha_l(\nu) - \alpha_r(\nu) = \sum_i \frac{RS_i}{\sqrt{2\pi} \cdot \sigma_i} e^{-\frac{(\nu-\nu_i)^2}{2\sigma_i^2}} \quad (3.9)$$

Note that the signs of the CD bands are the same as those of the corresponding rotational strengths.

We further note that according to Eq. (3.2) and (3.6), the value for  $V_{ij}$  is dependent on the dielectric constant  $D$ . For all the calculations that follow, we have assumed the value for the dielectric constant  $D = 2$ . This produces CD spectra similar in scale to those obtained from experiment<sup>8,9,11</sup>. Our assessment of a favorable comparison

between calculated and experimental CD is therefore based mainly on the ability of our procedure to reproduce the CD spectral shape.

### Determination of Electric Dipole Transition Moments (EDTMs)

The input parameters needed for the above calculations include the local EDTM orientation angles  $\theta_i$ , magnitudes  $|\boldsymbol{\mu}_i|$  (C m), and transition energies  $\varepsilon_i$  (J), in addition to the spectral standard deviations  $\sigma_i$  ( $\text{s}^{-1}$ ). The values of the orientation angles  $\theta_i$  were obtained from linear dichroism experiments performed on molecular crystals and oriented polymer films, in addition to quantum chemical calculations. We obtained values for  $|\boldsymbol{\mu}_i|$ ,  $\varepsilon_i$  and  $\sigma_i$  from spectral decomposition analyses of absorption data, which include prior experimental studies.

We determined the EDTM magnitude by integrating the experimental absorbance lineshape according to the Eq. (3.10)<sup>7</sup>.

$$|\boldsymbol{\mu}_i|^2 = \frac{3 \ln 10 c \epsilon \epsilon_0 \hbar}{\pi N_A} \int_{-\infty}^{\infty} \frac{\alpha_i(\nu)}{10 \nu} d\nu = \frac{3 \ln 10 c \epsilon \epsilon_0 \hbar}{\pi N_A} \cdot \frac{A_i}{\nu} \quad (3.10)$$

To determine the value of  $|\boldsymbol{\mu}_i|^2$  in units of  $\text{C}^2 \text{ m}^2$  using Eq. (3.2), we use the appropriate values and units for the absorption coefficient  $\alpha_i$  ( $\text{M}^{-1} \text{ cm}^{-1}$ ) as a function of the optical frequency  $\nu$  ( $\text{s}^{-1}$ ), the vacuum permittivity  $\epsilon_0 = 8.854 \times 10^{-12} \text{ kg}^{-1} \text{ m}^{-3} \text{ s}^2 \text{ C}^2$ , Planck's constant  $\hbar = 1.0546 \times 10^{-34} \text{ m}^2 \text{ kg s}^{-1}$ , the speed of light  $c = 2.99 \times 10^8 \text{ m s}^{-1}$ , and Avogadro's number  $N_A = 6.022 \times 10^{23} \text{ mol}^{-1}$ . Note that we used SI unit for all variables in our calculation except for the absorption coefficient  $\alpha_i$  ( $\text{M}^{-1} \text{ cm}^{-1}$ ).

The corresponding dimensionless oscillator strength is

$$f_i = \frac{4\pi m_e}{3\hbar e^2} v |\boldsymbol{\mu}_i|^2 = \frac{4m_e c \epsilon_0 \ln 10}{N_A e^2} \int_{-\infty}^{\infty} \frac{\alpha_i(\nu)}{10} d\nu \quad (3.11)$$

Here  $m_e = 9.109 \times 10^{-31}$  kg is the electron rest mass.

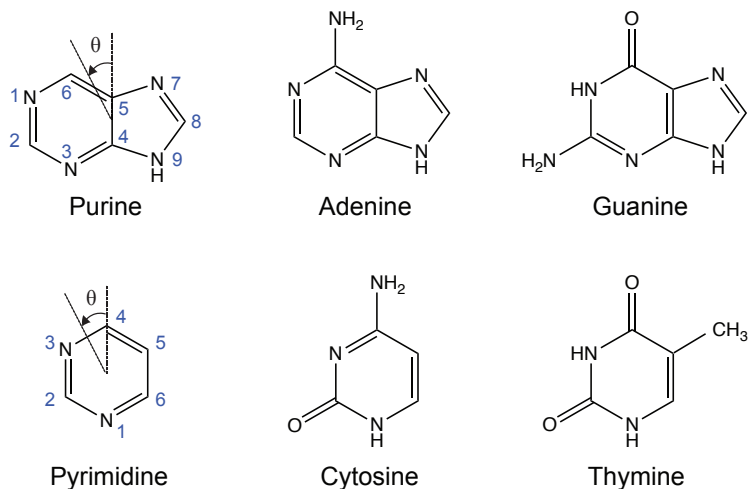
In the second equality of Eq. (3.10), we have defined the integrated area  $A_i = \int_{-\infty}^{\infty} \alpha_i(\nu) d\nu$ , which approximates the transition frequency as constant across the absorption lineshape. We modeled the experimental lineshape as a Gaussian

$$\alpha_i(\nu) = \frac{A_i}{\sqrt{2\pi} \cdot \sigma_i} e^{-\frac{(\nu - \nu_i)^2}{2\sigma_i^2}} \quad (3.12)$$

for which we have defined the standard deviation as  $\sigma_i$ .

### *EDTMs of natural bases*

In Figure. 3.2, we illustrate the definitions for the  $i$ th EDTM orientation angle  $\theta_i$  that we used in our calculations for the natural nucleic acid bases. These definitions are in accordance with the conventions established by DeVoe and Tinoco<sup>12</sup>.



**Figure 3.2:** Definitions for EDTM angles of natural orientation DNA bases.

The plane of the nucleobase is defined as the one containing atoms C4, C5 and C8 in the case of the purines (Adenine and Guanine), and atoms N1, C4 and C5 in the case of

the pyrimidines (Cytosine and Thymine). The center of the molecule is defined as the mid-point of atoms C4 – C5 in the case of purines or N1 – C4 in the case of pyrimidines. The lowest energy EDTMs lie within the plane of the nucleobases, and they are assumed to pass through the molecule center. The orientation angle  $\theta_i$  is defined as a counter-clockwise rotation relative to the above axes, as shown.

In Table 3.1-Table 3.4, we list the DNA natural bases EDTM parameters that we have used for our calculations. In all cases, we list the reference(s) from which we have obtained the relevant values. Although there is a well-established literature underlying the assignments we have made, we note that there does not appear to be uniform agreement for the correct orientation angles. We include other possible assignments in Appendix B.

**Table 3.1.** Electric dipole transition moment (EDTM) parameters used for Thymine (T).<sup>6</sup>

transition, $i$	$\varepsilon_i$ (cm <sup>-1</sup> )	$\lambda_i$ (nm)	$f_i$	$ \mu_i $ (D)	$ \mu_i /e$ (Å)	$\theta_i$ (°)	$\sigma_i$ (cm <sup>-1</sup> )
1	37 736	265	0.19	3.27	0.68	-19	2 022
2	46 512	215	0.1	2.14	0.45	71	3 146
3	51 282	195	0.1	2.04	0.43	-36	1 034
4	56 497	177	0.4	3.88	0.81	-20	1 475

**Table 3.2.** EDTM parameters used for Adenine (A).<sup>13</sup>

transition, $i$	$\varepsilon_i$ (cm <sup>-1</sup> )	$\lambda_i$ (nm)	$f_i$	$ \mu_i $ (D)	$ \mu_i /e$ (Å)	$\theta_i$ (°)	$\sigma_i$ (cm <sup>-1</sup> )
1	36 700	272	0.047	1.65	0.34	66	1 550
2	38 800	258	0.24	3.63	0.76	19	1 550
3	43 370	231	0.027	1.15	0.24	-15	700
4	46 840	213	0.14	2.52	0.52	-21	1 050
5	48 320	207	0.12	2.30	0.48	-64	1 000

**Table 3.3.** EDTM parameters used for Cytosine (C).<sup>6,14</sup>

transition, $i$	$\varepsilon_i$ (cm <sup>-1</sup> )	$\lambda_i$ (nm)	$f_i$	$ \mu_i $ (D)	$ \mu_i /e$ (Å)	$\theta_i$ (°)	$\sigma_i$ (cm <sup>-1</sup> )
1	37 594	266	0.145	2.86	0.60	25	1 592
2	43 478	230	0.0876	2.07	0.43	6	2 097
3	47 170	212	0.132	2.44	0.51	76	903
4	50 761	197	0.468	4.43	0.92	-27	3 206
5	60 606	165	0.134	2.17	0.45	0	1 550
6	66 667	150	0.187	2.44	0.51	60	1 550

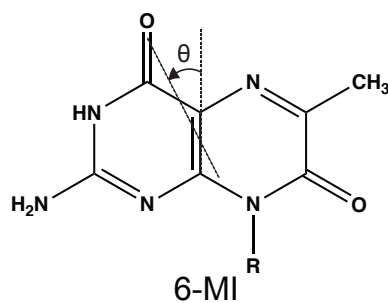
**Table 3.4.** EDTM parameters used for Guanine (G).<sup>6</sup>

transition, $i$	$\varepsilon_i$ (cm <sup>-1</sup> )	$\lambda_i$ (nm)	$f_i$	$ \mu_i $ (D)	$ \mu_i /e$ (Å)	$\theta_i$ (°)	$\sigma_i$ (cm <sup>-1</sup> )
1	36 765	272	0.132	2.76	0.57	35	1 550
2	40 323	248	0.192	3.18	0.66	-75	1 550
3	49 020	204	0.354	3.92	0.82	-75	1 550
4	53 476	187	0.386	3.92	0.82	-9	1 550
5	62 893	159	0.225	2.76	0.57	-4	1 550
6	64 935	154	0.103	1.84	0.38	-75	1 550
7	68 966	145	0.172	2.3	0.48	-4	1 550

***EDTMs of 6-MI***

The structure of the 6-MI Guanine analogue maintains the same Watson-Crick hydrogen-bonding pattern with the complementary cytosine base as does guanine. The lowest energy ( $\pi \rightarrow \pi^*$ ) electronic transitions of 6-MI are polarized within the plane of the polycyclic ring.<sup>15,16</sup>

Analogously to natural bases, we illustrate the definitions for the  $i$ th EDTM orientation angle  $\theta_i$  for 6-MI in Figure 3.3.

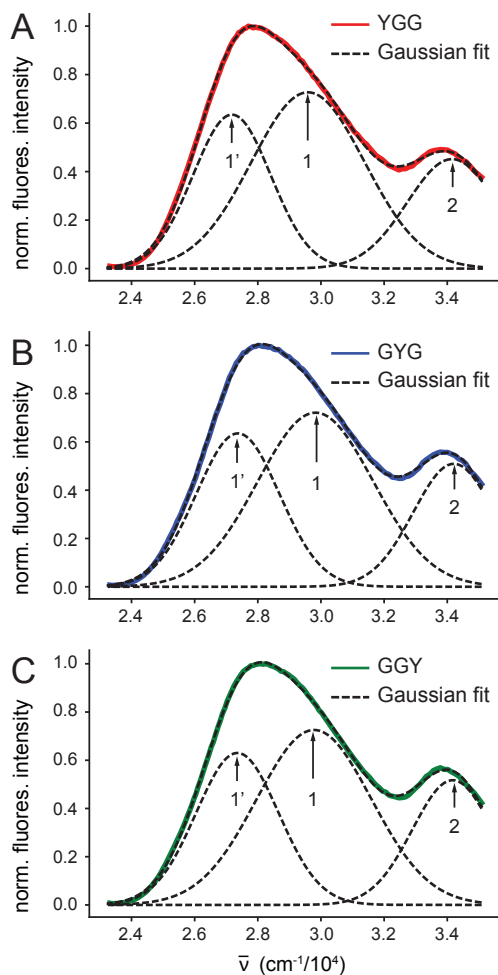


**Figure 3.3:** Definitions for EDTM angles of 6-MI Guanine fluorescent analogue

The peak absorption and fluorescence spectra of 6-MI in aqueous solution are known to exhibit inhomogeneous spectral line broadening, which depends on pH, and in 6-MI substituted DNA constructs on nucleic acid base sequence and secondary structure.<sup>17</sup> In 6-MI substituted ssDNA constructs, the 6-MI peak absorbance tends to shift to longer wavelengths in comparison to that of the 6-MI monomer in solution. This effect is more pronounced in 6-MI substituted dsDNA constructs than in ssDNA constructs.<sup>17</sup> Previous studies have shown that the observed spectral inhomogeneity is due to the presence of 6-MI conformational bi-stability corresponding to two equilibrium positions for the N3 proton, which forms one of the Watson-Crick hydrogen bonds between 6-MI and Cytosine (see Chapter II). The bi-stability of the protonated state of 6-MI gives rise to two distinct sub-populations of the lowest energy EDTM (which we refer to as transitions 1' and 1) with slightly different peak transition energies and Gaussian standard deviations.

In Figure 3.4, we show the excitation spectra of the three 6-MI substituted ssDNA constructs that we examined in the current work (YGG, GYG, GGY, see Chapter IV for base sequence details). These spectra were measured by detecting the 6-MI emission at 435 nm while varying the excitation wavelength. The excitation spectra have the advantage over absorption spectra that the underlying absorptive transitions of the 6-MI

probe can be isolated from those of the surrounding nucleic acid bases of the DNA construct. We note that the excitation spectral amplitudes vary with base sequence due to differences in the fluorescence quantum yield, while the spectral lineshape for these samples is approximately independent of base sequence.



**Figure 3.4:** Excitation spectra and Gaussian decomposition analysis of 6-MI substituted ssDNA constructs. The lowest energy features labeled 1' and 1 are due to the presence of two stable, spectrally distinct 6-MI protonation states within the nucleobase environment.

To determine the relative populations of the two protonated states of the 6-MI probe within the ssDNA constructs, we performed a Gaussian spectral decomposition analysis on each of these excitation spectra, which yielded the peak positions and line



widths of the underlying absorptive transitions. The results presented in Table 3.5 show that for these constructs the underlying 6-MI transition energies and widths are approximately independent of nucleic acid base sequence. We therefore used for our calculations the average values of the transition energy and standard deviations given in Table 3.6.

**Table 3.5.** Peak transition energies  $\varepsilon_i$  and standard deviations  $\sigma_i$  from Gaussian decomposition analysis of excitation-emission spectra for the 6-MI substituted DNA constructs studied in this work (see Chapter IV). These parameters were found to be independent of sequence context. We therefore used the averages of these values for our calculations, which are given in Table 3.6.

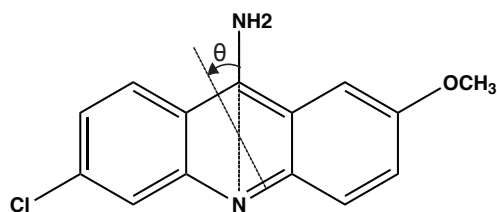
transition, $i$	YGG		GYG		GGY	
	$\varepsilon_i$ (cm <sup>-1</sup> )	$\sigma_i$ (cm <sup>-1</sup> )	$\varepsilon_i$ (cm <sup>-1</sup> )	$\sigma_i$ (cm <sup>-1</sup> )	$\varepsilon_i$ (cm <sup>-1</sup> )	$\sigma_i$ (cm <sup>-1</sup> )
1'	27 180	1 232	27 370	1 315	27 353	1 304
1	29 589	1 788	29 829	1 825	29 783	1 812
2	34 148	1 412	34 209	1 282	34 172	1 272

**Table 3.6.** EDTM parameters used for 6-methyl Isoxanthopterin (6-MI).<sup>15</sup>

transition, $i$	$\varepsilon_i$ (cm <sup>-1</sup> )	$\lambda_i$ (nm)	$f_i$	$ \mu_i $ (D)	$ \mu_i /e$ (Å)	$\theta_i$ (°)	$\sigma_i$ (cm <sup>-1</sup> )
1'	27 303	366	0.262	4.52	0.94	74.5	1 284
1	29 734	336	0.286	4.52	0.94	74.5	1 808
2	34 596	289	0.244	3.87	0.81	-25	1 452

### ***EDTMs of ACMA***

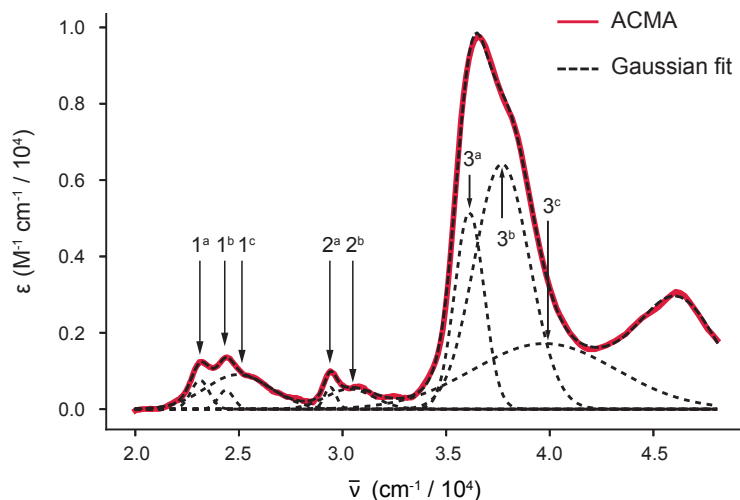
To determine the EDTM parameters for ACMA, we analyzed the absorption spectrum, which was previously reported by others.<sup>18,19</sup> We applied the reported extinction coefficient of ACMA in determining the magnitude of EDTM. ACMA EDTM angle definition is shown in Figure 3.5.



ACMA

**Figure 3.5:** Definitions for EDTM angles of ACMA.

We performed a Gaussian decomposition analysis on the ACMA absorption spectrum, as shown in Figure 3.6.



**Figure 3.6:** Absorption spectrum and Gaussian decomposition analysis of ACMA. The three lowest energy electronic transitions ( $S_0 \rightarrow S_1$ ,  $S_2$ ,  $S_3$ ) are each coupled to vibrational motions that give rise to vibronic sub-bands labeled  $1^a$ ,  $1^b$ , ...,  $2^a$ ,  $2^b$ , ..., etc.

We modeled the EDTM orientations based on the same short- and long-axis polarized assignments determined for quinacrine, for which the chromophore is chemically identical to ACMA and exhibits an absorption spectrum that is slightly shifted to lower energy. We assumed that the lowest energy electronic transition ( $S_0 \rightarrow S_1$ ) lies along the short-axis of the tricyclic ring system, and the next two electronic transitions ( $S_0 \rightarrow S_2$  and  $S_0 \rightarrow S_3$ ) lie along the long-axis of the tricyclic ring. From the ACMA

absorption spectrum, we see that each electronic transition exhibits a vibronic progression with vibrational spacing  $\sim 1,100 \text{ cm}^{-1}$ . We thus assigned transitions  $1^a - 1^c$  to the short-axis polarized  $S_0 \rightarrow S_1$  manifold, transitions  $2^a$  and  $2^b$  to the long-axis polarized  $S_0 \rightarrow S_2$  manifold, and transitions  $3^a - 3^c$  to the long-axis polarized  $S_0 \rightarrow S_3$  manifold. The resulting EDTM parameters are listed in Table 3.7.

**Table 3.7.** EDTM parameters used for 9-amino-6-chloro methoxyacridine (ACMA).

transition, $i$	$\varepsilon_i$ ( $\text{cm}^{-1}$ )	$\lambda_i$ (nm)	$f_i$	$ \mu_i $ (D)	$ \mu_i /e$ ( $\text{\AA}$ )	$\theta_i$ ( $^\circ$ )	$\sigma_i$ ( $\text{cm}^{-1}$ )
$1^a$	23 122	432	0.0037	0.58	0.12	0	439
$1^b$	24 372	410	0.002	0.42	0.09	0	371
$1^c$	24 956	401	0.0155	1.15	0.24	0	1 576
$2^a$	29 407	340	0.0018	0.36	0.08	90	285
$2^b$	30 347	330	0.0066	0.68	0.14	90	1 159
$3^a$	37 692	265	0.0929	2.29	0.48	90	1 331
$3^b$	36 118	277	0.0392	1.52	0.32	90	706
$3^c$	39 751	252	0.0647	1.86	0.39	90	3 435

### Circular Dichroism Calculation of the Dinucleotide

We applied the theoretical model described and EDTMs determined above to calculate the CD of the dinucleotides. We utilized the results of molecular dynamic (MD). CD is the net result of all conformations of dinucleotide in solution.

#### *MD simulation of the dinucleotide*

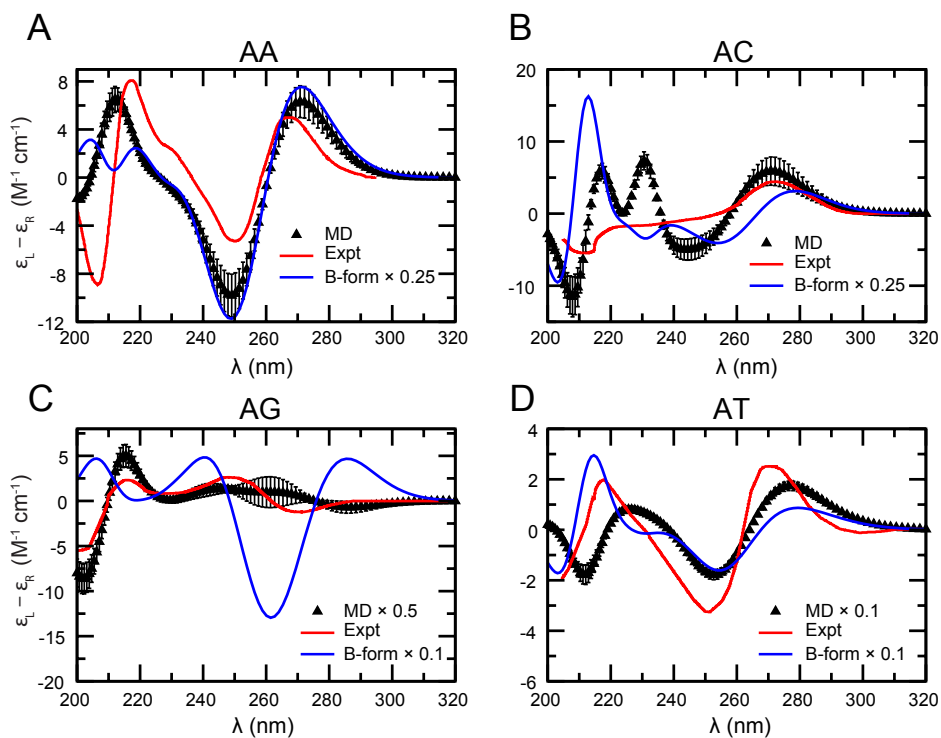
Molecular Dynamics (MD) simulations were performed using the GROMACS software program.<sup>20</sup> The simulation program “Nucleic Acid Builder” from Ambertools was used to generate the ideal B-forms of dinucleotides as the initial structures.<sup>21</sup> The dinucleotides were then solvated in a cubic box with the box length determined by the

largest distance between the atoms in the dinucleotides plus 40 Å, i.e. the relative distance between the edge of the simulation box and the dinucleotides were set to 20 Å. The Amber99+parmbsc0 force-field<sup>22</sup> was used for dinucleotides solvated in TIP3P water models.<sup>23</sup> Sufficient sodium ions were added to neutralize charges along the phosphate.<sup>24</sup> The solvated structure energy was then minimized using the Steepest Descent algorithm for 5000 steps. The whole system was then heated to T=300 K and equilibrated in the canonical (NVT) ensemble for 500 ps. This was followed by a secondary 500 ps equilibration in the isothermal-isobaric (NPT) ensemble using the BDP velocity rescaling thermostat<sup>25</sup> and Parrinello-Rahman barostat.<sup>26</sup> The MD production runs were then performed in the NPT ensemble using the Verlet integrator and LINCS constraints with a timestep of 2 fs. The production runs for the AA, AT, TA, and TT dinucleotides consisted of 1 μs and the trajectories were saved every 1 ps while production periods of 0.2 μs were used for the rest of the dinucleotides, where the trajectories were saved every 0.2 ps.

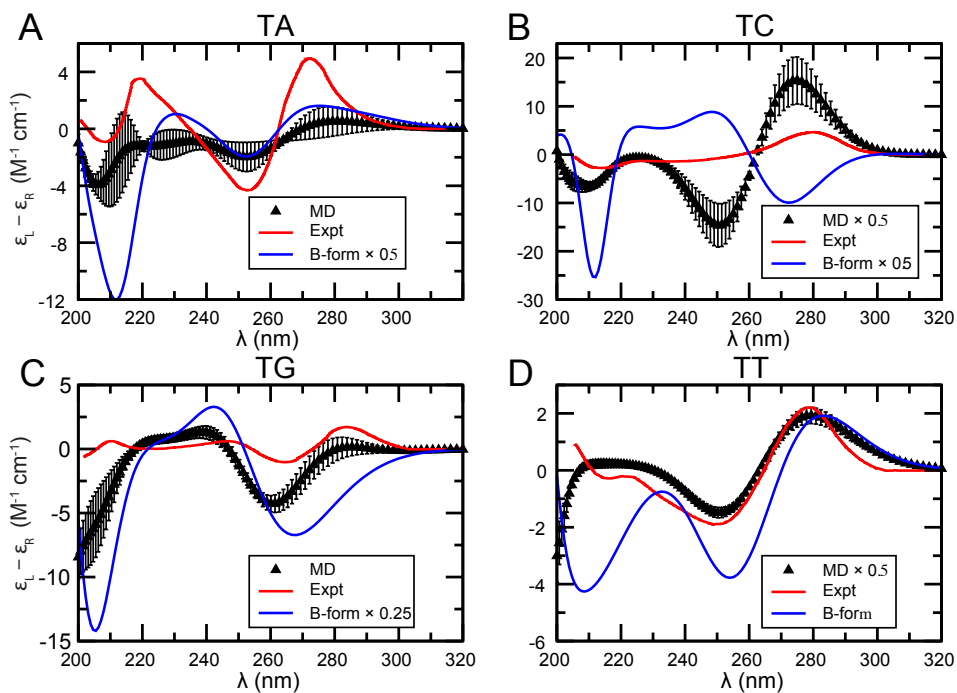
### ***CD calculation of the dinucleotide***

Overall, 1 million trajectory frames were used to obtain the average CD spectra reported for all dinucleotides. To report the statistical uncertainties in the CD spectra, 50 block averages were used to report the standard error of the mean, where each block consisted of 20,000 frames.

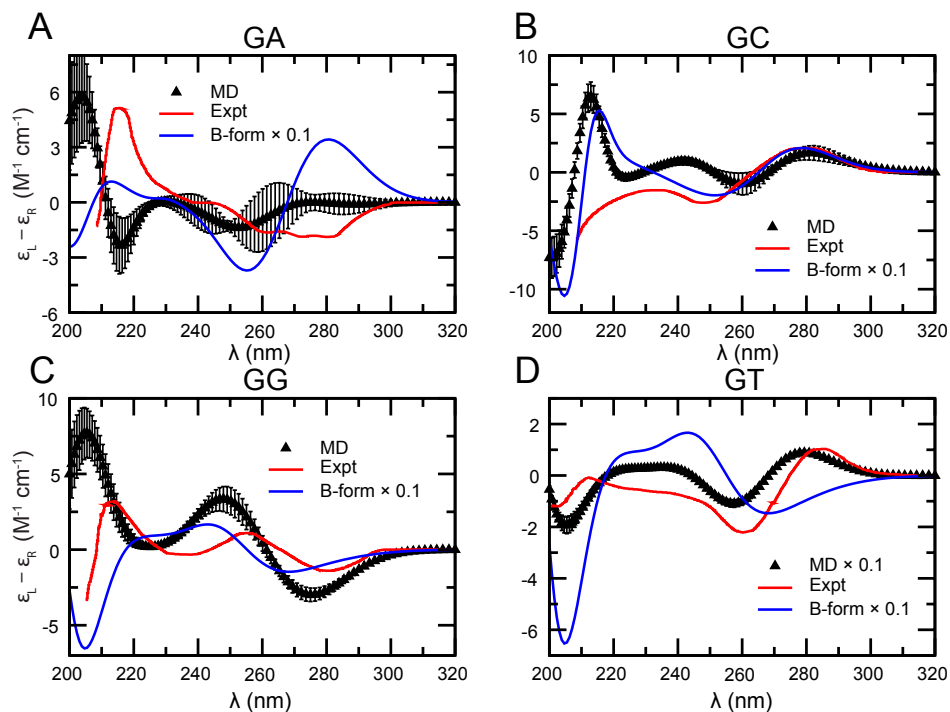
We obtained reasonable calculated CD results compared to the experimental results<sup>27</sup>. Here we show the results of all 16 dinucleotides in Figure 3.7 – Figure 3.10.



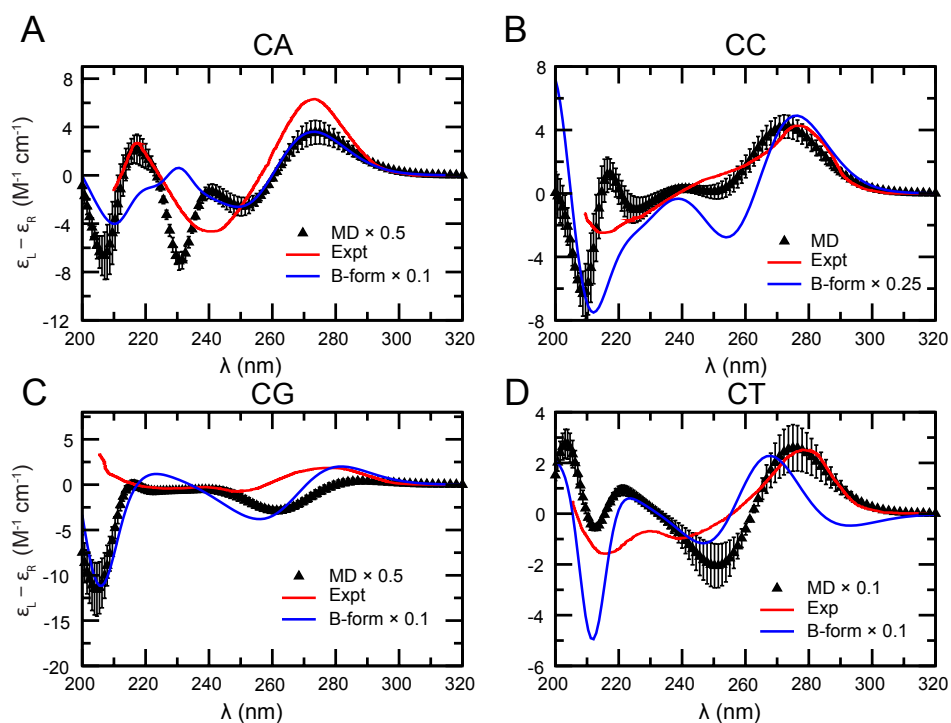
**Figure 3.7:** Circular dichroism of (A) AA, (B) AC, (C) AG, (D) AT, the polarity of each dinucleotide is from 5' to 3'.



**Figure 3.8:** Circular dichroism of (A) TA, (B) TC, (C) TG, (D) TT, the polarity of each dinucleotide is from 5' to 3'.



**Figure 3.9:** Circular dichroism of (A) GA, (B) GC, (C) GG, (D) GT, the polarity of each dinucleotide is from 5' to 3'.



**Figure 3.10:** Circular dichroism of (A) CA, (B) CC, (C) CG, (D) CT, the polarity of each dinucleotide is from 5' to 3'.

### **Bridge to Chapter IV**

In the next chapter, we will apply the CD calculation method presented in this chapter to calculate CD for 6-MI labeled DNA and DNA canonical dinucleotides.

## CHAPTER IV

# LOCAL NUCLEIC ACID BASE CONFORMATION REVEALED BY CIRCULAR DICHROISM STUDIES OF 6-METHYL ISOXANTHOPTERIN SUBSTITUTED DNA CONSTRUCTS

### Overview

Chapter IV presents the CD calculation conducted for 6-MI labeled DNA and DNA canonical dinucleotides. This chapter contains unpublished material co-authored with Neil P. Johnson, Peter H. von Hippel, and Andrew H. Marcus. Experimental CD and fluorometry were performed by Neil P. Johnson. We have found that single 6-MI substituted DNA constructs can be used to sensitively probe local nucleic acid base conformation. Qualitative agreement between calculations and experimental results reinforces the idea that the CD of 6-MI could be a useful probe of DNA secondary structure.

### Introduction

A useful strategy to study changes in the local conformations of nucleic acids is to employ fluorescent nucleic acid base analogues, which can be site-specifically positioned within a DNA construct such as a model replication fork. Base analogues such as 2-aminopurine [2-AP, an analogue of adenine (A)]<sup>1,2</sup> and 6-methyl Isoxanthopterin [6-MI, an analogue of guanine (G)]<sup>3-5</sup> absorb light at wavelengths longer than 300 nm, which is a spectral range where canonical protein and nucleic acid components are optically transparent. Because the electronic structures of such base analogue probes are typically sensitive to interactions with nearby chemical groups, site-specific spectroscopic



measurements of the optical properties of these analogues, such as absorbance, circular dichroism (CD) and fluorescence spectroscopy often depend sensitively on the details of the local DNA and protein environment.

In previous work, our groups have used excitation-emission fluorescence and CD spectroscopy to study the local conformations of DNA and protein-DNA complexes in which two adjacent (in the same strand) nucleotide residues have been substituted by ‘dimeric pairs’ of the fluorescent base analogue probe 6-methyl Isoxanthopterin (6-MI).<sup>1,4,6</sup> In those studies, the 6-MI dinucleotide served as an optical dimer probe of local base conformation, because the collective excitations (called excitons) of the 6-MI dinucleotide depends on the electronic coupling between the degenerate and non-degenerate electronic transitions of the monomer subunits. In a similar fashion, useful information about local base conformations is available from experiments in which a single base of a DNA construct is substituted with a fluorescent base analogue probe. In such cases, local interactions between the single-site-substituted analogue probe and its nearest neighbor bases affect the CD spectrum and the fluorescence quantum yield, which may be interpreted qualitatively in terms of local base stacking and unstacking conformations.

In the present study we have applied a simple theoretical model (see Chapter III) to aid our understanding of local interactions between a single 6-MI analogue probe<sup>4,6,7</sup> site-specifically positioned within a DNA construct that comprises its immediate environment. Specifically, we have studied the CD of 6-MI-substituted ss and dsDNA constructs in three different sequence contexts, and find that our model can be used to extract base-sequence-dependent information about the local conformation of the 6-MI probe as modulated by the local base or base-pair environment.

Our results indicate that information about local base conformation – in both single and duplex DNA strands – can be obtained through consideration of the non-degenerate couplings between the EDTMs of the 6-MI probe and those of adjacent canonical DNA bases.

## Materials and Methods

### Experimental procedures

#### *Preparation of DNA constructs labeled with 6-MI*

Unlabeled and 6-MI substituted DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA) and from Fidelity Systems (Gaithersburg, MD). The sequences and nomenclature of the principle DNA constructs used in this work are shown in Table 4.1. Solutions of oligonucleotides were prepared in an aqueous buffer solution containing 0.1 M NaCl and 10 mM NaPO<sub>3</sub> at pH 7.0, which were adjusted to 35 mM concentrations of the single strands using extinction coefficients provided by the manufacturer. Duplex DNA constructs were prepared by combining the single strand oligonucleotide solutions at equal concentrations and annealed by heating to 90 °C and cooling overnight in an insulated copper block. All experiments were performed at 20°C unless otherwise specified.

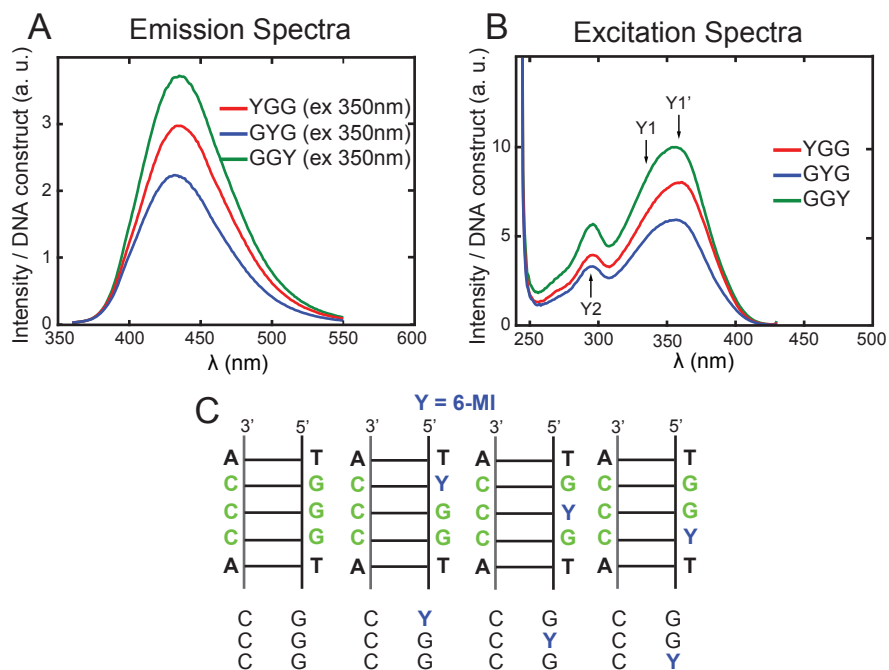
**Table 4.1.** Base sequences and nomenclature for the 6-MI-containing ssDNA constructs used in Chapter IV. The letter Y indicate the 6-MI probe. 6-MI labeled DNA constructs were also designed to form complementary sequences with CCC in inverse (3'-5' and 5'-3') orientations.

DNA construct	Nucleotide base sequence
CCC	3'-C GGA CCC ACG G-5'
GGG	5'-G CGT GGG TGC C-3'
YGG	5'-G CCT YGG TGC C-3'
GYG	5'-G CCT GYG TGC C-3'
GGY	5'-G CCT GGY TGC C-3'

**Spectroscopy.** We performed absorption, CD and excitation-emission spectroscopic measurements on both ss- and dsDNA constructs labeled with 6-MI and ACMA probes in defined positions. Fluorescence excitation and emission spectra were obtained using a Jobin-Yvon FluoroMax-3 spectrophotometer. For the DNA constructs containing 6-MI (shown as green, red, and blue curves in Figure 4.1), excitation spectra were recorded by varying the excitation wavelength over the 250 – 425 nm wavelength range while monitoring at the 435 nm emission peak. We measured the emission spectra of the 6-MI DNA constructs by exciting at the 350 nm peak absorbance and monitoring the fluorescence emission over the 350 – 550 nm range. For the ACMA-containing DNA construct (black curves), we varied the excitation over the 250 – 500 nm range, while detecting at 490 nm. We recorded emission spectra for the ACMA construct by exciting at the lowest energy peak absorbance at 435 nm, and detecting emission over the 450 – 600 nm range. We performed our absorption measurements using a Cary 3E UV-Vis spectrophotometer. We performed CD measurements using a Jasco model J-720 CD spectrophotometer. For the 6-MI labeled samples, we obtained absorbance and CD spectra over the 200 – 450 nm range. For samples containing ACMA, we measured absorbance and CD over the 300 – 600 nm range. We tested for reproducibility by repeating each measurement 3 – 4 times. The CD intensity parameters (the y-axes of the plotted CD spectra) represent the difference in absorption of the sample when excited by left- and right-circularly polarized light [ $\epsilon_L - \epsilon_R$  (in  $M^{-1} \text{ cm}^{-1}$ )] per mole of 6-MI probe chromophore. Note that all of the DNA constructs that we used employed a ‘monomer probe’ – i.e., one probe chromophore per DNA construct. All spectra were corrected for background contributions determined using the buffer solution alone.

## Results

In Figure 4.1, we show the excitation and emission spectra corresponding to the 6-MI-labeled ssDNA constructs. For the 6-MI labeled DNA constructs, the shapes of the excitation and emission spectra are independent of sequence context (YGG, GYG and GGY), although the relative intensities vary due to differences in fluorescence quantum yield. The spectral lineshape of the lowest energy absorptive feature, which is centered at approximately 355 nm, is inhomogeneously broadened. This inhomogeneity was studied previously in 6-MI-substituted ss and dsDNA constructs, and is due to the presence of two energetically distinct probe-nucleotide base environments that affect the degree of protonation of the 6-MI N3 hydrogen<sup>6</sup>.

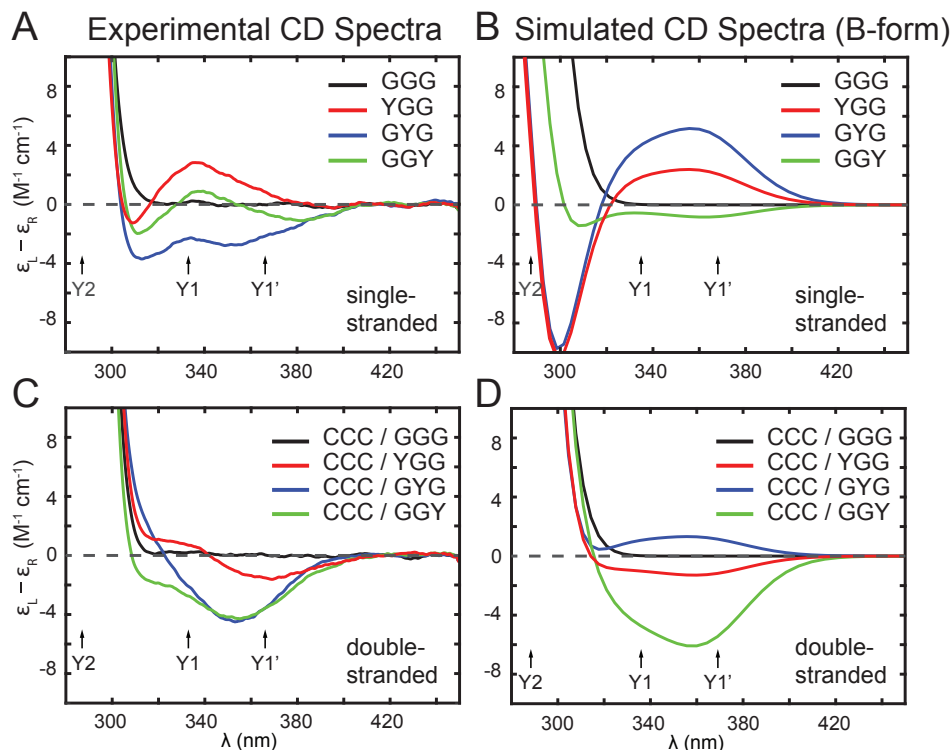


**Figure 4.1:** (A) Emission spectra of the YGG, GYG and GGY ssDNA constructs. 6-MI labeled ssDNA constructs were excited at 350 nm. (B) Excitation spectra of the YGG, GYG, and GGY ssDNA constructs. The detection wavelength for the excitation spectra was at 435 nm. Monomer electronic transitions were labeled with arrows to denote the transition energies. Y<sub>1</sub>, Y<sub>1</sub>' and Y<sub>2</sub> are transitions of 6-MI monomer. Y<sub>1</sub>, Y<sub>1</sub>' are two transitions with dipole moments that have identical orientation on 6-MI. Their transition energies depend on the local environment of 6-MI, including both the solution pH and the local DNA sequence. (C) Schematics of the central portions of the duplex DNA constructs near the spectral probes.

Prior work has shown that the two 6-MI transition energies are not correlated to differences in local base stacking conformation, but rather differences in the degree of 6-MI protonation (see Chapter II). From our Gaussian decomposition analyses of these spectra (See Figure. 3.4 and Table 3.5), we obtained EDTM parameters for the lowest energy transitions ( $S_0 \rightarrow S_1$ , centered at 366 nm,  $S_0 \rightarrow S_1$  centered at 336 nm, and  $S_0 \rightarrow S_2$  centered at 289 nm, see Table 3.6). Because the spectral widths and transition energies are experimentally indistinguishable for each of the three 6-MI-substituted DNA constructs, we used the average parameter values listed in Table 3.5 for our CD calculations. The EDTM orientations for these transitions were determined in previous work<sup>5</sup>. We assumed that the EDTM orientation for transition  $S_0 \rightarrow S_1$ , and  $S_0 \rightarrow S_1$  are identical.

We studied the sequence-dependence of the local base stacking conformations of 6-MI substituted ssDNA constructs by examining the CD of the three sequences: YGG, GYG and GGY. Each of these ssDNA constructs is an 11-mer mixed-base oligonucleotide, which has little tendency to form loops or other secondary structural motifs. Although the three-base nomenclature is useful to delineate the different samples we have studied, it does not emphasize the presence of a flanking thymine on both the 3' and 5' sides of the three-base sequence. We also measured the CD of the corresponding duplex DNA constructs CCC/YGG, CCC/GYG, and CCC/GGY, which were formed by hybridization of the 6-MI labeled strands with their complementary CCC strand. For control measurements, we examined a sample with sequence CCC/GGG.

In Figure 4.2 A and C, we present experimental CD spectra of the ss and ds DNA constructs, respectively, over the wavelength range of the 6-MI transition energies (280 – 445 nm, 22,000 – 36,000  $\text{cm}^{-1}$ ).



**Figure 4.2:** (A) Experimental and (B) simulated CD of ssDNA constructs, which assume that the nucleotides adopt the B-form conformation. A similar comparison is made for (C) experimental and (D) simulated CD of dsDNA. In panel B, the wavelengths of the uncoupled 6-MI transitions are indicated: Y1' (366 nm), Y1 (336 nm) and Y2 (289 nm) (see Table 3.6).

For each of the ss and ds DNA constructs, the placement of the 6-MI probe relative to its various neighboring bases and associated EDTMs gives rise to distinctly different CD spectra. As expected, our control GGG and CCC/GGG samples did not exhibit measurable CD at wavelengths greater than 300 nm. Thus, for samples containing the 6-MI probe, the optical activity at these wavelengths is due predominantly to the three 6-MI EDTMs (designated Y1' at 366 nm, Y1 at 336 nm and Y2 at 289 nm) and to their couplings to native base residues whose transitions occur at shorter wavelengths. In

Appendix C, we show these experimental (and simulated, see below) CD spectra over the broader spectral range 200 – 400 nm (25,000 – 50,000  $\text{cm}^{-1}$ ), which spans the transition energies of the native bases in addition to those of the 6-MI probe. At wavelengths shorter than 280 nm ( $> 32,000 \text{ cm}^{-1}$ ), the CD exhibits a prominent quasi-conservative Cotton effect centered at 268 nm, which arises due to the transitions and couplings of the native base residues within the 11-mer, and is consistent with an overall right-handed B-form conformation<sup>8</sup>. For completeness, we show in Appendix C the same CD spectra in Figure 4.2 as a function of wavenumber.

In Figure 4.2 B and D, we present our CD calculations of the 6-MI substituted ss and ds DNA constructs, respectively, for which we assumed the B-form conformation. We tested these calculations for self-consistency by varying the number of flanking bases in proximity to the 6-MI probes (see Appendix C). We found that in order to convergent results, it was necessary to include the couplings between transitions localized to bases separated by distances as large as half a helical turn. We thus included for all of our calculations all 11 bases for the ss DNA constructs, and all 22 bases for the ds DNA constructs. By following this protocol, we were able to obtain simulated CD spectra that agreed qualitatively with experiment, as we discuss further below. For all of the 6-MI substituted ss and ds DNA constructs that we studied, we found that both the shapes and magnitudes of the experimental CD exhibited sensitivity to the sequence-specific position of the 6-MI probe, and to the secondary structure of the DNA construct. We first consider the CD of the ss DNA constructs: YGG, GYG and GGY (see Figure 4.2 A). We note that each of these exhibited a negative amplitude ‘shoulder’ at  $\sim 305 \text{ nm}$ , which arises from the opposing rotational strength contributions of the Y2 and Y1 transitions. However, the dependence of the CD spectra on the sequence-specific position of the 6-MI probe can be

understood in terms of varying rotational strength contributions of the low energy Y1 and Y1' transitions. In the case of the YGG sequence (red curve), there is mainly a broad positive feature at  $\sim 335$  nm, which is due to overlapping positive rotational strength contributions of both the Y1 and Y1' transitions. On the other hand, the GGY sequence (green curve) exhibited a relatively weak CD spectrum, with a slightly positive feature at  $\sim 340$  nm (Y1 transition) and a slightly negative feature at  $\sim 370$  nm (Y1' transition). Finally, the GYG sequence (blue curve) exhibited two (unequal) negative amplitude features at  $\sim 330$  nm (Y1 transition) and at  $\sim 360$  nm (Y1' transition).

The spectral features we observed in the CD of the 6-MI substituted ss DNA constructs are also present in our calculations, which assumed the B-form conformation (see Figure 4.2 B). Yet, the agreement between experimental and simulated CD varies depending on the site-substitution position of the 6-MI probe. For example, the simulated YGG spectrum (red) exhibits two overlapping low-energy features with positive amplitude, in agreement with experiment. Similarly, the simulated CD of the GGY sequence (green) exhibits two weakly negative features, which is very similar to the experimental spectrum. We note that the slightly positive feature seen in the experimental CD at  $\sim 340$  nm is qualitatively similar to the slightly negative amplitude of this same feature in the simulated spectra. On the other hand, the experimental and simulated CD spectra of the GYG sequence (blue) are qualitatively different from one another. Most notably, the overlapping low-energy features in the experimental CD are strongly negative, while these same features appear to be strongly positive in the simulated spectra.

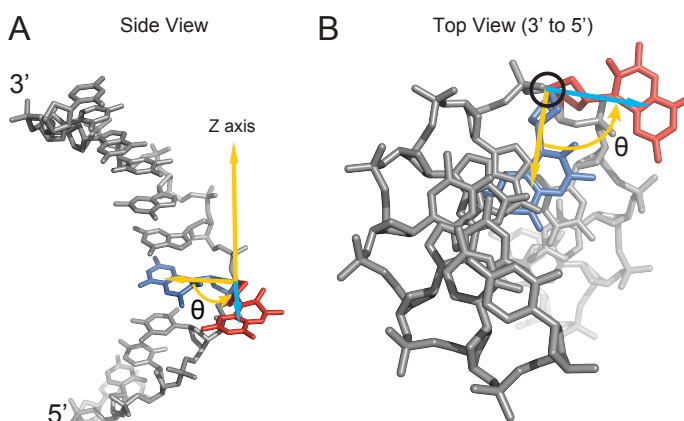
We next discuss the CD of the 6-MI substituted ds DNA constructs YGG/CCC, GYG/CCC and GGY/CCC, which are shown in Figure 4.2 C. The CD for these samples exhibited interesting similarities to those of the ss DNA constructs, and important



differences (compare to Figure 4.2 A). For each of the 6-MI substituted ds DNA constructs, the amplitudes of the low-energy Y1 and Y1' features at ~ 335 nm and ~ 370 nm, respectively, were more negative than the CD of the corresponding ss DNA constructs. This effect can also be seen in the simulated CD spectra of the YGG/CCC and GGY/CCC sequences, which assumed the B-form conformation and agree reasonably well with experiment (see Figure 4.2 D). However, similar to our results for the ss DNA GYG sequence, we obtained unfavorable agreement between experimental and simulated CD for the GYG/CCC sequence since the low energy Y1 and Y1' features exhibited negative amplitude in the experiment and positive amplitude in the simulated spectra.

It is perhaps surprising that we have obtained reasonable agreement between experimental and simulated CD spectra assuming the B-form geometry for the YGG and GGY ssDNA constructs, and for the YGG/CCC and GGY/CCC dsDNA constructs. In contrast, we found that we could not obtain, based on the B-form geometry, similarly favorable agreement between experimental and simulated CD for the GYG and GYG/CCC sequences. Without further analyses, we might interpret these results to suggest that the 6-MI probe experiences an approximately B-form local conformation within both ss and ds DNA constructs when the 6-MI probe base is positioned in between a flanking pyrimidine and a flanking purine (as is the case for the YGG and GGY sequences), but experiences a conformation that deviates from B-form when the 6-MI probe is positioned in between two flanking purines (as is the case for the GYG sequence). Such an effect could occur due to the presence of the additional atoms within the polycyclic ring of the 6-MI probe base in comparison to Guanine, which may lead to unfavorable base-stacking interactions with its two flanking purines.

To test whether our theoretical model can account for the sequence dependent CD spectra we observed in our experiments, we performed calculations as a function of the 6-MI rotation angle  $\theta$ , which is defined in Figure 4.3. This angle describes the rotation of the 6-MI probe base about the sugar-phosphate backbone (defined as the z-axis in Figure 4.3 A) relative to the B-form conformation. This rotation is affected while maintaining the atoms of the base within the same plane as for the B-form conformation. In addition, we assume that all of the remaining bases remain in the B-form geometry.

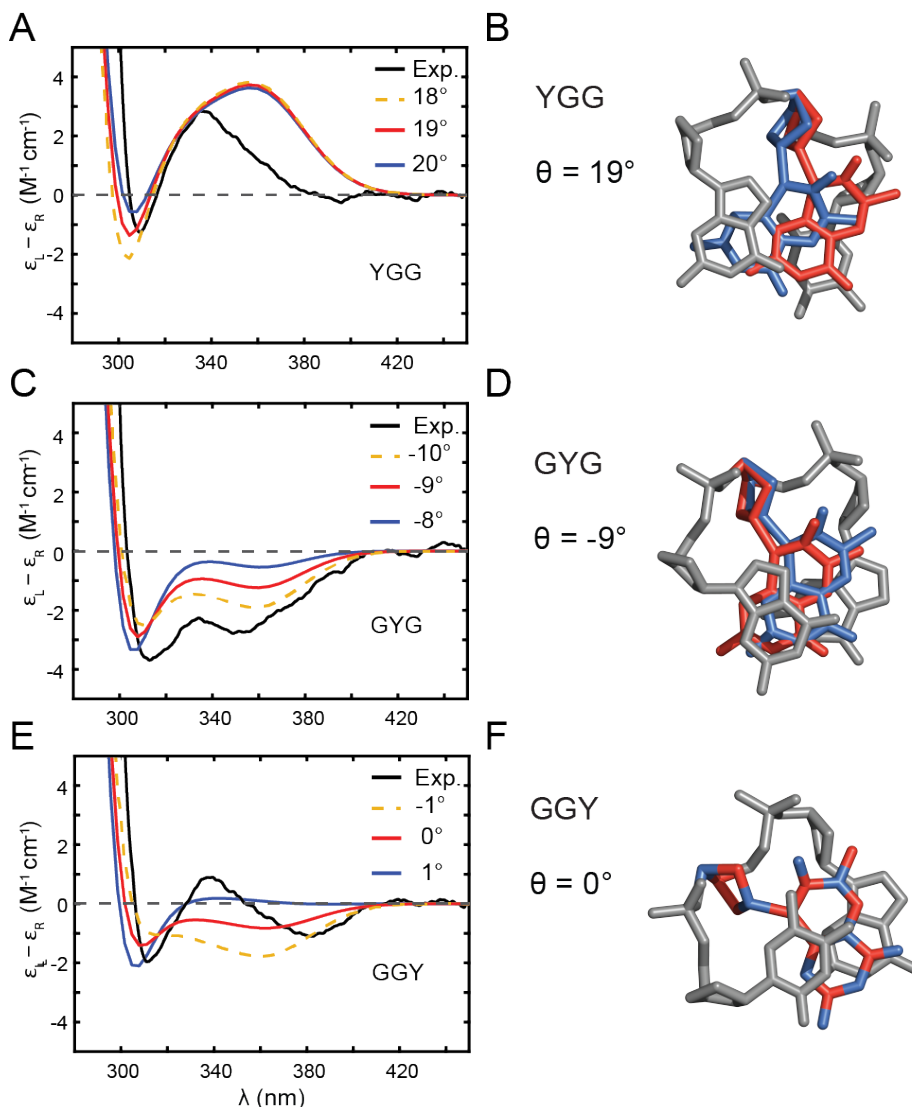


**Figure 4.3:** Molecular structural model of the ss oligonucleotide YGG with Y = 6-MI probe base (see Table 4.1 for complete sequences). The 6-MI probe base (shown in red) is rotated by the angle  $\theta$  relative to its B-form conformation (shown in blue). Views are shown from the side (**A**) and from the top (**B**).

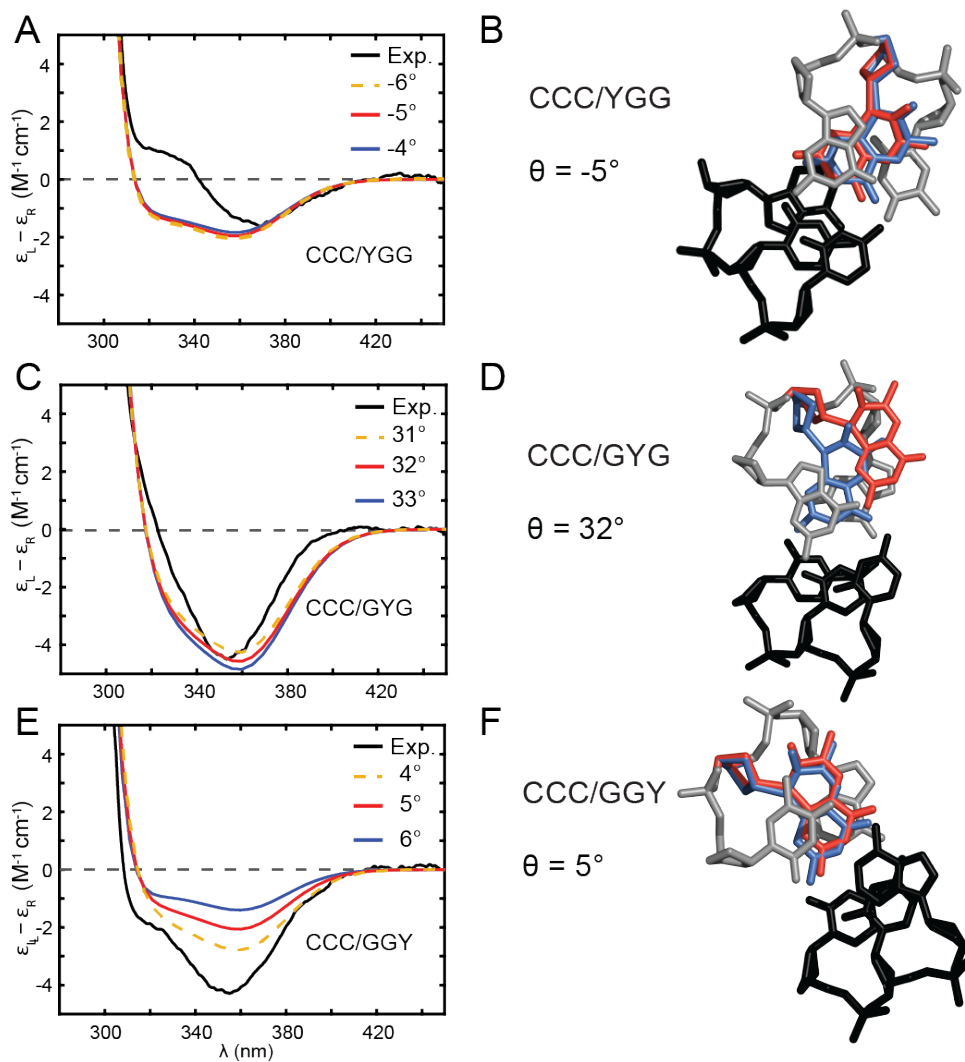
In Figure 4.4, we show the results of our calculations for each of the three 6-MI substituted ss DNA constructs in comparison to the experimental spectra. In each case, we show a family of simulated spectra to illustrate the sensitivity of these calculations to the rotation angle  $\theta$ . The dependence of these calculations over a broader range of angles is demonstrated in Appendix C.

For each of the 6-MI substituted ss DNA constructs, we were able to find an optimal 6-MI conformation that matched best with the experimental spectra. For both the GYG and GGY sequences, the simulated CD exhibited sensitivity to the rotation angles

close to the optimal angles that we obtained. Only a  $1^\circ$  difference in 6-MI conformation leads to a significant change that can invert the sign of CD. However, the simulated CD of the YGG sample did not exhibit this degree of sensitivity to the choice of the rotation angle  $\theta$ .



**Figure 4.4:** Comparison between experimental and simulated CD spectra of the 6-MI substituted ss DNA constructs (A) YGG, (C) GYG and (E) GGY. The corresponding optimized conformations are shown in panels (B), (D), (F), with the 3' end facing the top of the page. The 6-MI probe is shown in its B-form conformation in blue and in its rotated conformation in red.



**Figure 4.5:** Comparison between experimental and simulated CD spectra of the 6-MI substituted ds DNA constructs **(A)** CCC/YGG, **(C)** CCC/GYG and **(E)** CCC/GGY. The corresponding optimized conformations are shown in panels **(B)**, **(D)**, **(F)**, with the 3' end facing the top of the page. The 6-MI probe is shown in its B-form conformation in blue and in its rotated conformation in red. The opposing strand is shown in black.

From the above comparisons, we determined the angles most consistent with our experimental results. We note that in our model, any deviation of the 6-MI probe base angle from its B-form conformation corresponds to increasing its occlusion (i.e. stacking) with one of its flanking neighbor bases while decreasing its stacking with its other neighboring base. In the case of the ss YGG sample, the 19° rotation of the 6-MI probe base (shown in Figure 4.4 B) indicates that it becomes less stacked with its 3' flanking G

and more stacked with its 5' flanking T. In the case of the ss GYG sample, the  $-9^\circ$  rotation of the probe base indicates that it becomes more stacked with its 3' flanking G and less stacked with its 5' flanking G (see Figure. 4.4 E). Finally, for the GGY sample, we find that the comparison between experimental and simulated CD could not be improved by changing the 6-MI rotation angle (Figure. 4.4 F).

In the case of the ss CCC/YGG sample, the  $-5^\circ$  rotation of the 6-MI probe base (shown in Figure 4.5 B) indicates that it becomes more stacked with its 3' flanking G and less stacked with its 5' flanking T. In the case of the CCC/GYG sample, the  $32^\circ$  rotation of the probe base indicates that it becomes less stacked with its 3' flanking G and more stacked with its 5' flanking G (see Figure 4.5 E). Finally, for the CCC/GGY sample, the  $5^\circ$  rotation of the 6-MI probe base (shown in Figure 4.5 B) indicates that it becomes less stacked with its 3' flanking G and more stacked with its 5' flanking T. (Figure 4.5 F).

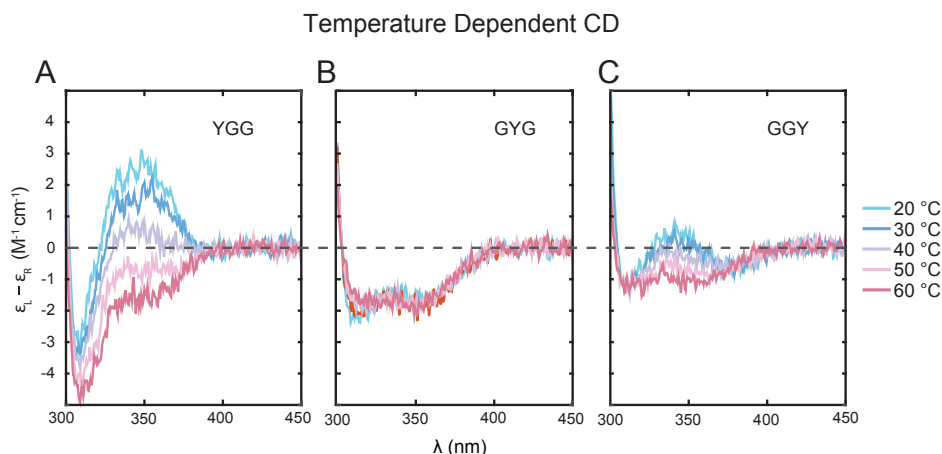
## Discussion

Our results suggest that the extent of base stacking of 6-MI probe is different among DNA constructs. It is especially interesting for the case of YGG and GGY. 6-MI in the YGG construct is flanked by 3' G and 5' T, while in GGY construct 6-MI is flanked by 3' T and 5' G. 6-MI conformation is not only dependent on the type of neighboring canonical bases, but also affected by the polarity of neighboring bases.

CD is extremely sensitive in GYG and GGY sequences around optimal results (see Figure 4.5). Only one degree difference in 6-MI conformation would cause significant change in simulated CD result and even flip the sign of CD. But in YGG, however, one-degree difference will not change CD significantly. If CD spectra of YGG, GYG and GGY reflect the secondary structure of ssDNA, then spectral shapes would be

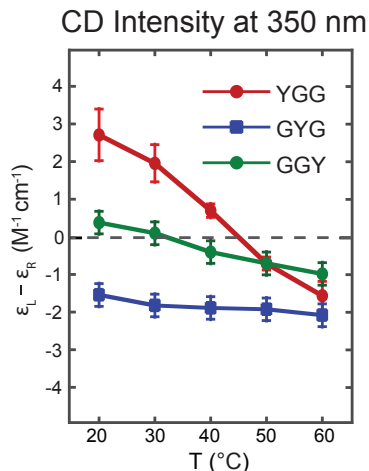
expected to change with temperature, as observed (see Figure 4.6). Interestingly the temperature dependent changes in the CD spectra also depend on the sequence context of the 6-MI probe. Larger spectral changes are observed for 6-MI residues at the edge of the GGG (YGG and GGY) tract while CD spectra of probe at the center (GYG) are nearly invariant with temperature. In Figure 4.7, we showed the CD intensity at 350 nm. The extent of CD temperature-dependence in three constructs is YGG > GGY > GYG.

The temperature dependent CD might be influenced by flanking T residues (...GTGGGTC...), which could act as flanking temperature-dependent hinges. This could explain the insensitivity of GYG to temperature compared to YGG and GGY. In other words, 6-MI stacked by Guanine is less flexible than 6-MI stacked by Thymine.



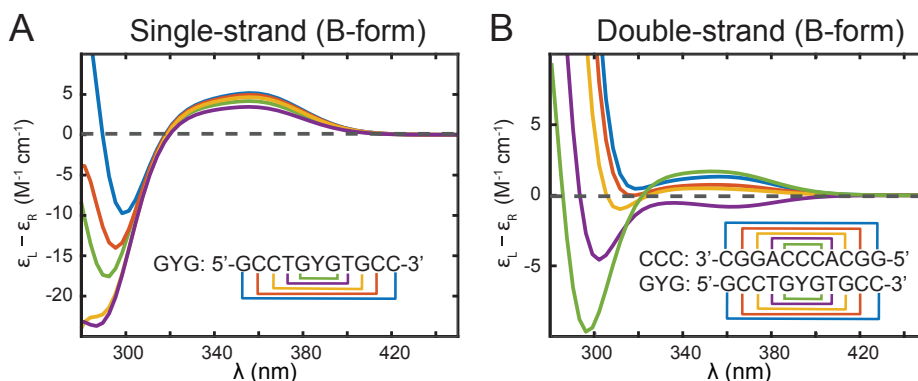
**Figure 4.6:** Temperature dependence of CD signal for the indicated single stranded construct **(A)** YGG, **(B)** GYG, **(C)** GGY. Temperature is indicated by the color code, 20 – 60 °C in 10° intervals.

However, the CD signal of the 6-MI probe of YGG and GGY respond differently to temperature. The "terminal" Guanine nucleotides are adjacent to Thymine residues with different polarity. Pyrimidine-Purines sequences in the dsDNA are more flexible than Purines-Pyrimidine sequences.<sup>9</sup> Our results might suggest that, likewise, TG sequence 5' to the GGG sequence is more flexible than the GT sequence at the 3' end although our oligonucleotides are single stranded.



**Figure 4.7:** CD intensity at 350 nm as a function of temperature from Figure 4.6.

Based on SAXS experiments, the mean number of stacked bases in  $da_{30}$  is 3 - 6 bases in 100 mM.<sup>10,11</sup> although this value could differ in our sequence, it gives an idea of the appropriate length of B form DNA for CD calculations. It was previously believed that nearest neighbor contributes most to CD signal. Yet our base-number dependent calculation shows otherwise. In Figure 4.8, we show CD calculation results when various number of 6-MI neighboring bases are included. Bases as far as half a turn away can have significant effect on final CD.



**Figure 4.8:** Calculations of the CD spectra of the GYG 6-MI substituted (A) ss and (B) ds DNA constructs (see Table 4.1). The calculations assumed B-form conformation, and were carried by including the couplings to successively larger numbers of flanking bases relative to the 6-MI probe, as indicated by the color code.

Discrepancy between calculated CD with experimental results could come from: 1. DNA can possess secondary structure in solution. Especially for ssDNA, the bases are flexible in solution. Fluctuations between different conformations are anticipated. CD signal is the average result of all stationary conformations at the given experimental condition. 2. Parameter (EDTM) uncertainty can also lead to the shift of peak position and coupling strength. We have carefully chosen the EDTM set. Nevertheless, incomplete assignment of EDTM and uncertainty of EDTM magnitude and energy contribute to discrepancy of simulated CD. 3. We have not considered non-zero intrinsic CD of any DNA bases. Although natural bases intrinsic CD are small enough to be negligible in our calculation, where strong CD peak comes from parallel interaction. 4. In our calculation, we adopt point dipole approximation, where each molecule is treated as a point. EDTM on one molecule have the same origin. However, distance and molecule are at the same order of scale. The failure of point dipole approximation may lead to complications of the coupling strength.

### **Conclusions and Future Direction**

In this chapter, we have presented the study of DNA local conformation by single 6-MI labeled DNA constructs facilitated by CD calculation. CD calculation adds to the level of sophistication in qualitatively analyzing 6-MI structural behavior in terms of local base stacking and unstacking conformations. Qualitative agreement between calculations and experimental results reinforces the idea that the CD of 6-MI could be a useful probe of DNA secondary structure. Such CD calculation can be easily applied to DNA constructs with known conformation. More sophisticated electronic coupling model,



extended dipole for instance, could be applied to improve the accuracy of CD calculated results.

### **Summary and Bridge to Chapter V**

In this chapter, we have presented the study of DNA local conformation using single 6-MI labeled DNA constructs. In the next chapter, we applied 6-MI to probe the ligand insertion of small molecules to duplex DNA.

CHAPTER V  
STUDIES OF FREE AND TETHERED INTERCALATOR MOLECULES  
INTERACTING WITH 6-MI LABELED DNA

**Overview**

Chapter V presents a study, using steady state fluorometry and circular dichroism (CD), of the interactions of the model intercalators Acridine Orange (AO) and 9-amino-6-chloromethoxyacridine (ACMA) with DNA constructs, which have been site-specifically labelled with the 6-Methyl Isoxanthopterin (6-MI) base analogue probe. This chapter contains unpublished material co-authored with Neil P. Johnson, Peter H. von Hippel and Andrew H. Marcus. Substantial exploratory experimental work involving the determination and interpretation of experimental spectra of probe labeled DNA, which was carried out by Neil P. Johnson, is included and interpreted in this chapter as well. In this work we have shown that, compared to guanine residues inserted at the same site, there is no significant preference for, or discrimination against, intercalation of AO moieties at or near the 6-MI labeled probe position. Thus, 6-MI can be used as a ‘non-discriminating’ reporter of local AO intercalation at and near guanine bases.

In this chapter we have also explored the potential of using single site-substituted 6-MI DNA constructs to measure interactions with tethered ACMA. Our results suggest that, unlike free ACMA that interacts with dsDNA by general intercalation, covalently-tethered ACMA displaces the base on the complementary strand across from the ACMA attachment site. Tethered ACMA and strategically positioned single site-substituted 6-MI DNA constructs may turn out to be an interesting model system for the study of

intercalated aromatic amino acid residues in DNA-protein interactions, and also perhaps of the enzymatic mechanisms used to ‘flip out’ bases in DNA repair processes.

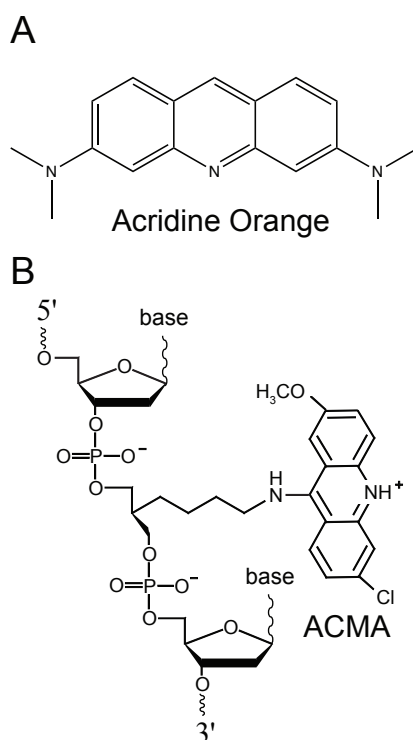
## Introduction

Intercalation occurs when a planar (aromatic) organic molecule binds between adjacent base pairs of duplex DNA, forming a sandwich-like structure<sup>1,2</sup>. The insertion of such an intercalator between two adjacent base pairs leads to significant local conformational changes, involving both the local unwinding and extension of the dsDNA structure, and can result in the perturbation of biological processes such as DNA transcription, replication and repair. Extensive studies have also focused on the development of potentially clinically useful intercalative drugs, since DNA is the primary target for many anticancer and antibiotic drugs.

Both experimental<sup>3,4,5</sup> and computational studies<sup>6,7,8,9</sup> have attempted to understand the kinetics and structural mechanisms of the insertion into duplex DNA of potential intercalating drugs. Nevertheless, it is clear that more incisive studies will be required to develop drugs with optimal base-pair sequence selectivity and reduced side effects. The ability to monitor intercalation at specific sites using well-characterized base analogue probes such as 6-MI should be helpful in moving forward with more incisive rational and novel drug design.

In prior work, the von Hippel and Marcus groups have used CD experiments performed on 6-MI dimer probes to study DNA and protein-DNA complexes<sup>10,11</sup>. Here we investigate the potential of single site-specifically-placed 6-MI probes to measure DNA local structural changes upon intercalation of a ligand near the probe base. Steady-state CD measurements and CD calculations (details presented in Chapter III and Chapter

IV) indicate that significant structure-dependent coupling occurs between the electronic states of the 6-MI probe and those of adjacent native bases (or of an adjacent intercalator probe) within a DNA construct. Hence, the CD spectrum of the low energy transitions of the 6-MI probe can be a sensitive measure of the local DNA conformation. Here we report studies of a strategically positioned single 6-MI residue in the presence and absence of a positionally-defined intercalator ligand.



**Figure 5.1:** Structure of **(A)** free intercalator Acridine Orange (AO) and **(B)** DNA backbone-tethered intercalator 9-amino-6-chloromethoxyacridine (ACMA).

Among the many different intercalating agents that have been studied<sup>12</sup>, acridine derivatives are among those that have been most thoroughly investigated as possible clinically relevant agents.<sup>13</sup> Acridine Orange (AO) is a well-studied DNA intercalator. It binds DNA via intercalation and stabilizes pigment-DNA complexes through charge neutralization of DNA backbone phosphate groups<sup>14</sup>. We have determined in this study

that 6-MI and AO (Figure 5.1 A) can serve as a Förster resonance energy transfer (FRET) donor-acceptor chromophore pair. A 6-MI molecule (FRET donor) when in its electronic excited state may transfer energy to an AO (FRET acceptor) molecule in proximity through the Förster non-radiative resonant dipole-dipole coupling mechanism. The intensity of fluorescence from the acceptor chromophore is highly dependent on its distance and orientation relative to the donor chromophore, thus making 6-MI a useful reporter of its interaction with AO at the 6-MI substitution site position. In addition, AO's fluorescence signal serves as an indicator of AO concentration and its mode of binding. In this work, we have characterized the FRET interactions between 6-MI and AO in DNA. We have then exploited this FRET donor-acceptor chromophore pair to confirm that 6-MI can be utilized as an excellent site-specific probe of the intercalation process.

We performed two types of measurements to study FRET interactions between 6-MI substituted DNA and acridine-type intercalators. In the first series of experiments we carried out titration studies of 6-MI-containing DNA constructs with AO free in solution. In a second set of experiments, we performed studies using 6-MI-labeled DNA constructs to which an acridine-type intercalator (ACMA) had been covalently attached to the DNA via a flexible linker, thus limiting the possible positions at which the ACMA probe could intercalate.

Since AO free in solution can, in principle, bind anywhere along the DNA substrate, titration studies with this ligand cannot provide defined local information about the details of the intercalation process. This uncertainty can be reduced by studying an intercalator that is site-specifically tethered to the DNA backbone. ACMA is a planar cation derivative of acridine that can be covalently attached to the DNA backbone via a

5-carbon linker (Figure 5.1 B). There is good evidence that free ACMA, like other polycyclic planar molecules that interact with duplex DNA via intercalation<sup>15</sup>, inserts itself between base pairs of dsDNA, thereby unwinding and elongating the DNA near the insertion site position.

It has been assumed that tethered ACMA will similarly intercalate into duplex DNA.<sup>16,17</sup> We attempted to characterize this putative intercalated structure by placing 6-MI residues into the dsDNA complex at various positions. Our results indicate that, unlike free ACMA, which interacts with dsDNA via intercalation, tethered ACMA displaces the base on the complementary strand across from the ACMA attachment site. This probe-labeled DNA construct may thus serve as a useful model system to study DNA-protein interactions involving the partial or total intercalation of aromatic amino acid residues, as well as the ‘base-flipping’ mechanisms in dsDNA that occur in biologically important processes such as the enzymic scanning mechanisms used to locate structural defects in DNA repair.

## **Materials and Methods**

### ***Chemicals***

6-MI-substituted DNA constructs were purchased from Integrated DNA Technologies (Coralville, IA) and from Fidelity Systems (Gaithersburg, MD). 6-MI ribonucleotide monophosphate (NMP) was obtained from Fidelity Systems and was used without further purification. ACMA-substituted DNA constructs were purchased from Glen Research Corporation. The sequences and nomenclatures of the DNA constructs we have used are shown in Table 5.1 and Table 5.2. Oligonucleotide concentrations were determined using extinction coefficients provided by the manufacturer. 6-MI NMP

concentrations were determined using extinction coefficients reported previously<sup>11</sup>. Duplex DNA constructs were formed and annealed by combining equimolar solutions of complementary strands, followed by heating the mixture to 90 °C and gradually cooling to room temperature. Solutions were buffered at pH 7.5-7.6 in 50 mM phosphate buffer containing 150 mM NaCl. All experiments were performed at 20°C. Acridine Orange hydrochloride hydrate was purchased from Sigma Aldrich and used as received (Product Number: 318337). AO stock solutions were made using the same buffer and salt conditions described above.

**Table 5.1.** Base sequences and nomenclature for the 6-MI-containing dsDNA constructs used in Chapter V. The letter **X** indicates the 6-MI probe. 6-MI-labeled DNA constructs were designed so that their sequences were complementary to one another in inverse (3'-5' and 5'-3') orientations.

DNA construct	Nucleotide base sequence
ds <b>TX</b> T	5'-CTAATCATTGT <b>X</b> TTCGGTCCTTGC-3' 3'-GATTAGTAACACAAGCCAGGAACG-5'
ds <b>TG</b> T	5'-CTAATCATTGT <b>G</b> TTCGGTCCTTGC-3' 3'-GATTAGTAACACAAGCCAGGAACG-5'
ds <b>G<b>X</b>G</b>	5'-CTAATCATTGG <b>X</b> GTCGGTCCTTGC-3' 3'-GATTAGTAACACAAGCCAGGAACG-5'
ds <b>GG</b> G	5'-CTAATCATTGG <b>G</b> GTCGGTCCTTGC-3' 3'-GATTAGTAACACAAGCCAGGAACG-5'

**Table 5.2.** Base sequences and nomenclature for the 6-MI- and ACMA-containing ssDNA constructs used in Chapter V. The letter **Y** indicate the 6-MI probe, and the letter **Z** indicates the ACMA probe. 6-MI and ACMA labeled DNA constructs were also designed to form complementary sequences with one another in inverse (3'-5' and 5'-3') orientations.

DNA construct	Nucleotide base sequence	DNA construct	Nucleotide base sequence
<b>CCC</b>	3'-C GGA <b>CCC</b> ACG G-5'	<b>YGG</b>	5'-G CCT <b>YGG</b> TGC C-3'
<b>GGG</b>	5'-G CGT <b>GGG</b> TGC C-3'	<b>GYG</b>	5'-G CCT <b>GYG</b> TGC C-3'
<b>CZC</b>	3'-C GCA <b>CZC</b> ACG G-5'	<b>GGY</b>	5'-G CCT <b>GGY</b> TGC C-3'

## *Spectroscopy*

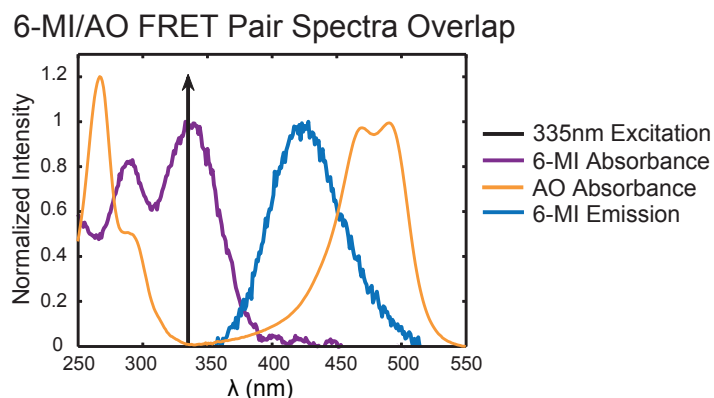
Absorption spectra were measured using a Cary 3E UV-visible spectrophotometer. Fluorescence spectra were measured in a Jobin-Yvon FluoroMax-3 spectrophotometer. All fluorescence titration spectra (shown in Figure 5.3 and Figure 5.4) were obtained using the same cuvette and the same slit setting. Circular dichroism spectra were measured using a Jasco model J-720 CD spectrophotometer with a temperature-controlled cell holder, as described previously.<sup>11</sup> CD spectra are defined in terms of the difference in extinction of the sample excited using left- and right-circularly polarized light:  $\epsilon_L - \epsilon_R$  (in  $M^{-1} \text{ cm}^{-1}$ ) per mole of 6-MI residue. All spectra were corrected for background contributions determined using the buffer solution alone.

## **Results**

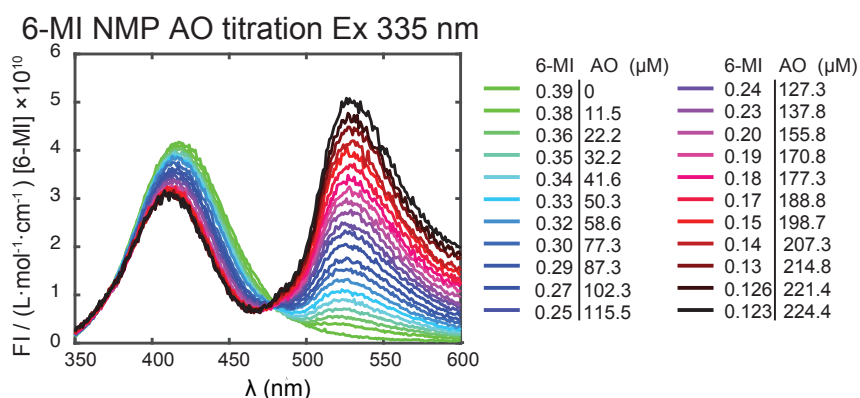
### **The 6-MI/AO FRET Pair**

6-MI and AO display significantly different absorbance spectra (Figure 5.2). At 335 nm (indicated by the black arrow in Figure 5.2), where 6-MI has a strong absorbance peak, the absorbance of AO is close to zero. Upon 335 nm excitation, the fluorescence spectrum of 6-MI significantly overlaps with the peak absorbance of AO, which spans the range 450-500 nm. Such spectral overlap of 6-MI emission with AO absorbance is a necessary condition for a FRET interaction to occur. We observed FRET between 6-MI ribonucleotide monophosphate (NMP) and AO in solution (Figure 5.3). Upon the addition of AO to the sample cuvettes and irradiation at 335 nm, fluorescence signals at 523 nm appeared while fluorescence signals at 435 nm decreased, indicating that AO was interacting with the 6-MI NMP via the FRET mechanism.





**Figure 5.2:** Spectral overlap of 6-MI NMP and AO monomer in aqueous (buffered) solution. The intensities of the absorbance spectra were normalized to the absorption peaks for easier observation of the spectral overlap. When illuminated at 335 nm, 6-MI is directly excited and resonantly transfers energy to AO, causing AO to fluoresce. A minimal fluorescence signal comes from direct AO excitation at 335 nm. The absorbance spectrum of AO was recorded using a 20  $\mu\text{M}$  sample concentration. The spectrum was normalized to its optical density (OD) at 492 nm. The 6-MI absorbance and emission spectra were also normalized by setting the spectral peak intensities to unity.












**Figure 5.3:** Fluorescence titration of AO into a solution of 6-MI NMP. (For buffer conditions see Materials and Methods.) Concentrations of 6-MI and AO monomers are indicated by color-coding (see figure). The intensities of the various spectral peaks were normalized to account for the changes in 6-MI NMP concentration due to the volume change of the solution because of AO stock solution addition.

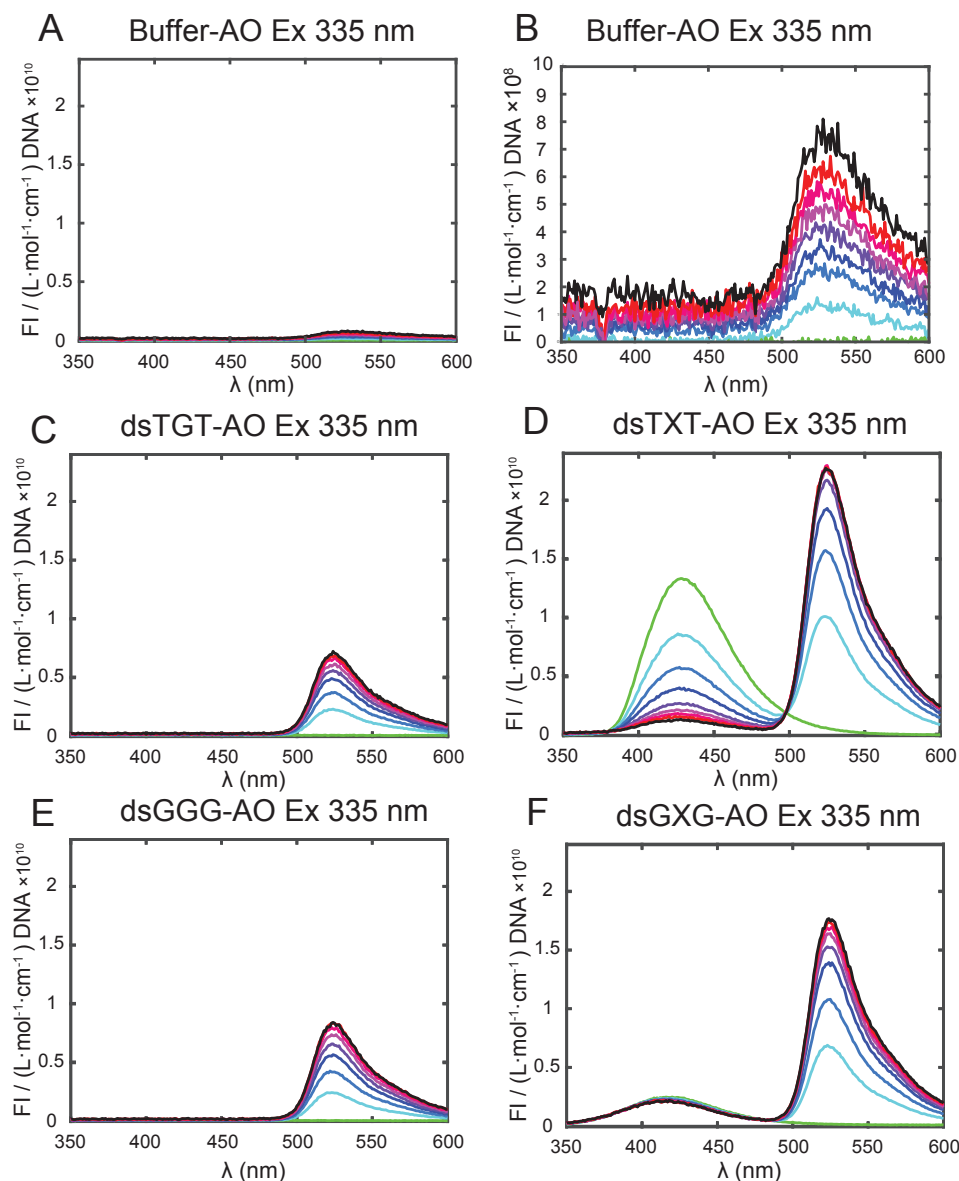
We then studied the FRET interactions between 6-MI substituted DNA and free AO in solution with further fluorescence titration experiments (Figure 5.4 and Table 5.3). Previous studies have shown that 6-MI fluorescence properties are dependent on the sequences of the bases within the oligonucleotide construct<sup>11,18</sup>. Bearing this in mind, we

conducted this experiment using two 6-MI substituted DNA constructs, which we refer to as dsTXT and dsGXG. These constructs were designed so that the flanking bases of the 6-MI probes were always either T or G. The full base sequences of these constructs are shown in Table 5.1.

Our nomenclature emphasizes that the 6-MI probe base was always flanked Ts or Gs, respectively. The remaining bases were identical between the two constructs. We also examined the fluorescence spectra of the unlabeled constructs dsTGT and dsGGG, which served as controls. We titrated AO stock solution into samples containing the dsTGT, dsGGG, dsTXT and dsGXG constructs, all dissolved in standard buffer solution. In each case, we measured the fluorescence spectra at an excitation wavelength of 335 nm, the spectra are shown in Figure 5.4. Free AO in buffer exhibited very low fluorescence (Figure 5.4 A), as expected due to its low extinction coefficient at 335 nm in the absence of DNA. Nevertheless, closer inspection showed (see Figure 5.4 B) that the free AO solution did exhibit some background fluorescence, which peaked at 523 nm and depended linearly on AO concentration.

**Table 5.3.** Concentration of Acridine Orange, dsDNA and the ratio of their concentrations in fluorescence titration experiment shown in Figure 5.3.

Spectrum Color	[AO] ( $\mu\text{M}$ )	[DNA] ( $\mu\text{M}$ )	[AO]/[DNA]
	0	4.00	0
	3.00	3.38	0.91
	5.21	2.93	1.82
	6.90	2.59	2.73
	8.22	2.32	3.64
	9.30	2.10	4.55
	10.19	1.91	5.45
	10.94	1.76	6.36
	11.57	1.63	7.27



**Figure 5.4:** Fluorescence titration experiments. **(A)** Fluorescence of buffer solution containing increasing AO concentration (control sample). The peak intensities were normalized to the AO concentrations present in the DNA-containing solutions to permit consistent comparisons of the fluorescence intensity with that of the other samples. **(B)** Same as **(A)**, but with an amplified y-axis scale. **(C)** dsTGT sample fluorescence with increasing AO concentration. **(D)** dsTXT sample fluorescence with increasing AO concentration. **(E)** dsGGG sample fluorescence with increasing AO concentration. **(F)** dsGXG sample fluorescence with increasing AO concentration. The concentrations of AO and dsDNA corresponding to each spectrum are shown in Table 5.3.

Both the dsTXT and dsGXG constructs exhibited fluorescence from the 6-MI probe, which peaked at 425 nm (green spectra in Figure 5.4 D and F) and was absent in the other samples. The dsGXG construct exhibited a much lower 6-MI fluorescence intensity than dsTXT, suggesting excited state deactivation (i.e., quenching) by the two flanking Guanine residues. Upon addition of AO to the sample cuvettes and irradiation at 335 nm, fluorescence signals at 523 nm appeared in both dsTXT and dsGXG constructs, indicating that AO was interacting with the 6-MI probes via the FRET mechanism. For the dsTXT sample, as the AO / DNA ratio was increased the fluorescence signal at ~425 nm from the 6-MI chromophore decreased, while the fluorescence at ~530 nm from the AO chromophore increased. At all concentrations, we detected the presence of a single isosbestic point at ~ 500 nm, which is consistent with the notion that the FRET signal reflects the presence of AO bound to DNA within the Förster radius defined by the 6-MI / AO interaction. We observed very similar results for the dsGXG sample. Nevertheless, the reduction of the 6-MI fluorescence was not as pronounced for the dsGXG sample as for the dsTXT sample.

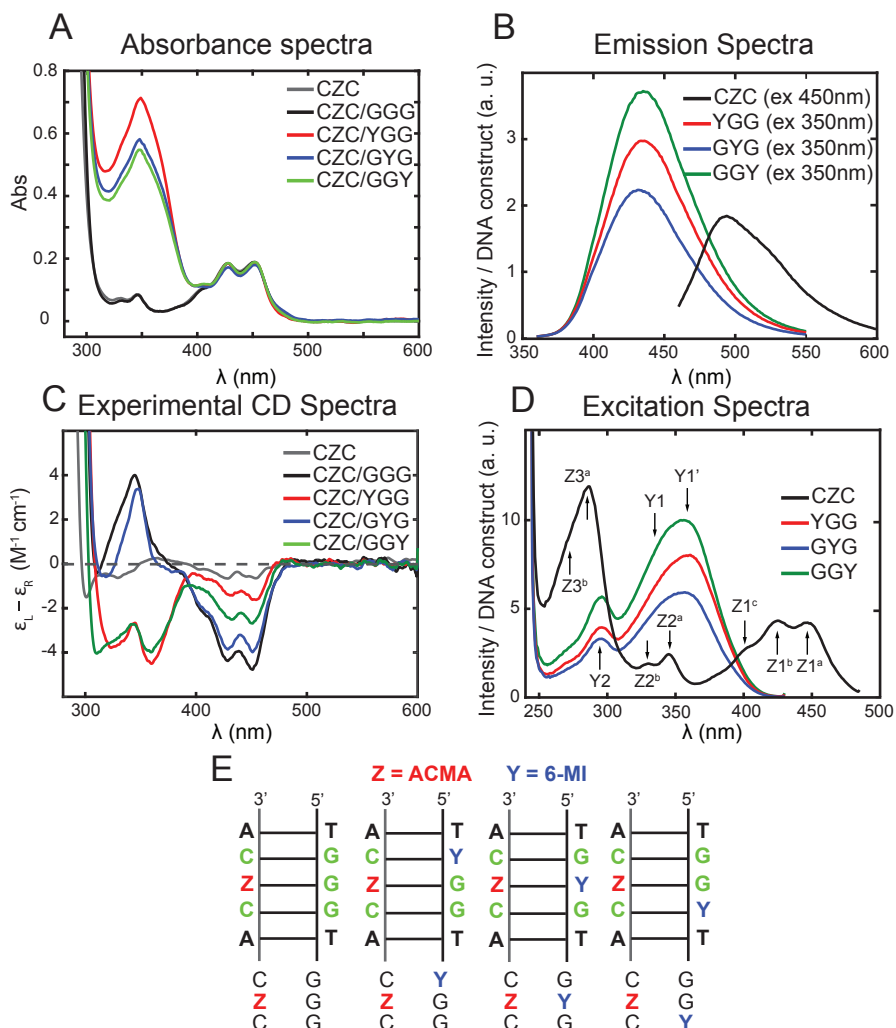
Stronger AO fluorescence was observed in the dsTXT and dsGXG samples than in the dsTGT and dsGGG samples (Figure 5.4 C and E). It is important to note that the dsTGT and dsGGG samples exhibited significantly stronger AO fluorescence than the control buffer sample that did not contain dsDNA (Figure 5.4 A). This enhanced fluorescence is due to the resulting background emission from direct excitation of AO at 335 nm of the AO-dsDNA intercalation complex. The AO emission spectrum is enhanced when AO interacts with dsDNA, serving as an indicator of intercalation binding.

## **Tethered ACMA and 6-MI Labeled DNA**

We next explored the potential use of single 6-MI labeled DNA to measure the conformation of tethered intercalator ACMA with reduced positional uncertainty relative to the DNA binding positions of free AO intercalator. We designed three 6-MI labeled ssDNA constructs in which three consecutive Guanine bases serve as controls, with each G residue substituted in turn by 6-MI in the probe-labeled constructs, (see Table 5.2).

ACMA has two low energy electronic transitions, which are observed in the absorbance spectrum near 350 nm and 450 nm (Figure 5.5). Each electronic transition in ACMA exhibits vibronic sub-bands with an approximate spacing of  $1250\text{ cm}^{-1}$ .<sup>19</sup> This feature is apparent also in the absorbance spectrum of ACMA labeled DNA. The 6-MI substituted dsDNA sequences CZC/YGG, CZC/GYG, CZC/GGY exhibited similar absorbance, emission and excitation spectra to those of non-6-MI labeled sequences (Figure 5.5 A, B, D), with insignificant energy shifts and variation of intensities due to 6-MI sequence-dependent behavior. The absorbance peak positions were invariant in the ssCZC and dsCZC/GGG structures (Figure 5.5 A). Peak electronic transition intensity at 450 nm and its associated vibronic structure, which are due to the ACMA absorption, were observed to be identical in the ssDNA and dsDNA structures, while the 350 nm band, which arises primarily from 6-MI absorption, exhibited hyperchromism in the duplex DNA complexes.

Figure 5.5 B shows fluorescence spectra of duplex DNA constructs with 6-MI at different positions across from the ACMA attachment site. Wavelengths for these experiments were chosen so that either 6-MI or ACMA were selectively probed. Interestingly, we observed significant differences between the YGG and GGY constructs, in which the 6-MI probes were both flanked by G and T, but with opposite polarity.



**Figure 5.5:** (A) Absorbance spectra of ACMA and 6-MI probe-labeled DNA constructs CZC, CZC/GGG, CZC/YGG, CZC/GYG, CZC/GGY with Z representing tethered ACMA and Y representing 6-MI. (B) Emission spectra of the CZC, YGG, GYG and GGY ssDNA constructs. CZC was excited at 450 nm. 6-MI labeled ssDNA constructs were excited at 350 nm. (C) CD spectra of CZC, CZC/GGG, CZC/YGG, CZC/GYG, and CZC/GGY. (D) Excitation spectra of CZC, YGG, GYG, and GGY. The detection wavelength for the excitation spectra was 435 nm for 6-MI labeled YGG, GYG and GGY, and 490 nm for ACMA-labeled CZC. Monomer electronic transitions were labeled with arrows to denote the transition energies. Y<sub>1</sub>, Y<sub>1</sub>' and Y<sub>2</sub> are transitions of 6-MI monomer. Y<sub>1</sub>, Y<sub>1</sub>' are two transitions with dipole moments that have identical orientation on 6-MI. Their transition energies depend on the local environment of 6-MI, including both the solution pH and the local DNA sequence. Z<sub>1</sub>, Z<sub>2</sub>, Z<sub>3</sub> are electronic transitions of ACMA. The superscripts a, b and c represent different vibronic modes within the same electronic transition. (E) Schematics of the central portions of the duplex DNA constructs in the vicinity of the spectral probes.

Circular dichroism (CD) spectra of the various duplex DNA constructs labeled on both strands are shown together with the spectrum of single-stranded CZC in Figure 5.5 C. The interaction of ACMA with ssDNA is shown to produce a significant CD signal, which has not been previously reported. The CD spectra exhibited bands from two electronic transitions at 450 nm and 350 nm, together with the vibrational progressions of these transitions (labeled in Figure 5.5 D). The planar ACMA molecule is achiral, so has no intrinsic optical activity. CD spectra of the tethered ACMA-ssDNA (CZC) molecule indicate that the chromophore experiences the asymmetric environment of ssDNA, which is consistent with the presence of stacking interactions between the chromophore and the DNA bases (Figure 5.5 C). The CD spectra of the double-stranded DNA constructs are more intense than those of the single-stranded components, both near 450 nm, where the signal is due to the ACMA alone, and near 350 nm where both ACMA and 6-MI can contribute. In addition, the negative CD signal associated with the 350 nm transition of the ssDNA construct changes sign in the duplex CZC/GGG construct.

## Discussion

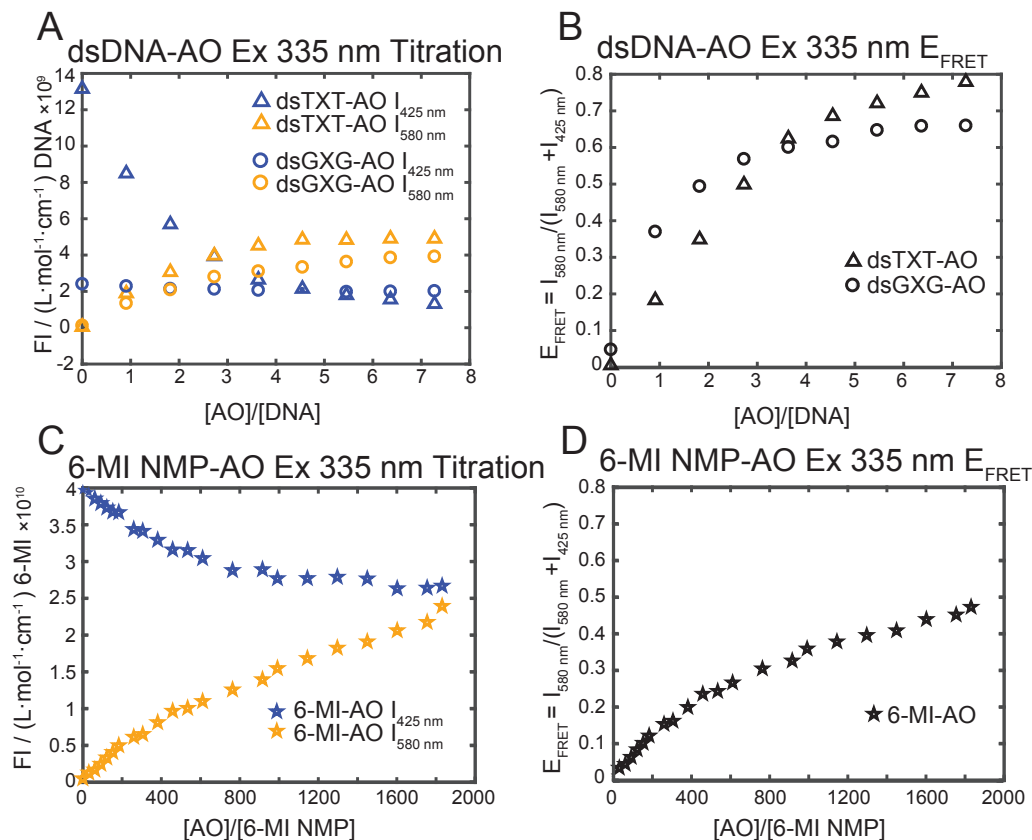
### The AO-dsDNA Intercalation Complex

To further investigate the presence of FRET between 6-MI and AO in different dsDNA samples, we examined the fluorescence intensity at two signature wavelengths, 425 nm and 580 nm. We assumed that the signal at 425 nm is representative of 6-MI fluorescence resulting from direct excitation of 6-MI at 335 nm. We assumed that the fluorescence at 580 nm originated from excited AO, resulting from FRET interactions between 6-MI and AO. Note that we did not use the peak intensity at 530 nm to define the AO excited state population. This is because 6-MI has a fluorescence ‘tail’

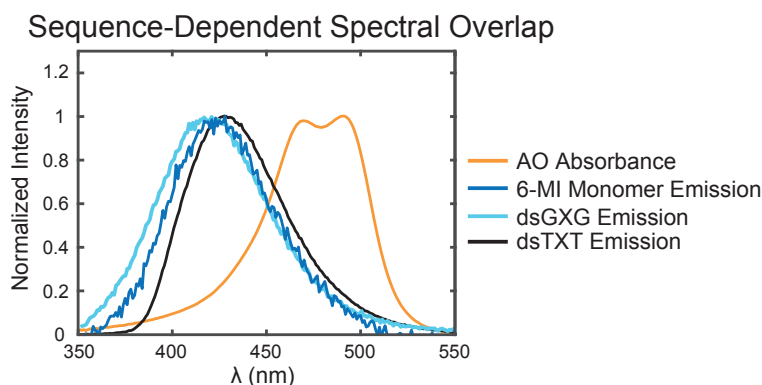
component between 500 nm and 550 nm, which could result in ‘background contamination’ of our FRET signal. We thus detect the AO fluorescence at 580 nm to eliminate the contribution of any spectral component that does not reflect the FRET interaction when evaluating AO fluorescence.

In Figure 5.6 A and B, we present the signature intensities and the associated FRET efficiency, respectively, from 6-MI and AO for both the dsTXT and dsGXG constructs upon titration with AO. We define the FRET efficiency as the ratio of intensities of the acceptor fluorescence and the total (donor + acceptor) fluorescence:  $E_{\text{FRET}} = I_{580}/(I_{425}+I_{580})$ . As the concentration ratio of AO to 6-MI-substituted DNA ( $[AO]/[DNA]$ ) was increased, we observed that the value of  $E_{\text{FRET}}$  increased for the dsTXT construct, while the value of  $E_{\text{FRET}}$  appeared to saturate at a plateau value for the dsGXG construct. In Figure 5.6 C and D, we present the signature intensities and associated  $E_{\text{FRET}}$  values from 6-MI and AO for the fluorescence titration experiment in Figure 5.3, where we examined the FRET efficiency between 6-MI NMP and AO. Despite the fact that 6-MI NMP has a much higher fluorescence quantum yield than 6-MI substituted into DNA at the same AO to 6-MI ratio, 6-MI labeled in dsDNA leads to higher FRET efficiency than 6-MI NMP in solution. This effect can be explained by the binding of AO to DNA and the inverse 6<sup>th</sup>-power dependence of the FRET efficiency on the donor-acceptor separation, which is associated with the Förster dipole-dipole coupling mechanism.<sup>20</sup> Binding of AO to dsDNA by intercalation results in much smaller distances between the 6-MI residue and AO than for unbound AO and 6-MI NMP, which are, of course, free to diffuse separately in solution.





**Figure 5.6:** Fluorescence titration experiments. **(A)** Intensities of the fluorescence peaks of the dsTXT and dsGXG constructs at the signature wavelengths 425 nm for 6-MI, and 580 nm for AO, versus  $[AO]/[DNA]$ . **(B)** dsTXT and dsGXG FRET efficiency ( $I_{580}/I_{425}+I_{580}$ ) at increasing  $[AO]/[DNA]$ . **(C)** 6-MI NMP intensity of fluorescence at signature wavelengths 425 nm for 6-MI, and 580 nm for AO, at increasing  $[AO]/[6-MI NMP]$ . **(D)** 6-MI NMP FRET efficiency at increasing  $[AO]/[6-MI NMP]$ .

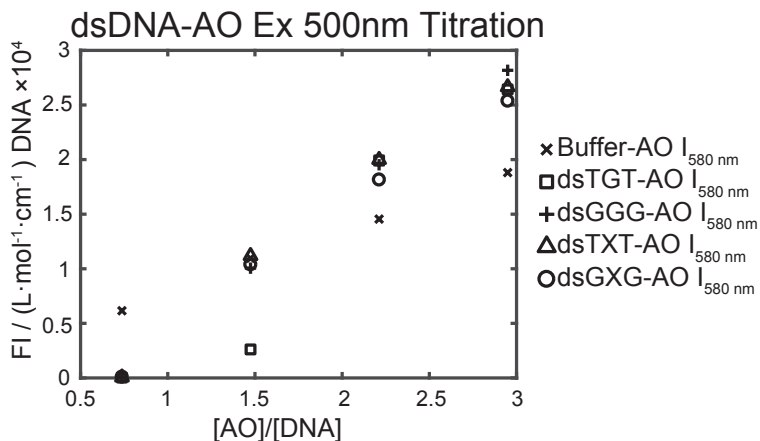


**Figure 5.7:** AO absorbance overlaps with the emission spectra of dsGXG, dsTXT and 6-MI NMP excited at 335 nm. The absorbance spectrum of AO was recorded at a 20  $\mu M$  sample concentration. The spectrum was normalized to its optical density (OD) at 492 nm. The concentrations of the dsGXG and dsTXT constructs were 4  $\mu M$ , while the concentration of the 6-MI NMP sample was 0.39  $\mu M$ . The emission spectra were normalized to the peak intensities, as previously.

Our observations of differences in FRET behavior between the two samples could be explained by the sequence-dependent excitation-emission-shift effect of 6-MI. As shown in Figure 5.7, the dsGXG and dsTXT constructs emit at different energies when excited at 335nm. The dsGXG construct emits at shorter wavelength than the dsTXT construct, thus having smaller spectral overlap with the AO absorption spectrum. Note that the 6-MI NMP peak emission wavelength falls between those of dsTXT and dsGXG. Both the dsGXG and dsTXT DNA constructs were 24 base-pairs in length, and thus could, in principle, intercalate up to 11 AO molecules. However, our observations indicate that FRET interactions between 6-MI and AO occurred at AO / DNA ratios smaller than unity. Thus, FRET interactions between AO and 6-MI were detected at anomalously low concentrations. There are two possible explanations for this observation. Either AO interacts preferentially with the 6-MI probe site within the dsDNA construct; or the Förster distance ( $R_0$ ) corresponding to the 6-MI / AO donor-acceptor pair is sufficiently large that all intercalation sites within the DNA duplex are close enough to allow for efficient energy transfer. Note that these two hypotheses are not mutually exclusive. We next tested these possibilities.

AO can be selectively excited at 500 nm. As shown in Figure 5.8, AO has higher fluorescence intensity when it interacts with any type of duplex DNA, in comparison to free AO in buffer solution. AO bound to DNA might have a higher fluorescence quantum yield due to its shielding from quenching interactions with the solvent, or from steric interactions between the AO probe and DNA that deactivate vibrational relaxation pathways. In fact, the AO fluorescence intensity is indistinguishable between 6-MI labeled DNA and canonical DNA. We thus conclude that 6-MI, in comparison to Guanine, has no significant relative preference (positive or negative) for intercalation or

near the 6-MI labeled site. This indicates that – in the context of its function as a base analogue to Guanine – 6-MI as an interaction partner with AO is likely very similar to G.

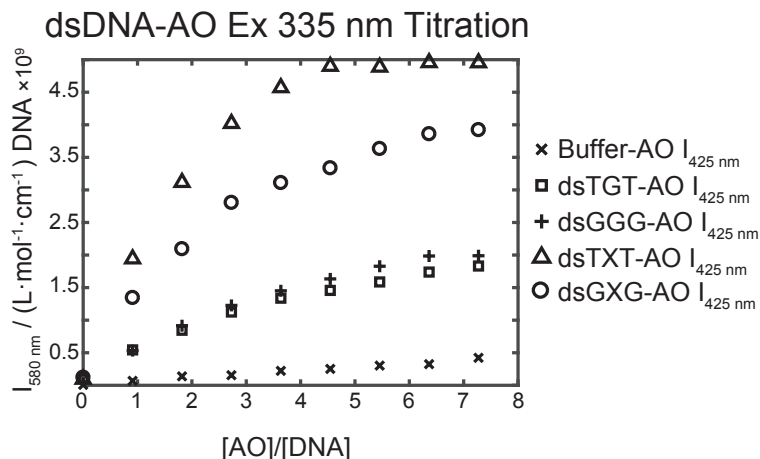


**Figure 5.8:** Fluorescence excited at 500 nm of AO in buffer solution containing dsTGT, dsGGG, dsTXT and dsGXG constructs. The fluorescence intensity of AO in buffer is normalized to be consistent with other samples containing duplex DNA at the same [AO]/[DNA] ratio.

However, since 6-MI exhibits sequence-dependent fluorescence behavior, it does have a significant effect on the AO FRET signal. As shown in Figure 5.9, when excited at 335 nm, AO in buffer solution exhibits a low fluorescence quantum yield. The fluorescence intensities of dsTGT-AO and dsGGG-AO complexes are much higher than those of uncomplexed AO. However, these intensities are significantly lower in comparison to their 6-MI-labeled counterparts dsTXT-AO and dsGXG-AO.

AO is known to form aggregates in solution at concentrations as low as 5  $\mu\text{M}$ .<sup>21,22</sup> Aggregation results in a shift of the absorbance spectrum of AO, as shown in Figure 5.10. As the AO concentration was increased, the peak absorbance at 495.8 nm decreased, while the shoulder of the absorbance at 473.4 nm increased. We fit the AO absorption spectra at different concentrations to three Gaussian peaks and assigned these,

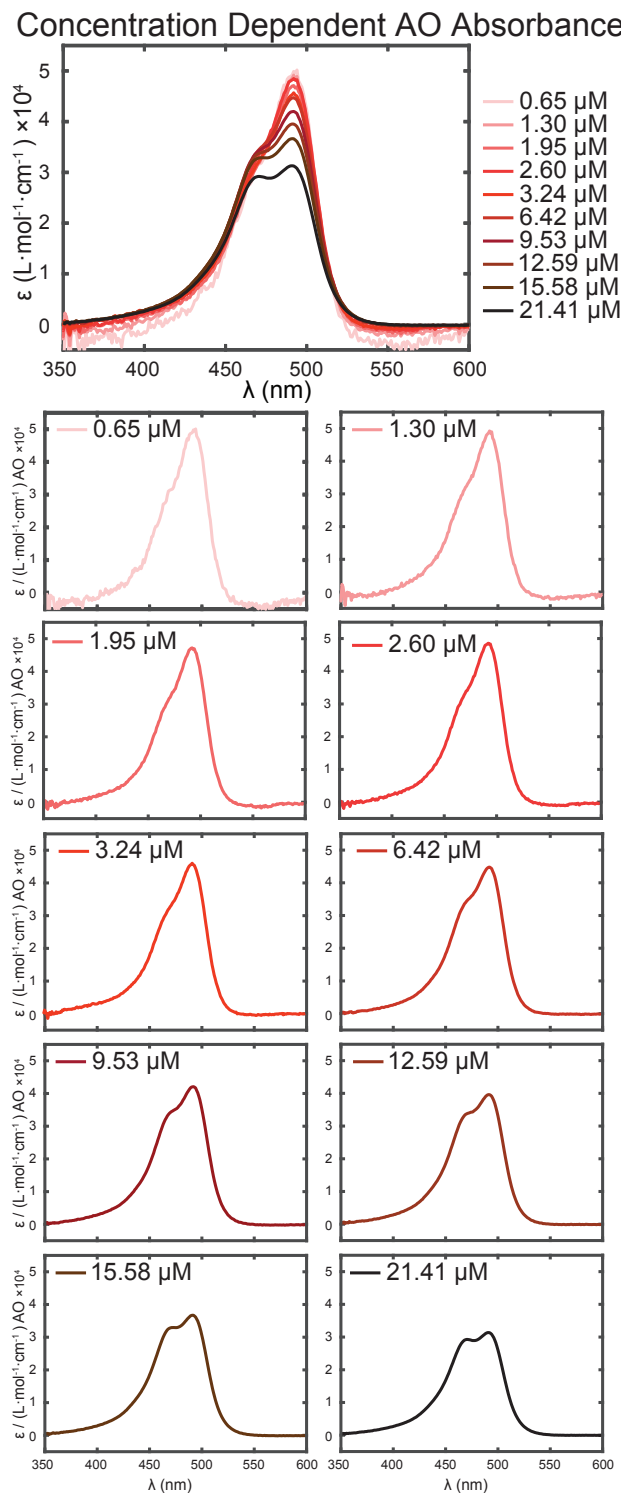
respectively, to represent the intensities of monomer, dimer, and higher aggregates of AO (for details see Appendix D).



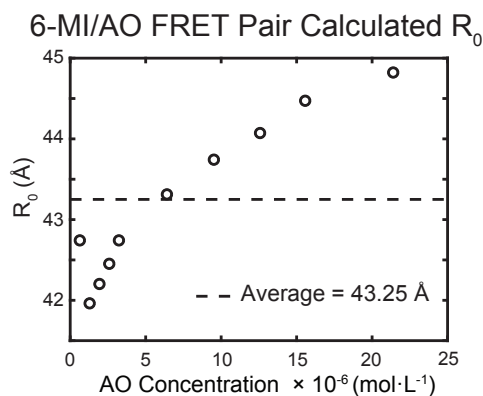
**Figure 5.9:** Fluorescence excited at 335 nm of AO in buffer containing dsTGT, dsGGG, dsTXT and dsGXG constructs. The fluorescence of AO in buffer solution is normalized to be consistent with other samples containing duplex DNA at the same [AO]/[DNA] ratio.

We calculated the Förster distance ( $R_0$ ) for this system based on the 6-MI NMP emission and AO monomer absorption spectra at different concentrations (shown in Figure 5.10). This calculation was performed using the PyMol Förster distance calculator. Details of the calculation are given in Appendix D. With AO concentrations ranging from 0.65 - 21.41  $\mu\text{M}$ , the value of the Förster distance  $R_0$  increased with AO concentration (Figure 5.11). However, the variation of  $R_0$  was small in comparison to the average value  $\langle R_0 \rangle = 43.25 \text{ \AA}$ . We note that the value of  $R_0$  for the Cy3/Cy5 FRET pair is around 50  $\text{\AA}$ . This range is significantly larger than the distance between two neighboring base pairs at 3.4  $\text{\AA}$ , but smaller than the total length of a 24-mer duplex DNA ( $\sim 78 \text{ \AA}$ ). Since 6-MI is positioned close to the middle of the two duplex DNA constructs, FRET does not provide a suitable way to distinguish between different intercalation sites of AO, based on the distance from 6-MI. Note that the change of  $E_{FRET}$  could also be due to the presence of an

antenna effect of multiple acceptors (AO) around the donor (6-MI NMP or 6-MI probe in DNA).<sup>23,24</sup>



**Figure 5.10:** The absorbance spectra of AO in buffer solution at concentrations ranging from 0.65  $\mu$ M to 21.41  $\mu$ M.



**Figure 5.11:**  $R_0$  calculated based on spectral overlap of the 6-MI NMP emission spectrum (shown in Figure 5.7) and the AO absorbance spectra in buffer solution at concentrations ranging from 0.65  $\mu\text{M}$  to 21.41  $\mu\text{M}$  (shown in Figure 5.10).

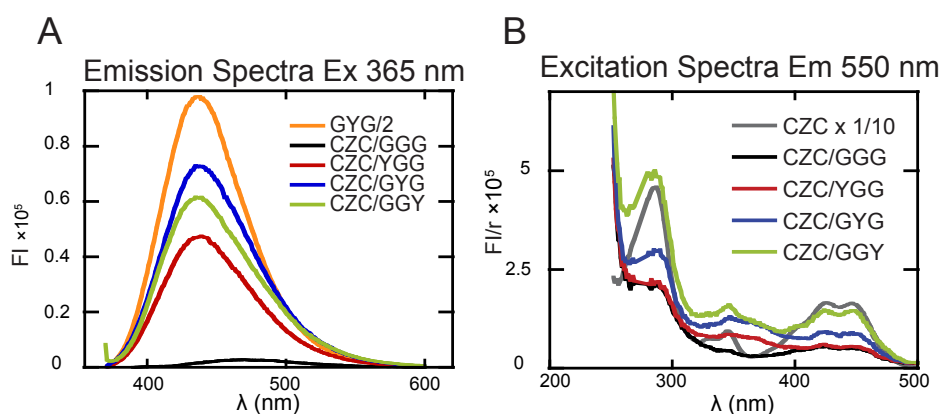
### Tethered ACMA and 6-MI labeled DNA

How might we interpret the circular dichroism spectra of the ACMA-labeled DNA constructs (Figure 5.5 D)? The increased chirality seen at 450 nm, where only ACMA absorbs, could reflect greater stacking of ACMA in the double-stranded molecule. In this scenario, the relative intensities of the 450 nm CD band would be expected to scale with stacking ( $\text{CZC/GGG} > \text{CZC/GYG} > \text{CZC/GGY} > \text{CZC/YGG}$ ). If this interpretation is correct, it would suggest that 6-MI at the 5' and 3' end of the GGG sequence inhibits insertion of tethered ACMA into the duplex, with a greater inhibition for the 5' substitution. Alternatively, non-degenerate coupling between ACMA and 6-MI could influence the CD in this region. ACMA might couple differently to the 6-MI residues in positions YGG and GGY as a result of the right-handed helical structure of the duplex DNA. This has been confirmed by calculated CD spectra (see Appendix D).

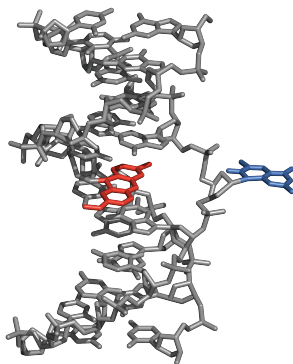
It is worth noting that the CD spectra of CZC/YGG and CZC/GGY have troughs of equal apparent intensities near 325 nm and 350 nm, thus resembling the CD pattern seen with single-stranded GYG, which suggests that this feature can be observed in the

absence of ACMA and may have its origin in the optical activity of 6-MI itself within the construct.

The CD spectra of the doubly labeled duplexes fall into two classes. First, the spectra of CZC/GGG and CZC/GYG, with a peak near 340 nm (where 6-MI and ACMA have electronic transitions) and troughs at 450 nm (signal primarily from ACMA) are nearly identical. Second, the spectra of CZC/GGY and CZC/YGG are similar in shape, with overlaying troughs at 300 – 400 nm and peaks at 450 nm.



**Figure 5.12:** (A) Emission spectra of CZC/GGG, CZC/YGG, CZC/GYG and CZC/GGY constructs excited at 365 nm. (B) Excitation spectra of CZC, CZC/GGG, CZC/YGG, CZC/GYG and CZC/GGY. The detection wavelength for the excitation spectra was 550 nm.



**Figure 5.13:** A model of the CZC/GYG construct with the 6-MI base ‘flipped out’. ACMA is the colored red, 6-MI is colored blue.

That observation that the CD spectra of CZC/GGG and GGG/GYG are almost identical is surprising. This suggests not only that the ACMA residue does not interact with 6-MI in the CZC/GYG complex, but that the 6-MI residue appears to be chirally ‘transparent’ in this construct. These results are compatible with the hypothesis that ACMA ‘pushes’ the 6-MI directly across from its attachment site out of the duplex. 6-MI has no intrinsic optical activity, and if extruded from the duplex and able to rotate freely it would have no CD signal and the CD of CZC/GYG would resemble that of the equivalent duplex DNA construct containing no 6-MI (CZC/GGG). This model (shown in Figure 5.13) is consistent with the fluorescence results. The overall fluorescence emission from 6-MI decreased by a factor of 2-3 in the duplex compared to the ssDNA; for example, the emission intensity of CZC/GYG is less than half of the GYG intensity (Figure 5.12 A). The relative fluorescence of the CZC/GYG > CZC/GGY > CZC/YGG constructs is different from the relative intensities of the equivalent ssDNA sequences, which decreases as GGY > YGG > GYG (Figure 5.5 B). The relative fluorescence intensity of GYG is the lowest of the single stranded constructs and CZC/GYG has the greatest emission intensity of the duplex constructs. This could be explained if the 6-MI residue were ‘flipped out’ of the duplex and exposed to solvent, rather than stacked with adjacent bases – an environment favoring nonradiative decay. This hypothesis has been confirmed by CD calculation, as is detailed in Appendix D.

Finally, it is worth noting that excitation at 365 nm (a wavelength at which the excitation of 6-MI is dominant and the contribution of ACMA is minimal) contributes to the fluorescence of ACMA at 550 nm (Figure 5.12 B), indicating energy transfer from 6-MI to ACMA in the double-strand DNA complex.



## Conclusions and Future Direction

Studies of the interactions between free intercalator AO with 6-MI-labeled DNA show that 6-MI can serve as a useful probe of the interactions of Guanine in studies of the mechanism(s) of intercalating ligand insertion at defined neighboring sites. Since duplex DNA, 6MI, and AO have distinct optical properties, we have been able to conduct spectroscopic experiments at different wavelengths to obtain information about the behavior of the different components involved in the intercalation complex. For example, in titration experiments we have isolated signals attributed to DNA, 6MI, AO, and the 6MI/AO complex. This information should be useful in future spectroscopic studies of such probe interactions at defined positions within dsDNA.

Additional AO titration experiments with a variety of DNA sequences are required to provide information about the stoichiometry of AO intercalation. Understanding intercalation kinetics is a difficult topic, and yet seems to be directly connected to DNA fluctuations. Both experimental<sup>3,4,5</sup> and computational studies<sup>6,7,8,9</sup> have been carried out to understand intercalation kinetics and the structural issues that guide these processes.. Significant questions pertain to the possible role(s) of DNA fluctuations during intercalation. For example, does the small ligand induce a DNA conformational change first, or do the relevant fluctuations exist and occur in the absence of the potential intercalator? It appears that 6MI and AO can serve as potentially useful candidate probe for single molecule FRET studies, which can then serve to obtain information about the kinetics of time-resolved AO intercalation into single DNA duplexes at 6MI labeled sites.

In MD modeling of the intercalation process, a base-pair ‘rolling’ mode towards the minor or major groove can be observed<sup>24-25</sup>. Such a base-pair rolling mode can also be

regarded as a DNA ‘bending’ mode. It may well be possible to monitor this bending motion using single molecule experiments. The 6-MI/AO FRET pair could be utilized as a powerful probe in single-molecule spectroscopy and CD calculations, and such experiments and calculations could be built upon our present studies. Some preliminary experiments have been performed and are illustrated in Appendix D.

In possible future smFLD (Florescence Linear Dichroism) experiments<sup>25</sup>, the acceptor (AO) signal would depend on the excitation of the donor chromophore (6MI), which is excited using polarized light. By detecting smFRET and smFLD signals, it might be possible to define important intermediate steps in the intercalation process.

Our groups have a long history of exploring ‘DNA breathing’ and related topics. It is our goal to elucidate this intuitive, yet sophisticated regime of DNA dynamics based on the insights gained from previous studies<sup>26,27</sup>. Intercalation can be considered as a perturbation (or consequence) of natural DNA breathing<sup>28,29</sup>. Thus, by exploring the properties of intercalation complexes we should be able to gather useful information to extend our mechanistic interpretations of both DNA fluctuations and intercalation. Finally, kinetic intercalation studies of planar aromatic compounds into specific sites in dsDNA constructs can also be regarded as models for biologically important stacking and intercalation interactions of defined aromatic amino acid residues of proteins with DNA<sup>30</sup>, and our group should be in a good position to interpret such results mechanistically. We look forward to future studies along these lines.

## CHAPTER VI

### CONCLUDING REMARKS

#### Summary

The work I presented in this dissertation focuses on the study of DNA local base conformation using 6-methyl Isoxanthopterin (6-MI) labeled DNA. By studying the spectroscopic properties of 6-MI within its surrounding DNA environment, we can learn about nucleic acid base stacking that play an important role in biological processes.

We first examined the conformational heterogeneity of 6-MI in various sequence contexts of duplex and ssDNA constructs. We implemented a simple two-state model to interpret the excitation–emission peak shift (EES) measurements of 6-MI. We obtained excellent agreement between theory and experiment. Our pH-dependent EES measurements of the 6-MI NMP in solution show that these states can be identified as protonated and deprotonated forms of the 6-MI fluorescent probe. Our results suggest that proton transfer in 6-MI-substituted DNA constructs is coupled to conformational heterogeneity of the probe base and can be interpreted to suggest that Watson–Crick base pairing between 6-MI and its complementary cytosine in duplex DNA involves a “low-barrier-hydrogen-bond”.

We next utilized these findings in the study of DNA local conformation facilitated by circular dichroism (CD). We have studied the CD of 6-MI-substituted ss and dsDNA constructs in different sequence contexts. In order to quantitatively analyze CD spectra, we have constructed a simple theoretical model to aid in our understanding of local interactions between a fluorescent base analogue site-specifically positioned within a DNA construct and the native bases and ligands that can comprise its immediate environment. Qualitative

agreement between CD calculations and experimental results reinforces the idea that 6-MI could be a useful probe of DNA secondary structure. Such CD calculation can be easily applied to DNA constructs with known conformation. Our results suggest that the extent of base stacking of 6-MI probe is different among DNA constructs.

We next applied 6-MI to probe the ligand insertion of small molecules to duplex DNA, further extending the potential of 6-MI as a useful reporter of local nucleic acid base conformation. We have determined in this study that 6-MI and AO can serve as a Förster resonance energy transfer (FRET) donor-acceptor chromophore pair. We have shown that, compared to guanine residues inserted at the same site, there is no significant preference for, or discrimination against, intercalation of AO moieties at or near the 6-MI labeled probe position. Thus, 6-MI can be used as a ‘non-discriminating’ reporter of local AO intercalation at and near guanine bases. We have also explored the potential of using single site-substituted 6-MI DNA constructs to measure interactions with tethered ACMA. Our results suggest that, unlike free ACMA that interacts with dsDNA by general intercalation, covalently-tethered ACMA displaces the base on the complementary strand across from the ACMA attachment site.

Throughout this dissertation, the application possibilities of 6-MI as a fluorescent base analogue have been expanded and have set important foundations for future research.

### **Future Direction**

A number of future research projects could be built on the work presented in this dissertation.

Future time-correlated single photon counting (TCSCP) experiments will be needed to map out the structure of 6-MI potential energy surface. Preliminary results show

individual 6-MI fluorescent lifetime components vary at different wavelengths (emission or excitation) and pH conditions. Multiple lifetime components with varying distributions would confirm the validity of the two-state (or even multiple-state models). Further time-resolved anisotropy experiments should reveal the number of different relaxation modes that exist in sequence-dependent 6-MI systems, and their relative time scales.

The 6-MI/AO FRET pair characterization grants possible future smFLD (Fluorescence Linear Dichroism) experiments, in which the acceptor (AO) signal would depend on the excitation of the donor chromophore (6MI), which is excited using polarized light. By detecting smFRET and smFLD signals, it might be possible to define important intermediate steps in the intercalation process.

The simple theoretical model we constructed to quantitatively analyze CD spectra could be applied to other biological systems of different DNA constructs and fluorescent probe. For example, such CD calculation could be applied to the system of gp32 protein and dimer 6-MI labeled ssDNA. We have observed CD signal change with the addition of gp32. It could come from the nondegenerate electronic coupling of tryptophan and 6-MI. We could learn how tryptophan site might be involved in gp32 protein binding to 6-MI labeled site.

More advanced exciton coupling theory could be incorporated to the CD calculation protocols and computational modules presented in this work to evaluate the adequacy of applying point dipole approximation in the CD calculation.

Finally, we are involved in on-going collaborations in which we have applied extended dipole model to combine the MD simulation and the CD calculation. CD offers more comprehensive experimental results for simulation results to compare to. As CD is

the net result of all conformation of dinucleotide in solution, CD calculation might be utilized to investigate the dynamic of dinucleotide in solution.

## APPENDIX A

### SUPPLEMENTARY INFORMATION FOR CHAPTER II: SEQUENCE-DEPENDENT CONFORMATIONAL HETEROGENEITY AND PROTON TRANSFER REACTIVITY OF THE FLUORESCENT GUANINE ANALOG 6-METHYL ISOXANTHOPTERIN (6-MI) IN DNA

#### *Homogeneous One-State Model for Excitation-Emission Shift (EES) Experiments*

We first consider a chemical system that exists as a single optical conformation with ground and excited electronic states that couple to nuclear degrees of freedom. The system is depicted in Figure 2.2A using displaced harmonic functions of a single generalized coordinate  $Q$ . Ground and excited state free energy surfaces are, respectively, described by

$$G_{G1}(Q) = \frac{1}{2}\alpha_1 Q^2 + G_{g1} \quad (\text{A.1})$$

and

$$G_{E1}(Q) = \frac{1}{2}\alpha_1(Q - d_1)^2 + G_{e1}. \quad (\text{A.2})$$

In Eq. (A.1) and (A.2),  $\alpha_1$  is a polarizability that characterizes the curvature of the free energy surfaces. For simplicity, we have assumed these to be the same in the ground and excited states. The displacement  $d_1$  of the excited state surface relative to the ground state is a measure of the coupling between electronic excitation and vibrational motions, and it is related to the reorganization energy according to  $\lambda_1 = \frac{1}{2}\alpha_1 d_1^2$ .

The energy gap between ground and excited surfaces is thus parameterized by the coordinate  $Q$

$$E_{ex}(Q) = G_{E1}(Q) - G_{G1}(Q). \quad (\text{A.3})$$

Substitution of Eqs. (A.1) and (A.2) into Eq. (A.3) yields

$$E_{ex}(Q) = G_{e_1g_1} + \lambda_1 - 2\lambda_1 Q/d_1, \quad (\text{A.4})$$

where we have set  $G_{e_1g_1} = G_{e_1} - G_{g_1}$ . Equation (A.4) describes the energy of an optical field needed to resonantly excite the system, which depends on the nuclear coordinate  $Q$ . Because the peak absorption may vary from system to system, we define the rescaled excitation energy

$$X(Q) \equiv E_{ex}(Q) - E_{ex}(0) = -2(\lambda_1/d_1)Q, \quad (\text{A.5})$$

which has its origin at the absorption peak energy  $E_{ex}(0)$ .

Immediately after electronic excitation, we assume that the system undergoes rapid vibrational relaxation so that the peak emission energy  $E_{em}^{max}$  occurs at  $Q = d_1$ .

$$E_{em}^{max} = G_{E_1}(d_1) - G_{G_1}(d_1) = -\lambda_1 + G_{e_1g_1} \quad (\text{A.6})$$

The ‘excitation-emission shift’ (i.e., the EES) is the difference in energy between the peak emission, and the resonant excitation, which is again parameterized by  $Q$

$$\begin{aligned} Y(Q) \equiv E_{ex}(Q) - E_{em}^{max} &= -2(\lambda_1/d_1)Q + 2\lambda_1 \\ &= X(Q) + 2\lambda_1. \end{aligned} \quad (\text{A.7})$$

The second equality in Eq. (A.7) makes the substitution given by Eq. (A.5).

Equation (A.7) shows that for the homogeneous one-state model, the excitation-emission shift  $Y(Q)$  is a linear function of  $X(Q)$ , with unit slope and y-intercept equal to twice the reorganization energy (i.e. the Stokes’ shift). A plot of the Eq. (A.7) versus the rescaled excitation energy is shown in Figure 2.2 B.

### ***Heterogeneous Two-State Model for Excitation-Emission Shift (EES) Experiments***

We next generalize the above model to describe a heterogeneous mixture of multiple conformations, each determined by a different ground and excited free energy surface. We



consider the simplest case of a two-state system, which is depicted in Figure 2.2 C. The ground and excited free energy surfaces of sub-state 2 are given by

$$G_{G2}(Q) = \frac{1}{2}\alpha_2(Q - x)^2 + G_{g2} \quad (\text{A.8})$$

and

$$G_{E2}(Q) = \frac{1}{2}\alpha_2(Q - x - d_2)^2 + G_{e2}, \quad (\text{A.9})$$

respectively, while those of sub-state 1 are given by Eqs. (A.1) and (A.2). In Eqs. (A.8) and (A.9), the parameter  $x$  is the displacement of the ground free energy surface of sub-state 2 relative to that of sub-state 1. The parameter  $d_2$  specifies the displacement of the excited state surface of sub-state 2 relative to its ground state.

We assume that the system is in thermal equilibrium, so that each sub-state may be assigned a  $Q$ -dependent Boltzmann weight

$$P_{1(2)}(Q) = \exp[-G_{G1(2)}(Q)/k_B T]/Z(Q), \quad (\text{A.10})$$

where  $Z(Q) = P_1(Q) + P_2(Q) = 1$  is the partition function. We note that in this model, only the ground state energies contribute to the Boltzmann weights  $P_1$  and  $P_2$ , so that the possibility of photochemical processes that alter these values are not considered. In this two-state system, the energy gap between ground and excited free energy surfaces is a Boltzmann weighted sum of contributions from sub-states 1 and 2

$$\begin{aligned} E_{ex}(Q) &= P_1(Q)[G_{E1}(Q) - G_{G1}(Q)] + P_2(Q)[G_{E2}(Q) - G_{G2}(Q)] \\ &= P_1(Q)[-2(\lambda_1/d_1)Q + \lambda_1 + G_{e1g1}] \\ &\quad + P_2(Q)[-2(\lambda_2/d_2)Q + \lambda_2 + G_{e2g2} + 2(\lambda_2/d_2)x]. \end{aligned} \quad (\text{A.11})$$

In Eq. (A.11), we have applied the definitions  $G_{e_2g_2} = G_{e_2} - G_{g_2}$  and  $\lambda_2 = \frac{1}{2}\alpha_2d_2^2$ . The rescaled excitation energy is thus given by

$$X(Q) = -2[P_1(Q)(\lambda_1/d_1) + P_2(Q)(\lambda_2/d_2)]Q. \quad (\text{A.12})$$

The peak emission energy is also a Boltzmann weighted sum of terms. Emission from sub-state 1 occurs with energy gap defined at  $Q = d_1$ , and from sub-state 2 with energy gap at  $Q = x + d_2$ . The peak emission energy is therefore given by

$$\begin{aligned} E_{em}^{max}(Q) &= P_1(Q)[G_{E_1}(d_1) - G_{G_1}(d_1)] \\ &+ P_2(Q)[G_{E_2}(x + d_2) - G_{G_2}(x + d_2)] \\ &= P_1(Q)[- \lambda_1 + G_{e_1g_1}] + P_2(Q)[- \lambda_2 + G_{e_2g_2}]. \end{aligned} \quad (\text{A.13})$$

We calculate the EES by subtracting Eq. (A.13) from Eq. (A.11)

$$\begin{aligned} Y(Q) &\equiv E_{ex}(Q) - E_{em}^{max}(Q) \\ &= -2 \left[ \frac{P_1(Q)\lambda_1}{d_1} + \frac{P_2(Q)\lambda_2}{d_2} \right] Q + 2P_1(Q)\lambda_1 + 2P_2(Q)\lambda_2(1 + x/d_2) \\ &= X(Q) + 2P_1(Q)\lambda_1 + 2P_2(Q)\lambda_2(1 + x/d_2). \end{aligned} \quad (\text{A.14})$$

In the final equality of Eq. (A.14), we have made use of the expression for the rescaled excitation energy given by Eq. (A.12).

Equation (A.14) is the most general expression for the EES of a two-state system. The limiting behavior of Eq. (A.14) is determined by considering the cases where the population resides entirely in sub-state 1, or in sub-state 2. In the former case ( $P_1 = 1, P_2 = 0$ ), Eq. (A.14) reduces to the linear expression for a homogeneous one-state system, as given by Eq. (A.7). In the latter case ( $P_1 = 0, P_2 = 1$ ), Eq. (A.14) reduces to the linear expression

$$Y(Q) = -\frac{2\lambda_2}{d_2}Q + 2\lambda_2(1 + x/d_2) = X(Q) + 2\lambda_2(1 + x/d_2). \quad (\text{A.15})$$

Equation (A.15) shows that when population resides entirely in sub-state 2, the EES is a linear function of the reduced excitation energy with unit slope and y-intercept  $2\lambda_2(1 + x/d_2)$ , which is vertically displaced from that of the opposite limit, as shown in Figure 2.2 D.

By inspection of Eq. (A.14), one can see that the effect of tuning the excitation energy across the absorption band maximum (by varying the parameter  $Q$  from positive to negative values) is to adjust the relative contribution of sub-states 1 and 2 to the EES. The EES thus undergoes an energy-dependent transition from the linear behavior described by Eq. (A.7) to the linear behavior described by Eq. (A.15), as shown in Figure 2.2 D for three different energetic conditions. The mid-point of the transition occurs at the value of  $Q = d_c$ , where the ground free energy surfaces of sub-states 1 and 2 cross. At this point in the EES plot  $[X(d_c), Y(d_c)]$ , the populations of the sub-states 1 and 2 are exactly equal (i.e.  $P_1 = P_2 = 1/2$ ). The Cartesian coordinates of the transition point are given by

$$X(d_c) = -\left(\frac{\lambda_1}{d_1} + \frac{\lambda_2}{d_2}\right) d_c \quad (\text{A.16})$$

and

$$Y(d_c) = -\left(\frac{\lambda_1}{d_1} + \frac{\lambda_2}{d_2}\right) d_c + \lambda_1 + \lambda_2(1 + x/d_2) \quad (\text{A.17})$$

The two-state expression for the EES given by Eq. (A.14) can be greatly simplified by making the additional assumptions  $\lambda_1 = \lambda_2 = \lambda$ , and  $d_1 = d_2 = d$ . Under these conditions, Eq. (A.14) reduces to Eq. (2.1) of the main text and Eq. (A.16) and (A.17) take on the forms  $X_c = -2(\lambda/d)d_c$  and  $Y_c = X_c + \lambda[2 + (x/d)]$ .

### ***Formulas for Free Energies of Transition and Activation***

The crossing point is  $Q = d_c$  is related to the free energies of transition and activation through the Marcus formulas<sup>1-2</sup>. Under this condition, the energies of the two ground free energy surfaces are equal, i.e.  $G_{G1}(d_c) = G_{G2}(d_c)$ . This condition when applied to Eqs. (A.1) and (A.8) can be used to derive the Marcus formulas described by Eqs. (2.2) and (2.3) of the main text.

### ***Two-State Model Fitting Procedure for EES Data***

As described in the main text, we analyzed our EES data for both 6-MI NMP samples in aqueous solution, as well as 6-MI substituted ssDNA, using Eq. (2.1) of the main text. Equation (2.1) makes use of the assumptions  $\lambda_1 = \lambda_2 = \lambda$ , and  $d_1 = d_2 = d$ . For our analysis of 6-MI substituted dsDNA constructs, it was necessary to allow for  $\lambda_1 \neq \lambda_2$ , in which case we fit our data to Eq. (A.14).

For all of the analyses performed in this work, we made the simplifying assumption  $d_1 = d_2 = d = 1$ , so that the parameters  $x$ ,  $\lambda_{1(2)}$ , and  $d_c$  can be considered to be measured in units of  $d$ . We determined the value of the reorganization energy  $\lambda_1$  by extrapolating the EES data to a diagonal line in the limit of large excitation energy [ $X \equiv E_{ex}(Q) - E_{ex}(0)$ ]. This line has the form described by Eq. (A.7), with y-intercept equal to the Stokes' shift, which is equal to twice the value of  $\lambda_1$ . For cases in which  $\lambda_1 \neq \lambda_2$ , we treated  $\lambda_2$  as an adjustable variable in the expression given by Eq. (A.15).

In fitting our EES data to Eq. (2.1) [or alternatively to Eq. (A.14)], it was necessary to determine the Boltzmann weights  $P_1(Q)$  and  $P_2(Q) = 1 - P_1(Q)$ , which are given by Eq. (A.10). The Boltzmann weights are functions of the parameters that characterize the ground state free energy surfaces of sub-states 1 and 2, which are given by Eqs. (A.1) and (A.8), respectively. These surfaces are completely specified by the parameters  $x$ ,  $\lambda_{1(2)}$ , and

$\Delta G^\circ = G_{g2} - G_{g1} = G_{g2}$  (setting  $G_{g1} = 0$ ). We further made use of the fact that the polarizability may be written  $\alpha_{1(2)} = 2\lambda_{1(2)}/d^2$ . In practice, once the value(s) of  $\lambda_{1(2)}$  were determined by the extrapolation procedure described above, we then self-consistently varied the parameters  $x$  and  $\Delta G^\circ$  to obtain the best fit of Eq. (2.1) or Eq. (A.14) to the experimental EES data. These optimized values were then substituted into Eq. (2.2) to determine the crossing point  $d_c$ , which was in turn used in Eq. (2.3) to determine the free energy of activation  $\Delta G^\ddagger$ .

**Table A.1.** Parameters for fitting EES plot of 6-MI NMP.

pH (cm <sup>-1</sup> )	<sup>a</sup> $\Delta G^\circ$ (cm <sup>-1</sup> )	<sup>a</sup> $\Delta G^\ddagger$ (cm <sup>-1</sup> )	$x$ (cm <sup>-1</sup> )	$\lambda_1$	$d_c$ (cm <sup>-1</sup> )	$\lambda_2$	<sup>b</sup> $K_{eq}$
5	900	2104	0.21	2829	0.86	2829	0.01 ± 0.00
6	863	1965	0.21	2811	0.84	2811	0.02 ± 0.01
7	638	1169	0.21	2815	0.65	2815	0.05 ± 0.02
7.6	525	794	0.22	2870	0.53	2870	0.06 ± 0.02
8	350	448	0.21	2872	0.40	2872	0.18 ± 0.07
8.3	200	211	0.21	2867	0.27	2867	0.38 ± 0.15
8.5	200	211	0.21	2868	0.27	2867	0.38 ± 0.15
9	-350	99	0.21	2862	-0.19	2862	5.4 ± 2.1
9.5	-500	276	0.21	2868	-0.31	2868	11.2 ± 4.3
10	-450	207	0.21	2868	-0.27	2868	8.8 ± 3.4
10.3	-325	109	0.21	2887	-0.17	2887	5.7 ± 4.8

<sup>a</sup> Standard free energies of reaction ( $\Delta G^\circ$ ) and free energies of activation ( $\Delta G^\ddagger$ ) defined in Eqs. (2.2) and (2.3). Uncertainties of  $\Delta G^\circ$  and  $\Delta G^\ddagger$  are ± 100 cm<sup>-1</sup>.

<sup>b</sup> Equilibrium constant for the reaction [6-MI (protonated)] ⇌ [6-MI<sup>-</sup> (deprotonated)].

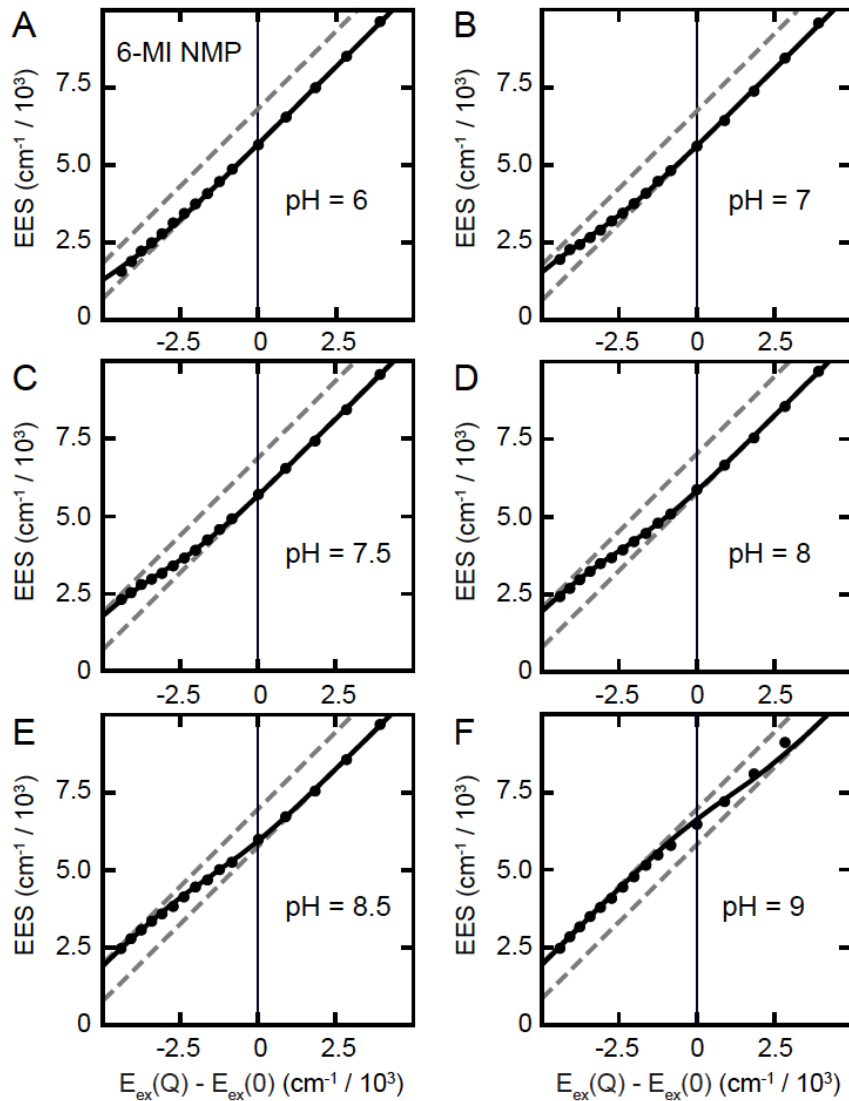
**Table A.2.** Parameters for fitting EES plot of ssDNA <sup>a</sup>

	$\Delta G^\circ$ (cm <sup>-1</sup> )	$\Delta G^\ddagger$ (cm <sup>-1</sup> )	$x$	$\lambda_1$ (cm <sup>-1</sup> )	$d_c$	$\lambda_2$ (cm <sup>-1</sup> )	$K_{eq}$
<b>pH 7.6</b>							
6-MI	600	1010	0.22	2829	0.60	2829	0.06 ± 0.02
GXG	750	1038	0.31	2419	0.66	2419	0.03 ± 0.01
GXXG	360	381	0.31	2329	0.40	2329	0.18 ± 0.07
TXT	800	1159	0.30	2534	0.68	2534	0.02 ± 0.01
TXXT	900	1264	0.32	2650	0.69	2650	0.01 ± 0.00
TXG	600	632	0.31	2521	0.54	2521	0.06 ± 0.02
TXTXG	600	735	0.31	2495	0.54	2495	0.06 ± 0.01
<b>pH 6</b>							
6-MI	913	2157	0.21	2827	0.87	2827	0.01 ± 0.00
GXG	700	900	0.32	2454	0.61	2454	0.03 ± 0.01
GXXG	550	521	0.32	2450	0.46	2450	0.07 ± 0.03
TXT	850	1228	0.31	2534	0.70	2534	0.02 ± 0.01
TXXT	975	1479	0.31	2741	0.73	2741	0.01 ± 0.00
TXG	600	732	0.31	2521	0.54	2521	0.06 ± 0.02
TXTXG	700	921	0.31	2495	0.61	2495	0.03 ± 0.01

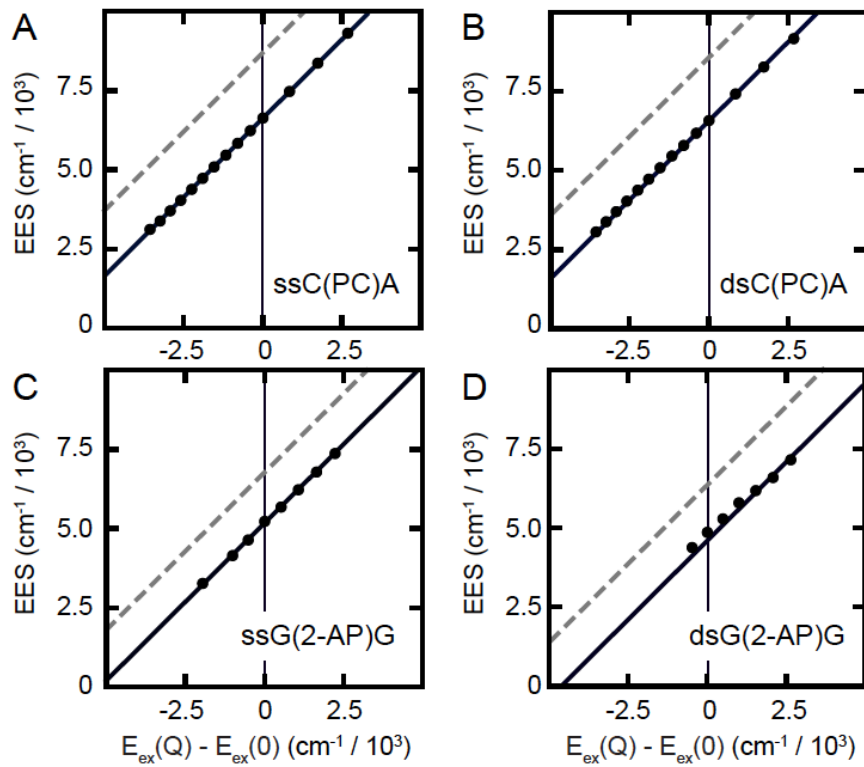
<sup>a</sup> Abbreviations as in Table A.1.**Table A.3.** Parameters for fitting EES plot of dsDNA

	$\Delta G^\circ$ (cm <sup>-1</sup> )	$\Delta G^\ddagger$ (cm <sup>-1</sup> )	$x$	$\lambda_1$ (cm <sup>-1</sup> )	$d_c$	$\lambda_2$ (cm <sup>-1</sup> )	$K_{eq}$
<b>pH 7.6</b>							
GXG	250	253	0.32	1987	0.36	2500	0.30 ± 0.11
GXXG	500	594	0.31	2238	0.52	2800	0.09 ± 0.03
TXT	900	1420	0.31	2304	0.79	2800	0.01 ± 0.00
TXXT	450	668	0.23	2374	0.54	2800	0.11 ± 0.04
TXG	250	293	0.23	2100	0.37	2700	0.30 ± 0.11
TXTXG	550	914	0.23	2350	0.62	2800	0.07 ± 0.03
<b>pH 6</b>							
GXG	250	253	0.31	2101	0.35	2500	0.30 ± 0.11
GXXG	550	663	0.31	2376	0.53	2800	0.07 ± 0.03
TXT	950	1184	0.4	2284	0.72	2500	0.01 ± 0.00
TXXT	550	810	0.25	2437	0.58	2800	0.07 ± 0.03
TXG	289	347	0.23	2300	0.39	2800	0.25 ± 0.09
TXTXG	640	1004	0.26	2350	0.65	2800	0.05 ± 0.02

<sup>a</sup> Abbreviations as in Table A.1.

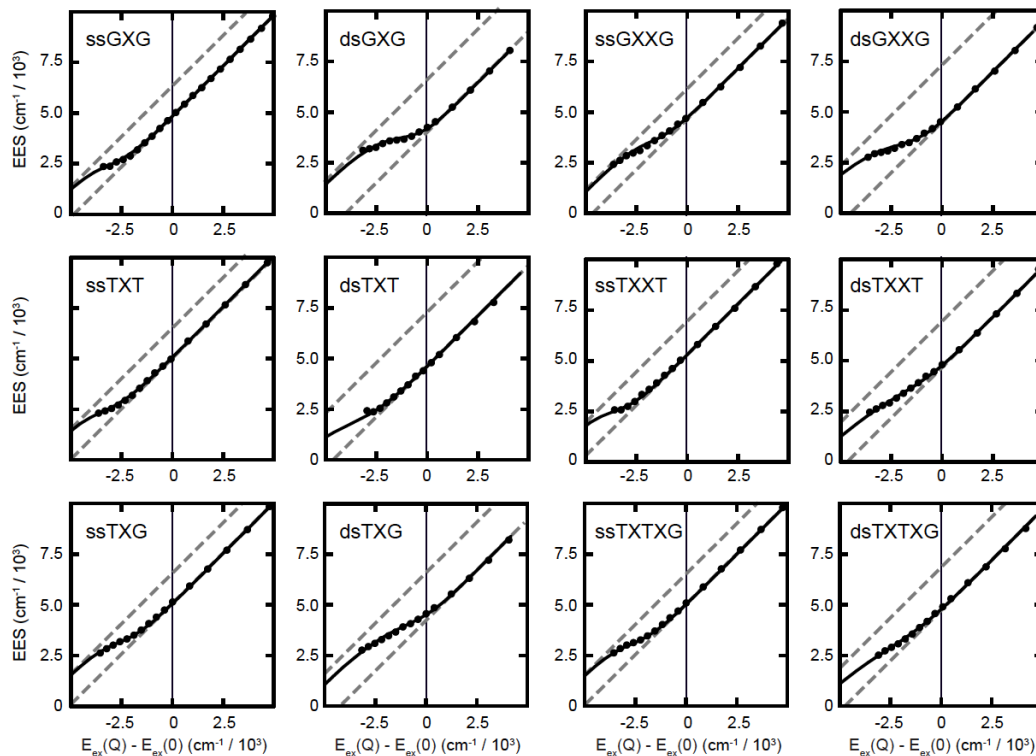


**Figure A.1** EES data of 6-MI nucleotide monophosphate (NMP) in aqueous solution, as a function of rescaled excitation energy and pH. (A-F) pH = 6, 7, 7.5, 8, 8.5, and 9, respectively. Solid curves are fits to the data using Eq. (2.1) of the main text. The values of the optimized parameters are given in Table A.1.

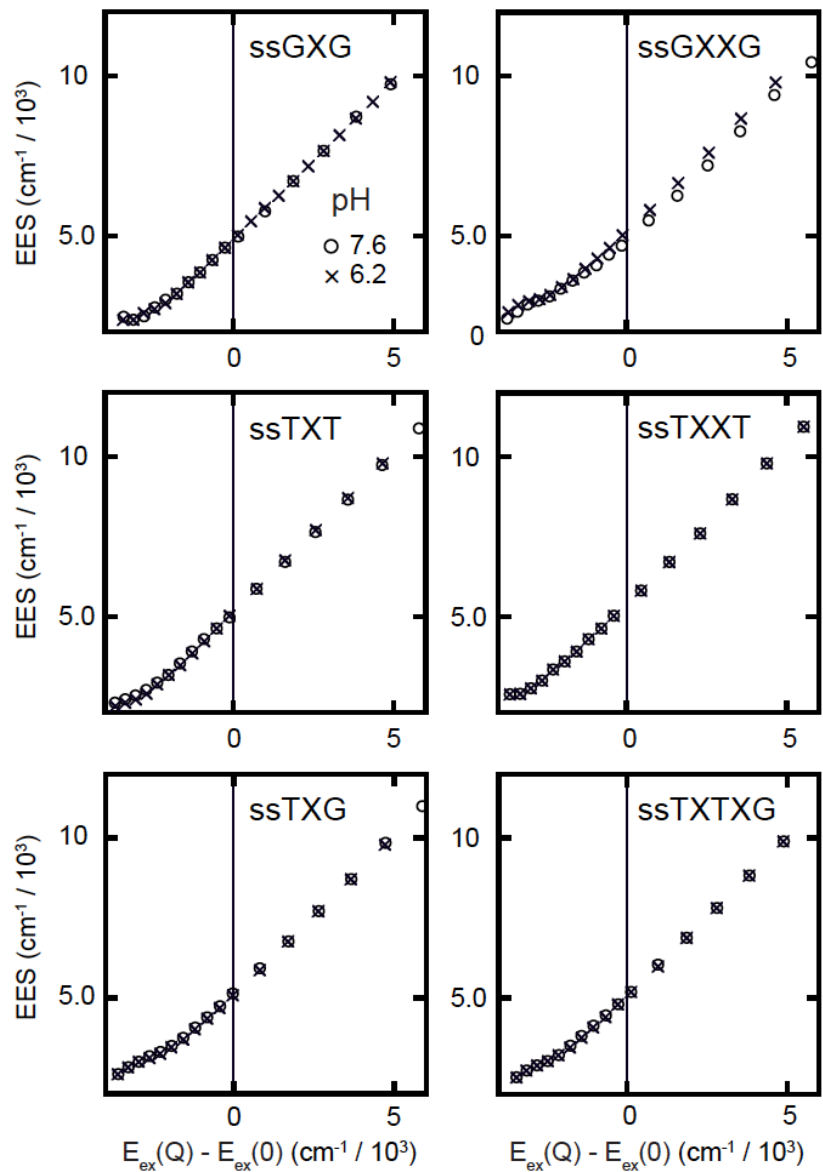


**Figure A.2** EES data on pyrrolocytosine (PC) and 2-aminopurine (2-AP) substituted DNA, as a function of rescaled excitation energy. **(A)** ssC(PC)A, **(B)** dsC(PC)A, **(C)** ssG(2-AP)G and **(D)** dsG(2-AP)G. pH = 7.5, 20 °C.

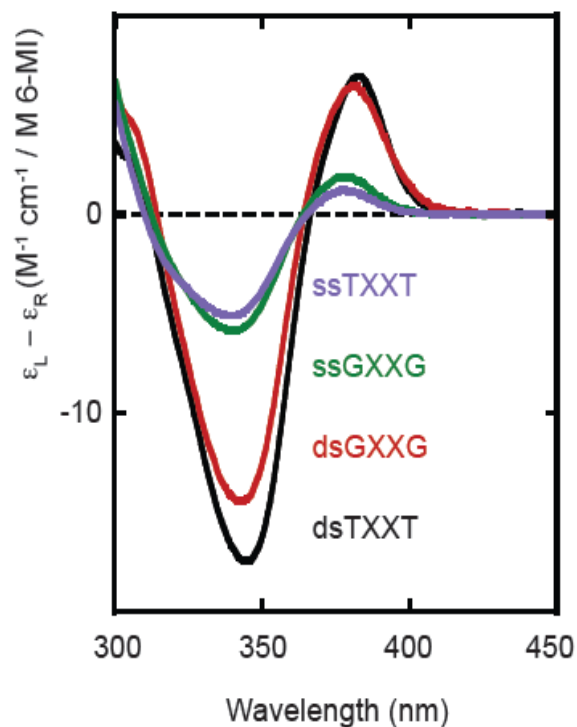




**Figure A.3** EES versus excitation energy for 6-MI substituted ss and dsDNA constructs. Black circles, experimental results; black line theoretical fit; red and green lines, EES of protonated and deprotonated 6-MI probe, respectively. pH = 7.5, 20 °C. Solid curves are fits to the data using Eq. (2.1) of the main text for the ssDNA substrates, and using Eq. (A.14) for the dsDNA substrates. The values of the optimized parameters are given in Tables A.2 and A.3.



**Figure A.4** EES data versus rescaled excitation energy for 6-MI substituted ssDNA, as a function of pH = 6.2 (crosses), 7.6 (circles).



**Figure A.5** CD spectra of 6-MI residues in various dsDNA constructs. dsTXXT (black), dsGXXG (red), ssTXXT (magenta), ssGXXG (green). DNA sequences and nomenclature as in Fig. 1. The units of CD are given in terms of  $\epsilon_L - \epsilon_R$  (in  $M^{-1} \text{ cm}^{-1}$ ) per mole of 6-MI residue.

In duplex DNA the CD of adjacent 6-MI residues have a bimodal shape centered at the absorption maximum that is characteristic of B form DNA<sup>3,4</sup>. CD measurements show that the 6-MI dimer is stacked in a right helical conformation in duplex DNAs (Figure A.5). ssTXXT and ssGXXG also exhibited some exciton coupling characteristic of base stacking, although this signal was less than for the corresponding dsDNA.

## APPENDIX B

SUPPLEMENTARY INFORMATION FOR CHAPTER III: THEORETICAL  
ASPECTS OF CIRCULAR DICHROISM CALCULATIONS*Alternative EDTMs Assignments for Natural Bases***Table B.1.** Comparison of experimental electronic transition moment in Thymine and Derivatives. The EDTMs assignments highlighted in gray are chosen in calculation.

Compound	Transitions	Transition energy center (cm <sup>-1</sup> ) / (nm)	oscillator strength	Orientation * (°)	Dipole radius (Å)	ref
1-methylthymine	1	36630(273)	0.19	14±2 or -		ref <sup>1</sup>
	2	48309(207)	0.28	19±2 -76±2 or 71±2		
1-Methyluracil	1	36364(275)	0.195	-9		ref <sup>2</sup>
	2	46948(213)	0.260	-53 or 59		
	1	37500(266.7)	0.19	0±1 or 7±1		ref <sup>3</sup>
Thymine	1	37594(266)		-31±6 or 51±20		ref <sup>4</sup>
	1	36000(277.8)		-12		ref <sup>2,5</sup>
Thymidine	S <sub>0</sub> → S <sub>1</sub>	37500(266.7)	0.24		0.77	ref <sup>6</sup>
Thymine or Uracil	1	265	3.27	-19	0.68	ref <sup>7</sup>
	2	215	2.14	71	0.45	
	3	195	2.04	-36	0.43	
	4	177	3.88	-20	0.81	

**Table B.2.** Comparison of experimental electronic transition moment in Adenine and Derivatives. The EDTMs assignments highlighted in gray are chosen in calculation.

Compound	Transitions	Transition energy center (cm <sup>-1</sup> ) / (nm)	oscillator strength	Orientation (°)	Dipole radius (Å)	ref
9-methyladenine	1	36364(275)	0.10	83		ref <sup>8,9</sup>
	2	37736(265)	0.20	25		
	3	46948(213)	0.25	-45		
	4	49020(204)	0.11	15		
	5	53476(187)	0.30	72		
	6	57143(175)	~0.1	~45		
	7	62500(160)	0.23	6		
	8	68027(147)	~0.1	~-45		
	1	275		-3±3		ref <sup>10</sup>
	1	36710(272.4)	0.047	66±7	0.34	ref <sup>11</sup>
	2	38820(257.6)	0.24	19±7	0.75	
	3	43370(230.6)	0.027	-15±6	0.24	
	4	46840(213.5)	0.14	-21±7	0.52	
5	~48320(207.0)	~0.12	-64±10	0.48		
Adenosine	S <sub>0</sub> → S <sub>1</sub>	36700(272.5)	0.05			ref <sup>6</sup>
	S <sub>0</sub> → S <sub>2</sub>	38800(257.7)	0.24			

**Table B.3.** Comparison of experimental electronic transition moment in Cytosine and Derivatives. The EDTMs assignments highlighted in gray are chosen in calculation.

Compound	Transitions	Transition energy center (cm <sup>-1</sup> )/(nm)	oscillator strength	Orientation (°)	Dipole radius (Å)	ref
1-methylcytosine	1	37736(265)		12±3		ref <sup>12</sup>
	2	43478(230)		-11 to 9 (-1)		
Cytosine	1	37736(265)		9 or 51		ref <sup>12</sup>
	2	43478(230)		4 to 56		
	1	37594(266)	0.14	6		ref <sup>13</sup>
	2	43478(230)	0.03	-46		
	3	47170(212)	0.13	76		
	4	50761(197)	0.36	-27 or 86		
	5	60606(165)	0.15	~0 or ~60		
	6	66667(150)	0.20	~0 or ~60		
	1	37000 to 41000		14±1		ref <sup>14</sup>
	2	(270.3 to 243.9) 44000(227.3)		-5±3		
	1	272	2.86	6	0.60	ref <sup>7</sup>
	2	231	2.07	-46	0.43	
	3	214	2.44	76	0.51	
4	196	4.43	-27	0.92		
5	163	2.17	0	0.45		
6	151	2.44	60	0.51		
Cytidine	1	37313(268)		25±3 or -		ref <sup>4</sup>
	2	41667(240)		46±4 6±4 or - 27±3		
	S <sub>0</sub> → S <sub>1</sub>	36800(271.7)	0.21			ref <sup>6</sup>

**Table B.4.** Comparison of experimental electronic transition moment in Guanine and Derivatives. The EDTMs assignments highlighted in gray are chosen in calculation.

Compound	Transitions	Transition energy center (cm <sup>-1</sup> ) /(nm)	oscillator strength	Orientation (°)	Dipole radius (Å)	ref	
9-Ethylguanine	1	36700(272.5)	0.14±0.01	-4±3 or		ref <sup>15</sup>	
	2	40200(248.8)	0.21±0.01	35±4			
	3	48600(205.8)	0.38±0.03	-75±3			
	4	53500(186.9)	0.42±0.03	-71±4 or -			
	5	63000(158.7)		79±4			
	6	65000(153.8)		-9±4 or 41			
	7	69000(144.9)		~0 or ~40			
	1	35971(278)		~0 or ~40		ref <sup>12</sup>	
	2	39526(253)		44±5 or - 14±5 115±10 or 95±10			
	Guanine Hydrochloride	1	36300(275.5)	0.13±0.01	2 or 77		ref <sup>15</sup>
		2	40600(246.3)	0.22±0.01	-82 or -14		
		3	51800(193.1)	0.7	-75 or -17		
	Guanine	1	35714(280)		4±3 or -		ref <sup>4</sup>
		2	40323(248)		61±4 31±3 or - 88±4		
1		272	2.76	-4	0.57	ref <sup>7</sup>	
2		248	3.18	-75	0.66		
3		204	3.92	-75	0.82		
4		187	3.92	-9	0.82		
5		159	2.76	-4	0.57		
6		154	1.84	-75	0.38		
7		145	2.30	-4	0.48		
Guanosine		1	36765(272)	0.15	-24		ref <sup>16</sup>
		2	40650(246)	0.24	88		
		3	50000(200)	0.40	86		
		4	53763(186)	0.48	-8 to 44		
		S <sub>0</sub> → S <sub>1</sub>	36700(272.5)	0.19		0.69	ref <sup>6</sup>
	S <sub>0</sub> → S <sub>2</sub>	40300(248.1)	0.21		0.69		

*EDTMs Assignment of Acridine Orange*

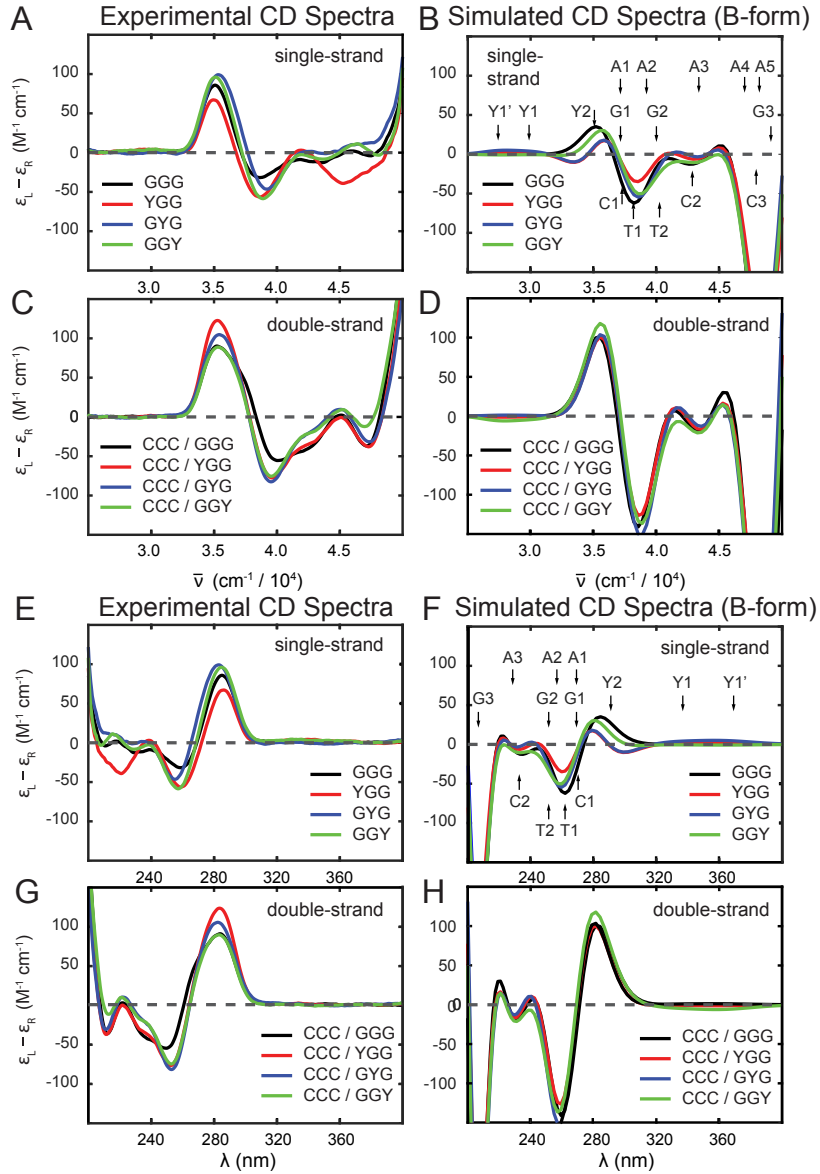
**Table B.5.** Comparison of experimental electronic transition moment in Acridine Orange. The EDTMs assignments highlighted in gray are chosen in calculation.

Compound	Transitions	Transition energy center (cm <sup>-1</sup> )/(nm)		oscillator strength	Orientation (°)	Dipole radius (Å)	ref	
Acridine Orange	1 2 3	19920(502) 30303(330) 36764(272)			-78±2 -20±2 -67±2		ref <sup>7</sup>	
	1 2 3	19920(502) 30303(330) 36764(272)	19500(512.8) 19970(500.8) 21100(473.9) 21300(469.5) 22650(441.5) 22670(441.1) 24190(413.4) 24040(416.0)		-87±2 0 90 0 90 0 90 -8±3 -65±2		ref <sup>8</sup>	
Monomer	1 2 3 4	492 294 261 230			90 0 90 -		ref <sup>9</sup>	
	Dimer	1 2 3 4 5	540 458 294 261 230		1.00 9.18 2.40 10.29 2.50	0 90 0 90 -		

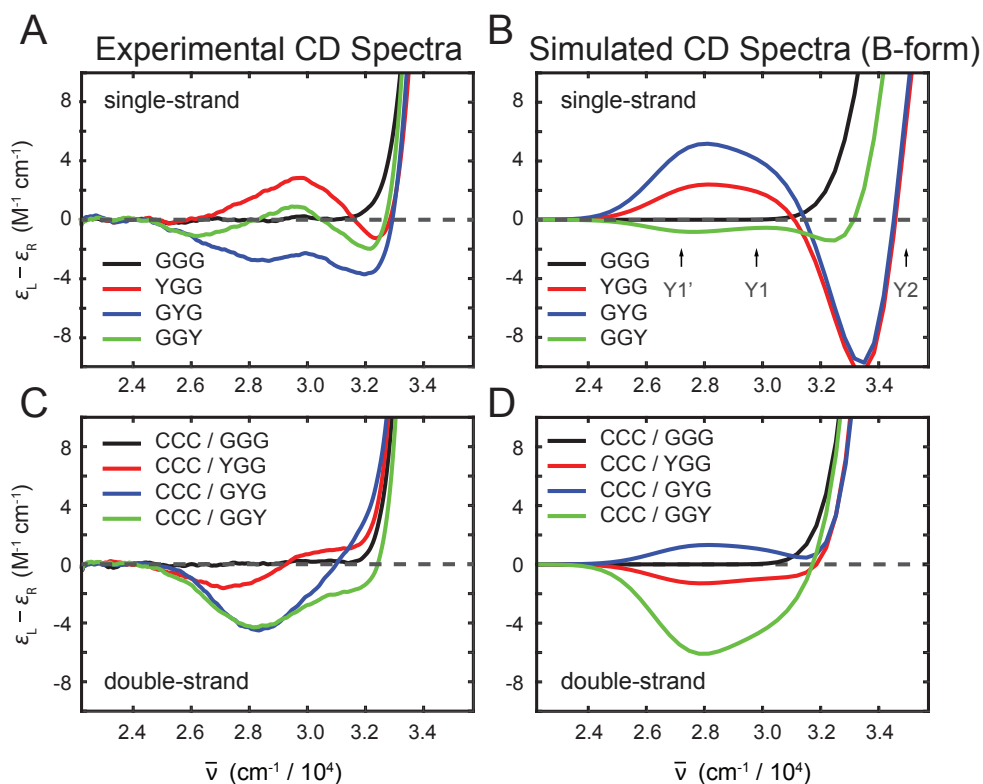


APPENDIX C

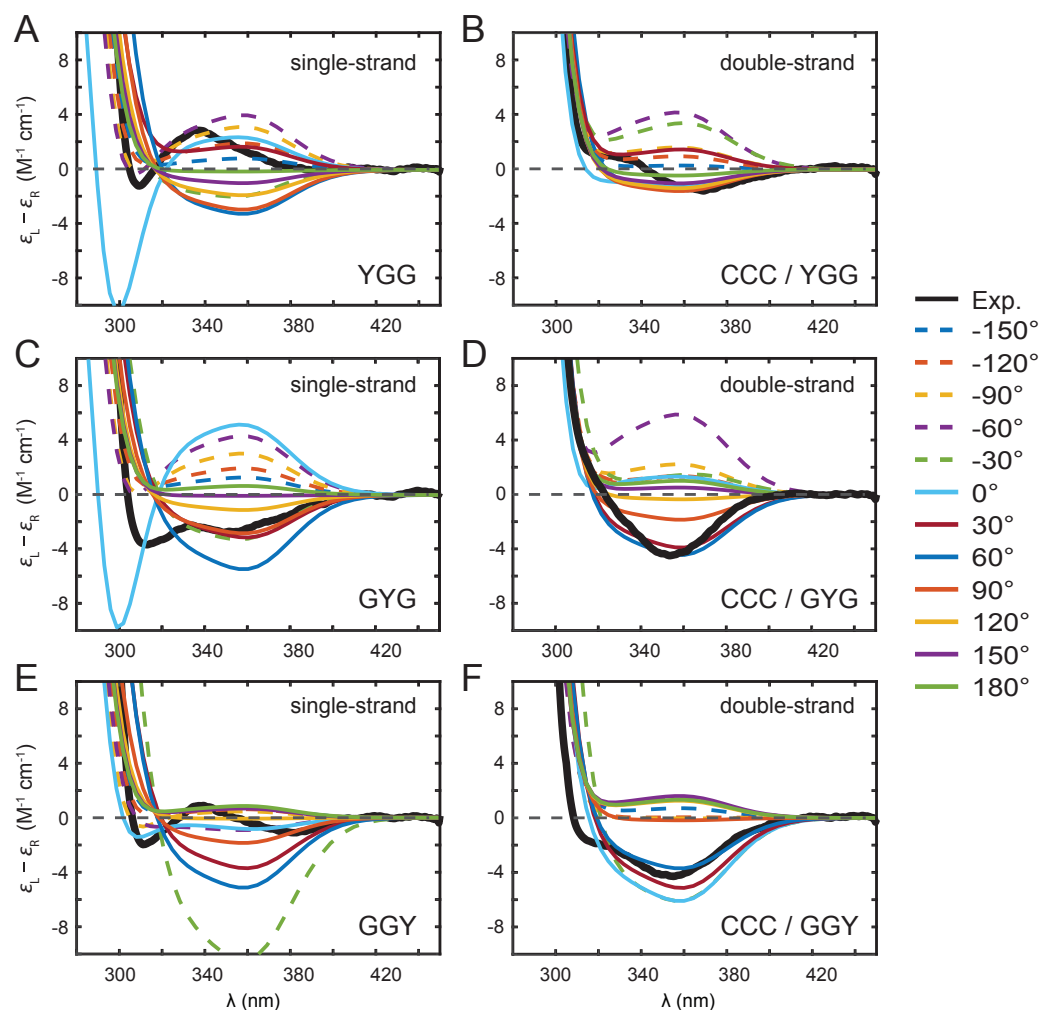
SUPPLEMENTARY INFORMATION FOR CHAPTER IV: LOCAL NUCLEIC ACID BASE CONFORMATION REVEALED BY CIRCULAR DICHROISM STUDIES OF 6-METHYL ISOXANTHOPTERIN SUBSTITUTED DNA CONSTRUCTS



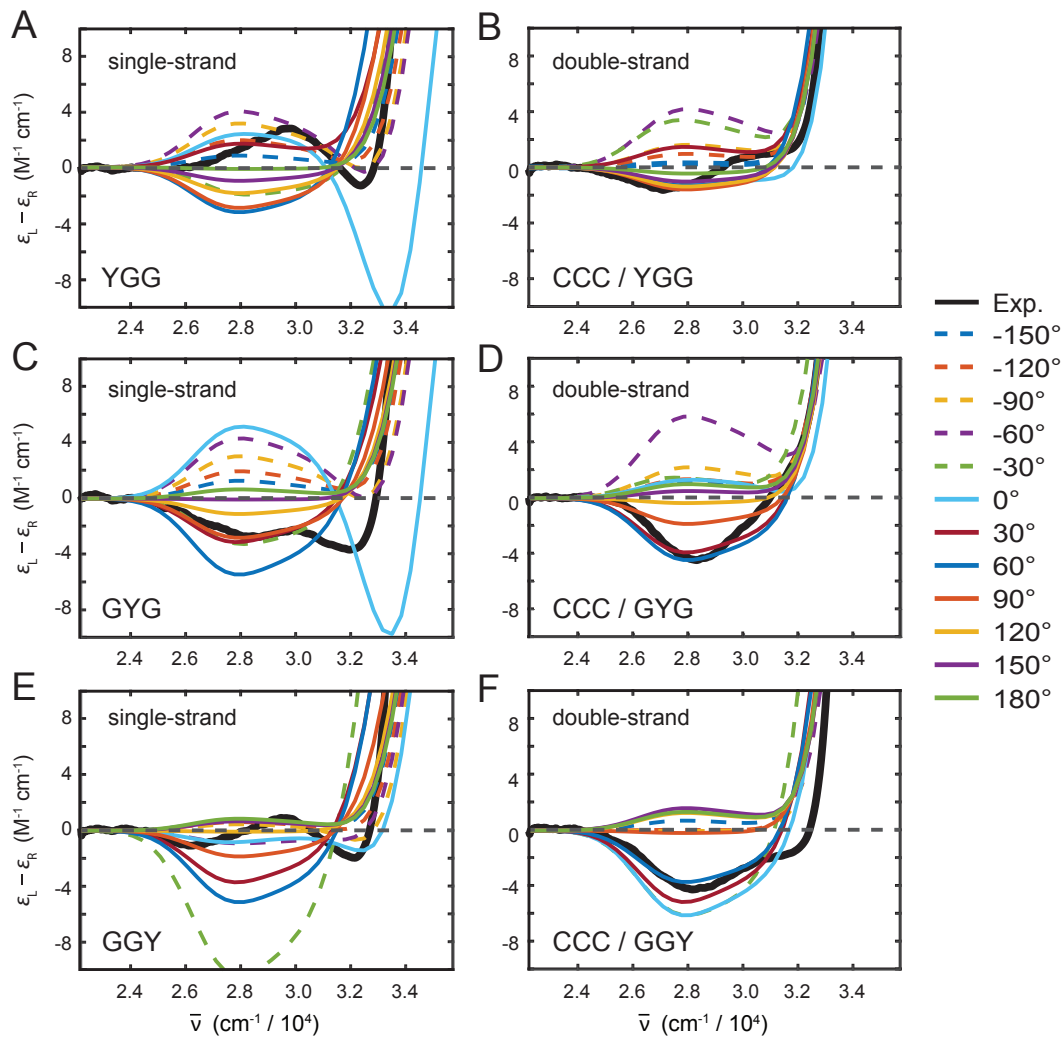
**Figure C.1:** Comparison between experimental CD spectra (panels (A), (C), (E) & (G)) and model calculations (panels (B), (D), (F) & (H)) of 6-MI substituted ss and ds DNA constructs, which are described in Table 4.1 of the main text. The spectra are shown in terms of wavenumbers (panels (A) – (D)) and wavelength (panels (E) – (H)), over the full range of 6-MI probe and nucleotide transition energies. The site EDTM transition energies used in our calculations are indicated by vertical arrows in panels B and F, where Y = 6-MI, A = adenine, C = cytosine, G = guanine, and T = thymine. The values for the EDTM parameters are given in Chapter III.



**Figure C.2:** (A) Experimental and (B) simulated CD of ssDNA constructs, which assume that the nucleotides adopt the B-form conformation. A similar comparison is made for (C) experimental and (D) simulated CD of dsDNA. In panel B, the wavenumbers of the uncoupled 6-MI transitions are indicated: Y1' ( $27303\text{ cm}^{-1}$ ), Y1 ( $29734\text{ cm}^{-1}$ ) and Y2 ( $34596\text{ cm}^{-1}$ ) (see Table 3.6). The spectra are shown in terms of wavenumbers.



**Figure C.3:** Comparison between experimental (thick black curves) and simulated (colored) CD spectra of the 6-MI substituted ss (panels A, C & E) and ds (panels B, D & F) DNA constructs as a function of the 6-MI probe rotation angle. The probe rotation angle is defined in Figure 4.3 of the main text.



**Figure C.4:** Comparison between experimental (thick black curves) and simulated (colored) CD spectra of the 6-MI substituted ss (panels A, C & E) and ds (panels B, D & F) DNA constructs as a function of the 6-MI probe rotation angle. Same as Figure C.3, but with the horizontal axis in wavenumber.

## APPENDIX D

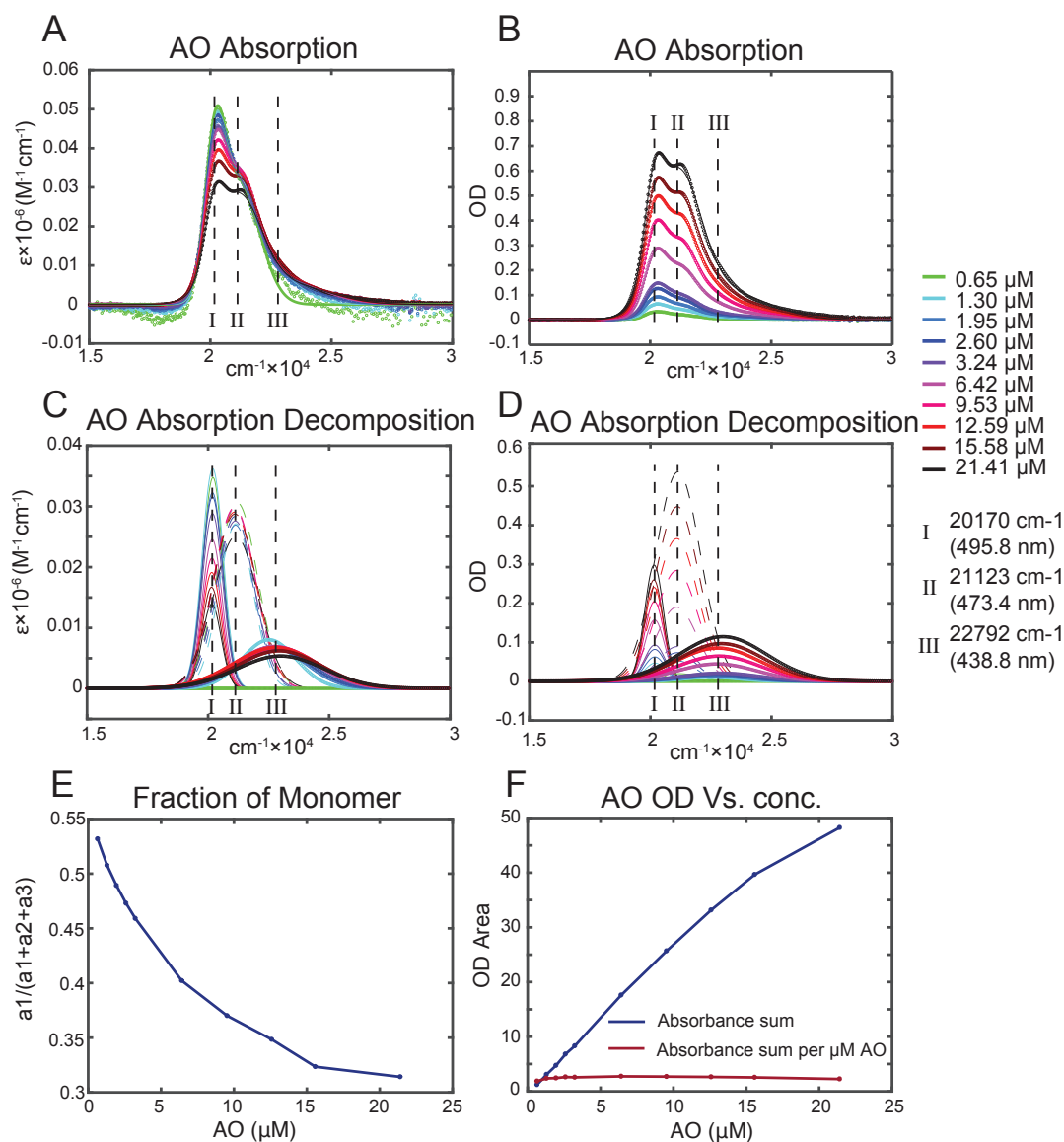
### SUPPLEMENTARY INFORMATION FOR CHAPTER V: STUDIES OF FREE AND TETHERED INTERCALATOR MOLECULES INTERACTING WITH 6-MI LABELED DNA

#### ***D.1 Acridine Orange Absorbance Gaussian Decomposition***

We decomposed the AO absorbance spectra at different concentrations to three Gaussian peaks, and assigned them to the abundance of monomer ( $f_1$ ), dimer ( $f_2$ ), and higher order aggregates ( $f_3$ ). The results of our decomposition analysis are summarized in Figure D.1 and Table D.1.

$$\begin{aligned}f_1(\nu) &= a_1 \times e^{-\left(\frac{\nu-b_1}{c_1}\right)^2} \\f_2(\nu) &= a_2 \times e^{-\left(\frac{\nu-b_2}{c_2}\right)^2} \\f_3(\nu) &= a_3 \times e^{-\left(\frac{\nu-b_3}{c_3}\right)^2} \\f(\nu) &= f_1(\nu) + f_2(\nu) + f_3(\nu)\end{aligned}$$

As the concentration of AO was increased, the relative proportions of the three components changed. Most importantly, the fractional monomer number density decreased with increasing AO concentration (see Fig D.1 E). Thus, one cannot apply the usual Beer-Lambert law to AO since both the extinction coefficient and the value of  $\lambda_{\max}$  was dependent on AO concentration. As a result, we did not determine the AO concentration by its absorption. Instead, we prepared AO solutions with solid AO in a volumetric flask. We found that although the shape of AO absorbance spectrum depended on concentration, the total absorbance [i.e., the integrated optical density (OD) from 350 nm to 600 nm] was linearly proportional to AO concentration at values as high as approximately 15  $\mu\text{M}$  (see Figure D.1 F).



**Figure D.1:** Gaussian decomposition analysis of Acridine Orange (AO) absorbance. (A) AO absorbance spectra and Gaussian fits at various concentrations (indicated by colors). (B) Same as (A), but with original OD at implied concentrations. (C) AO molar absorption spectra at different concentrations decomposed into three Gaussian peaks. (D) AO absorption spectra with original OD decomposed into three Gaussian peaks. (E) Number fraction of monomer at different AO concentrations. (F) Integrated OD across the spectral range 350 – 600 nm were examined at different concentrations. The averages of three peaks were labeled by dashed lines on top of the absorbance spectra.

**Table D.1** AO absorbance fitting parameters

#	[AO] ( $\mu\text{M}$ )	$a_1$	$a_2$	$a_3$	$b_1$ ( $\text{cm}^{-1}$ )	$b_2$ ( $\text{cm}^{-1}$ )	$b_3$ ( $\text{cm}^{-1}$ )	$c_1$ ( $\text{cm}^{-1}$ )	$c_2$ ( $\text{cm}^{-1}$ )	$c_3$ ( $\text{cm}^{-1}$ )
1	0.65	0.0226	0.0199	0	20200	21200	0	585.1	1185	0
2	1.30	0.0471	0.0353	0.0104	20190	21150	22520	619.2	989	1582
3	1.95	0.0629	0.0525	0.0132	20180	21130	22670	614.8	1072	1939
4	2.60	0.0821	0.0737	0.0178	20180	21110	22700	608.6	1107	2066
5	3.24	0.0934	0.0894	0.0207	20180	21120	22910	604.8	1129	2249
6	6.42	0.1578	0.1903	0.0443	20170	21090	22760	579.2	1195	2308
7	9.53	0.2048	0.2842	0.0644	20160	21090	22840	565.6	1225	2385
8	12.95	0.2410	0.3658	0.0850	20150	21090	22790	555.7	1240	2455
9	15.58	0.2602	0.4478	0.0965	20140	21110	22940	544.7	1276	2466
10	21.41	0.2982	0.5368	0.1142	20150	21140	23000	549.3	1312	2543
Ave ( $\text{cm}^{-1}$ )					20170 $\pm$ 20	21123 $\pm$ 34	22792 $\pm$ 150			
Ave (nm)					495.79	473.43	438.75			

## D.2 Determination of Förster Distance ( $R_0$ )

Both theoretical and experimental methods were utilized to determine the Förster distance ( $R_0$ ) of FRET donor-acceptor pair 6-MI NMP and AO. We calculated  $R_0$  based on the spectral overlap using the PyMol Förster distance calculator (See [https://pymolwiki.org/index.php/Forster\\_distance\\_calculator](https://pymolwiki.org/index.php/Forster_distance_calculator) ). In these calculations, we used the value  $\Phi = 0.7$  for the fluorescence quantum yield<sup>1</sup> of the donor (6-MI) in the absence of the acceptor (AO).

We measured  $R_0$  experimentally by fluorescence titration shown in Figure 5.3. The probability of an excited 6-MI molecule transferring its electronic excited state energy by the Förster dipole-dipole coupling mechanism is dependent upon the distribution of AO molecules within its immediate vicinity. In this measurement, energy transfer and other interactions between AO molecules are not considered.<sup>2,3</sup> We assume isotropic distributions for both donors and acceptors. The decay of the donor intensity with increasing acceptor concentration is dependent on the Förster distance  $R_0$ . The following equations show the analytical form for the donor excited state population decay, to which we fit our experimental data. Here  $F_D$  is the 6-MI fluorescence intensity in the absence of AO in solution, and  $F_{DA}$  is the 6-MI fluorescence intensity in the presence of AO at different concentrations. The ratio  $F_{DA} / F_D$  is a function of  $\gamma$ .  $\gamma = A/A_0$  is the ratio of the acceptor concentration ( $A$ ) and the critical concentration ( $A_0$ ). Examples of our fitting results are shown in Figure D.2. Förster distance  $R_0$  (Å) is related to the critical concentration  $A_0$  (mol/L).

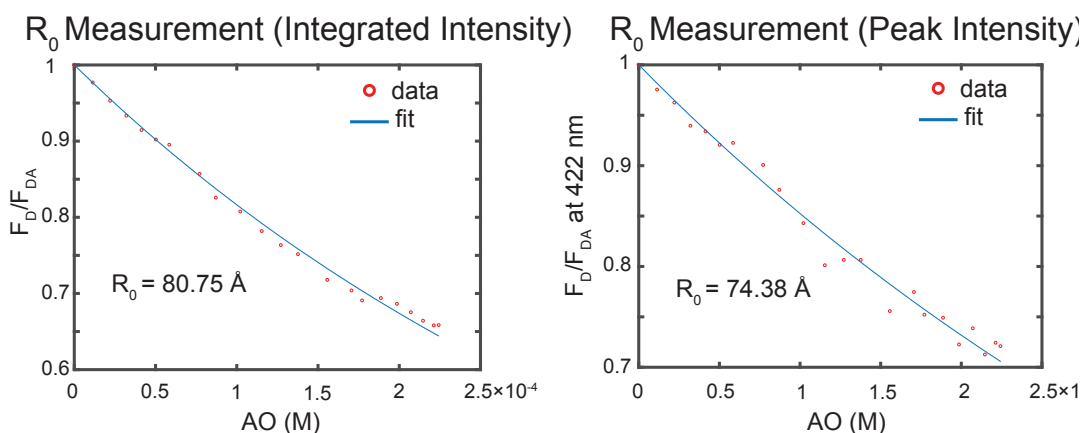


$$\frac{F_{DA}}{F_D} = 1 - \sqrt{\pi} \exp(\gamma^2) [1 - \text{erf}(\gamma)]$$

$$\text{erf}(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^\gamma \exp(-x^2) dx$$

$$\gamma = \frac{A}{A_0}$$

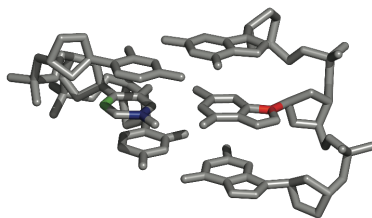
$$A_0(M) = 447 / R_0^3(\text{\AA})$$



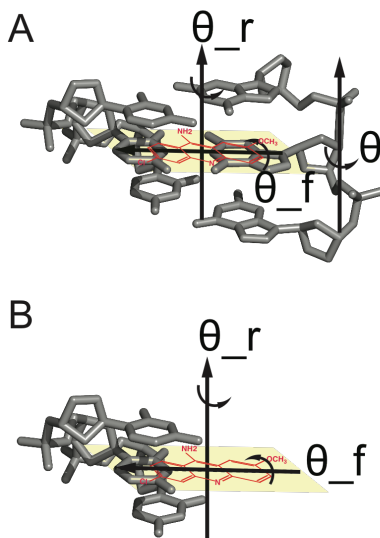
**Figure D.2:** Examples of Förster distance ( $R_0$ ) measurement fitting results. We used either integrated intensity (from 400-450nm) or peak intensity at 422nm to evaluate donor fluorescence intensity. The fluorescence titration results correspond to these two plots was shown in Figure 5.3 in main text.

### D.3 Circular Dichroism Calculation of ACMA Labeled Constructs

Experimental results are compatible with the hypothesis that ACMA ‘pushes’ the 6-MI directly across from its attachment site out of the duplex (see Figure 5.12 in Chapter V). Following this hypothesis, we have calculated CD of ACMA tethered DNA constructs. We applied the CD calculation presented in Chapter III and IV to the 6-MI and ACMA labeled DNA constructs. Assignments of the ACMA electric dipole transition moments are not available from the literature. We therefore derived its parameters based on ACMA’s similarity to quinacrine. Detailed descriptions can be found in Appendix B.



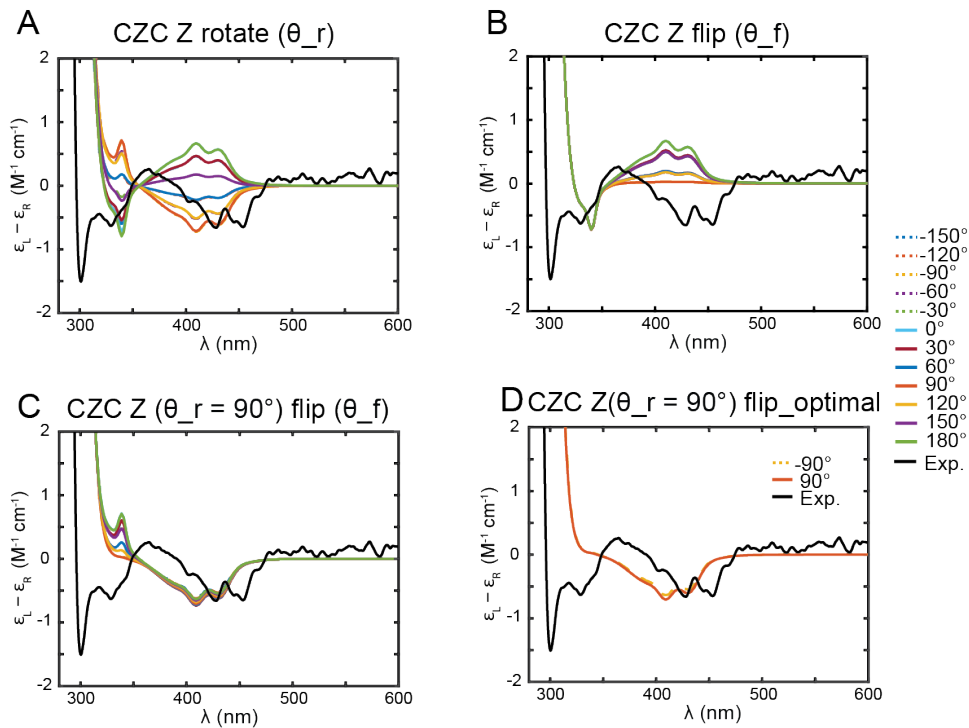
**Figure D.3:** The initial placement of ACMA on the CCC/GGG constructs.



**Figure D.4:** Angles corresponds to three conformational freedoms.

We began these calculations by placing ACMA in an arbitrary position, where ACMA is parallel to the base pairs. We used three ‘anchor’ atoms to place ACMA (see the three colored atoms in Figure D.3) on top of the pseudo base pair. Note that the base pair we used as the target to place AMCA does not exist in our DNA constructs. We next carried out CD calculation exploring three conformational freedoms, i.e. three angles that contribute to ACMA conformational change. These three angles ( $\theta_r$ ,  $\theta_f$  and  $\theta$ ) are illustrated in Figure D.4. We assumed that conformation of the 6-MI probe base was unaltered by the presence of the ACMA and retained the conformation presented in Chapter III. However, for the CZC/GYG construct when 6-MI is at the complementary position of ACMA, the 6-MI was shown to be ‘flipped out’.

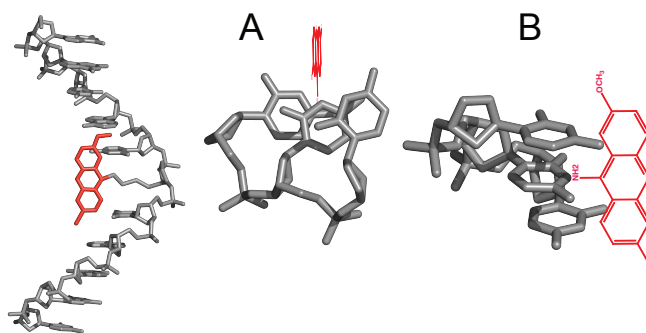
## CZC



**Figure D.5:** Comparison between experimental and simulated CD spectra of the ACMA tethered ss CZC construct as a function of ACMA rotation angles. **(A)** CD calculated results as a function of  $\theta_r$ . **(B)** CD calculated results as a function of  $\theta_f$ . **(C)** CD calculated results as a function of  $\theta_f$ , when  $\theta_r = 90^\circ$ . **(D)** Optimal results achieved.  $\theta_r = 90^\circ$ ,  $\theta_f = \pm 90^\circ$ .

Single stranded CZC experimental CD have a negative vibrational feature around 450 nm, a negative feature around 300 nm, and a slightly positive signal around 375 nm. We found that the calculated CD signal at 450 nm were greatly affected by ACMA rotation ( $\theta_r$ ) around the Z-axis, as shown in Figure D.5 A. The rotational ( $\theta_f$ ) motion of ACMA around the long-axis also affected the 450 nm feature but did not result in a negative peak as observed experimentally (see Figure D.5 B). With  $\theta_r = 90^\circ$ , an optimized feature at 450 nm was achieved, which compared favorably to experimental observations (see Figure D.5 C). With  $\theta_r = 90^\circ$ , the calculated CD are shown as a function as  $\theta_f$  (Figure D.5 C). The optimal results (shown in Figure D.5D and Figure

D.6) were achieved when ACMA sits orthogonal to the pseudo base pair. It is reasonable to assume ACMA in ssDNA has more flexibility than ACMA in dsDNA, with no constraints from the complementary bases. It is possible that the ss CZC construct has multiple ACMA conformational states that cause a negative net CD signal at ~300nm, which we were not able to model using only the two adjustable degrees of freedom.

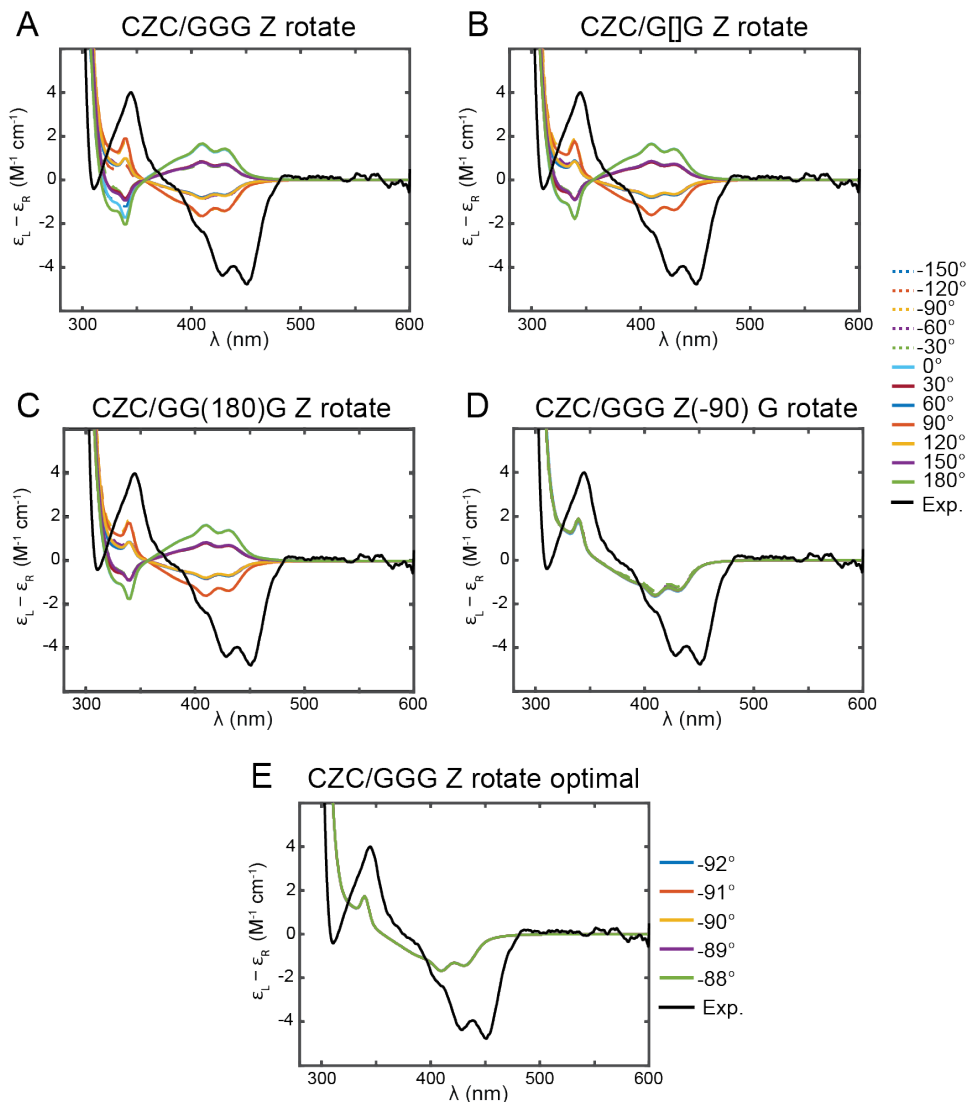


**Figure D.6:** The ss CZC conformation that produces optimal CD result, as shown in Figure D.5D, where  $\theta_r = 180^\circ$ ,  $\theta_f = \pm 90^\circ$ . **(A)** Top view of CZC from 3' to 5', only nearest neighbors of ACMA were shown. **(B)** Side view of CZC

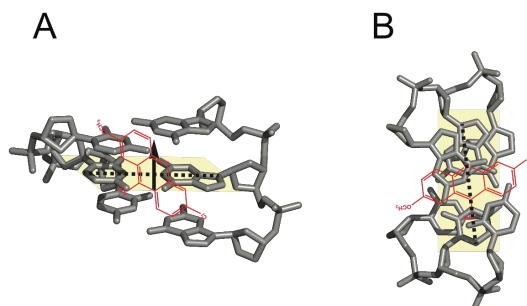
### ***CZC/GGG***

With the more constrained duplex DNA conformation, only  $\theta_r$  and  $\theta$  were allowed to change in this set of CD calculation involving CZC/GGG. Figure D.7 A shows the calculated CD results of the CZC/GGG construct when the B-form conformation was assumed for the canonical bases. We found that the calculated CD is highly dependent on the angle  $\theta_r$ . We also calculated the CD in the absence of the ACMA complementary Guanine in the same manner (see Figure D.7 B). Insignificant differences between Figure D.7 A and Figure D.7 B suggest that the presence of the complementary Guanine does not affect the calculated CD signal in the 280 nm – 600 nm region. We found that the calculated CD did not change significantly when the ACMA complementary Guanine rotated around the backbone (Figure D.7 D). The complementary Guanine of ACMA had

little effect on the net CD spectrum. Our best results were achieved when the ACMA was positioned perpendicular to the pseudo base pair, with  $\theta_r \approx 90^\circ$  (see Figure D.8).



**Figure D.7:** Comparison between experimental and simulated CD spectra of the ACMA tethered CZC/GGG construct as a function of ACMA rotation angles. **(A)** CD calculated results as a function of  $\theta_r$ . **(B)** CD calculated results as a function of  $\theta_r$ , in the absent of ACMA complementary guanine. **(C)** CD calculated results as a function of  $\theta_r$ , when  $\theta = 180^\circ$ . **(D)** CD calculated results as a function of  $\theta$ , when  $\theta_r = 90^\circ$ . **(E)** Optimal results achieved.  $\theta_r$  ranges from  $-88^\circ$  to  $-92^\circ$ ,  $\theta_f = 0^\circ$ ,  $\theta$  does not show strong correlation with CD calculated results.



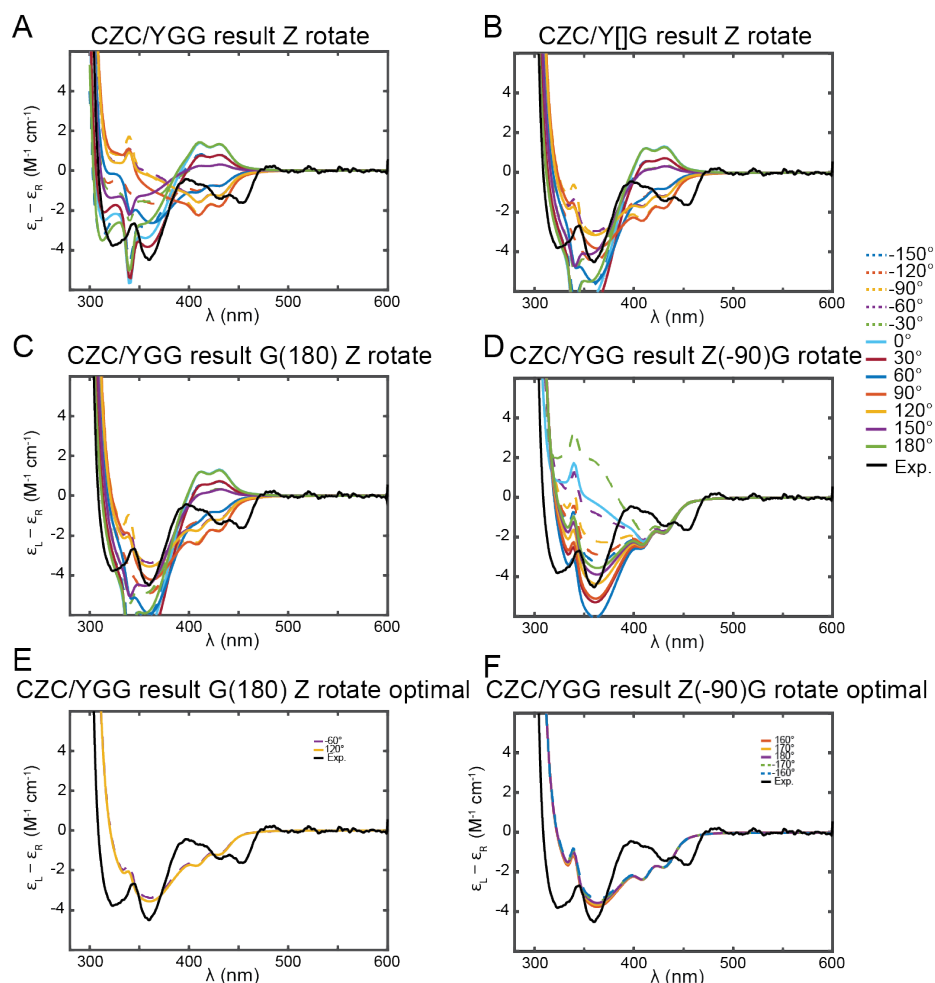
**Figure D.8:** The CZC/GGG conformation that produces optimal CD result, as shown in Figure D.7 E, where  $\theta_r = 90^\circ$ . **(A)** Side view of CZC/GGG **(B)** Top view of CZC/GGG from 3' to 5', only nearest neighbors of ACMA were shown.

### ***CZC/YGG***

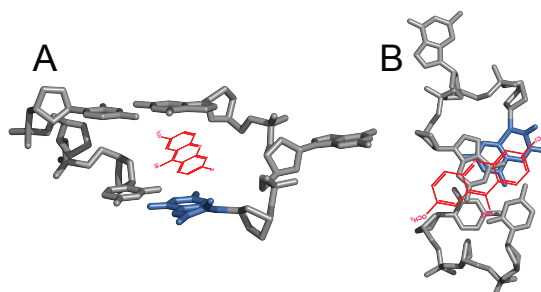
Unlike the CZC/GGG construct, the complementary Guanine has a large effect on the calculated CD of the CZC/YGG construct around 350 nm (see Figure D.9 A B, and D). The most favorable results were achieved when complementary Guanine is ‘flipped out’ (see Figure D.9 E, and F, and Figure D.10). Again, we discovered that ~450nm CD feature are mostly dependent on  $\theta_r$ .

### ***CZC/GYG***

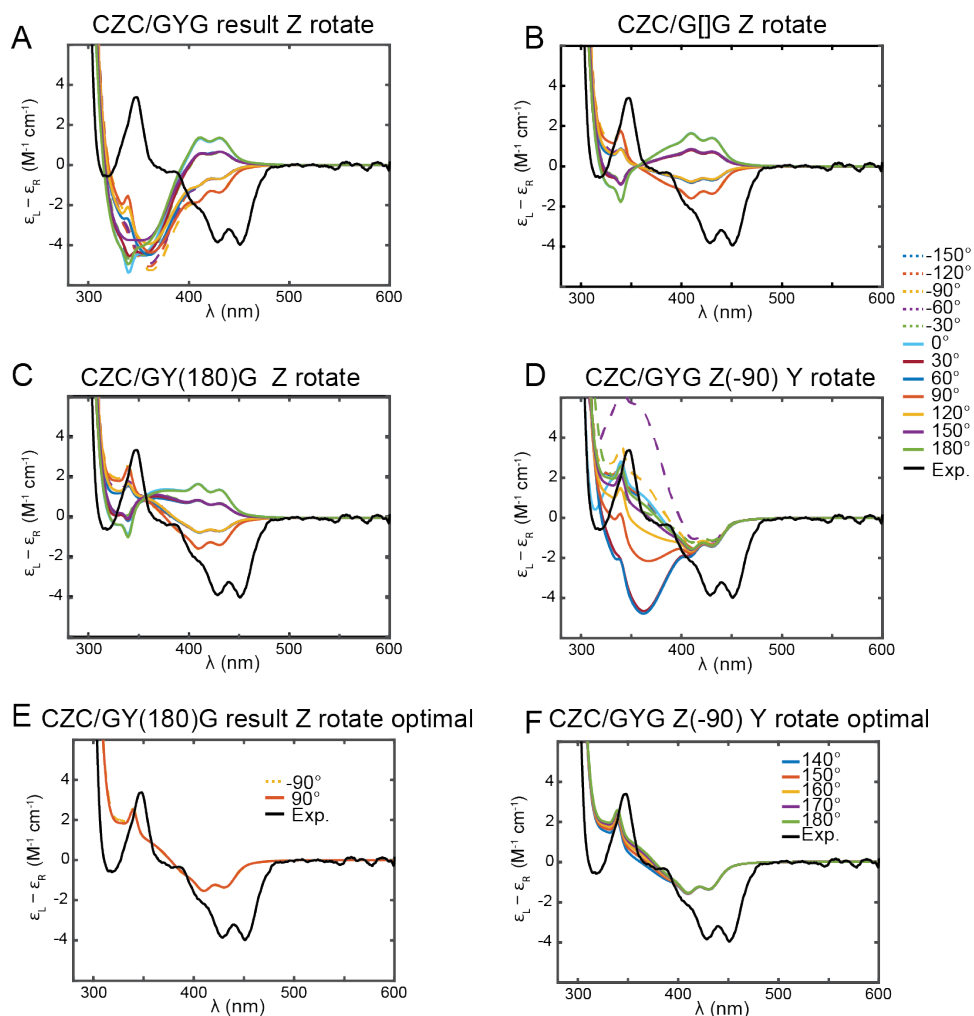
The CD features around 450 nm are predominantly dependent on ACMA conformation (see Figure D.11 A, B, C). 6-MI conformation ( $\theta$ ) changes CD significantly at 300-400 region (Figure D.11 D). Optimal results were achieved when complementary 6-MI is “flipped out” (see Figure D.11 E, and F, and Figure D.12).



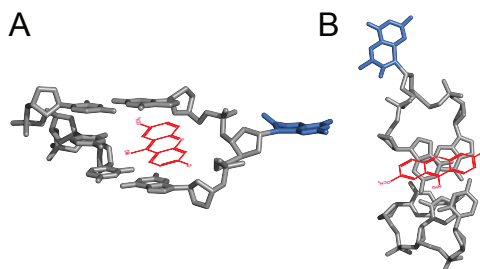
**Figure D.9:** Comparison between experimental and simulated CD spectra of the ACMA tethered CZC/YGG construct as a function of ACMA rotation angles. **(A)** CD calculated results as a function of  $\theta_r$ . **(B)** CD calculated results as a function of  $\theta_r$ , in the absent of ACMA complementary guanine. **(C)** CD calculated results as a function of  $\theta_r$ , when  $\theta = 180^\circ$ . **(D)** CD calculated results as a function of  $\theta$ , when  $\theta_r = 90^\circ$ . **(E)** Optimal results achieved when  $\theta_r$  ranges from  $-60^\circ$  or  $120^\circ$ ,  $\theta_f = 0^\circ$ ,  $\theta = 180^\circ$ . **(F)** Optimal results achieved when  $\theta_r = -90^\circ$ ,  $\theta_f = 0^\circ$ ,  $\theta = 160^\circ$ ,  $170^\circ$ ,  $180^\circ$ ,  $-170^\circ$  or  $-160^\circ$ .



**Figure D.10:** The CZC/YGG conformation that produces optimal CD result, as shown in Figure D.7 E, where  $\theta_r = 120^\circ$  ( $-60^\circ$ ),  $\theta_f = 0^\circ$ ,  $\theta = 180^\circ$ . **(A)** Side view of CZC/YGG **(B)** Top view of CZC/YGG from 3' to 5', only nearest neighbors of ACMA were shown.



**Figure D.11:** Comparison between experimental and simulated CD spectra of the ACMA tethered CZC/GYG construct as a function of ACMA rotation angles. **(A)** CD calculated results as a function of  $\theta_r$ . **(B)** CD calculated results as a function of  $\theta_r$ , in the absent of ACMA complementary 6-MI. **(C)** CD calculated results as a function of  $\theta_r$ , when  $\theta = 180^\circ$ . **(D)** CD calculated results as a function of  $\theta$ , when  $\theta_r = 90^\circ$ . **(E)** Optimal results achieved when  $\theta_r = \pm 90^\circ$ ,  $\theta_f = 0^\circ$ ,  $\theta = 180^\circ$ . **(F)** Optimal results achieved when  $\theta_r = -90^\circ$ ,  $\theta_f = 0^\circ$ ,  $\theta = 140^\circ, 150^\circ, 160^\circ, 170^\circ, 180^\circ$ .

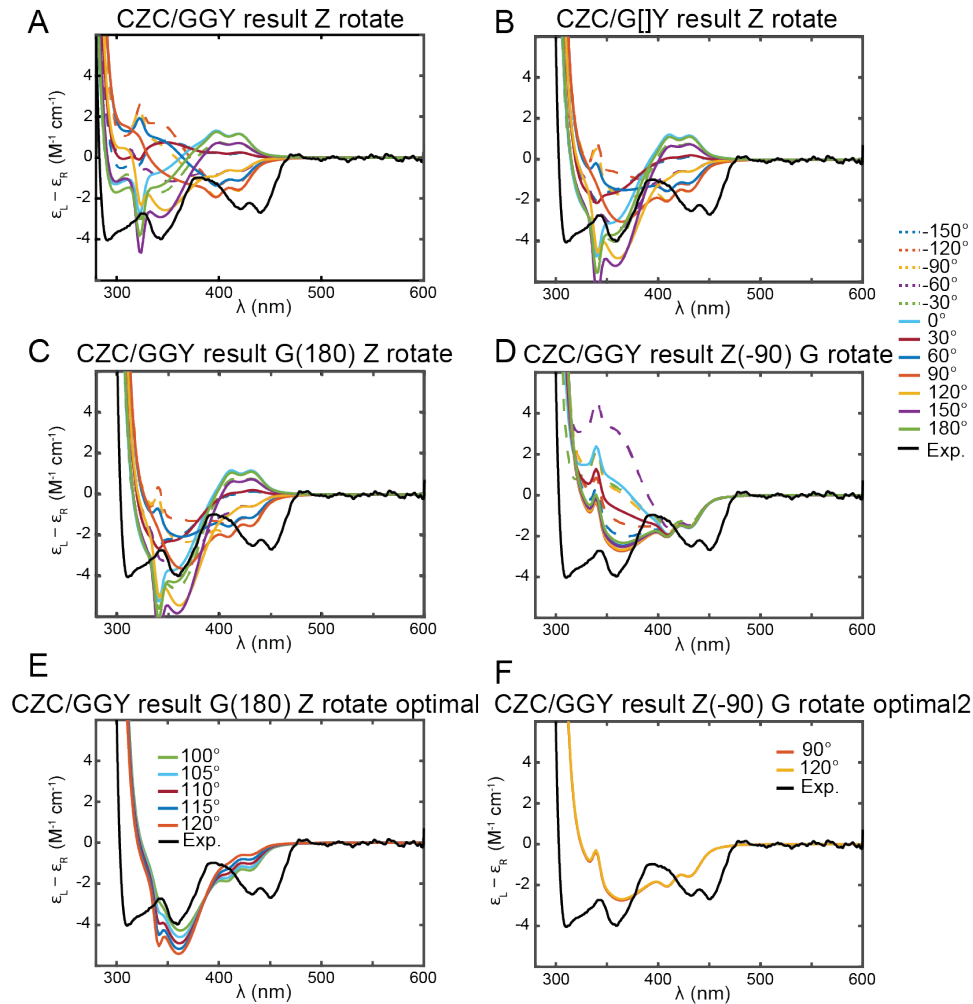


**Figure D.12:** The CZC/GYG conformation that produces optimal CD result, as shown in Figure D.11 E, where  $\theta_r = 90^\circ$  ( $-90^\circ$ ),  $\theta_f = 0^\circ$ ,  $\theta = 180^\circ$ . **(A)** Side view of CZC/GYG **(B)** Top view of CZC/GYG from 3' to 5', only nearest neighbors of ACMA were shown.

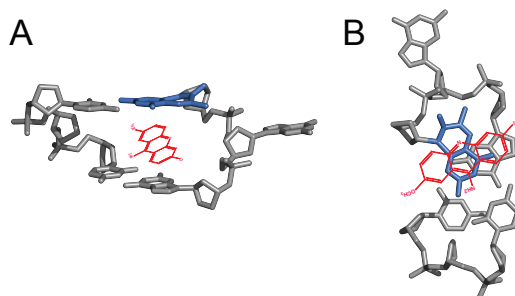


## CZC/GGY

Note the similarity of experimental CD between CZC/GGY and CZC/YGG. The optimal orientation of ACMA are also similar. The sharp peak at around 300 nm is caught by the calculated CD, when  $\theta_r$  is about 120 (-60°) (see Figure D.13 E). All ACMA EDTM orientation are either along the short axis or long axis of ACMA. It is impossible to distinguish of conformation at  $\theta_r=120^\circ$  or  $-60^\circ$ , by CD calculation alone.



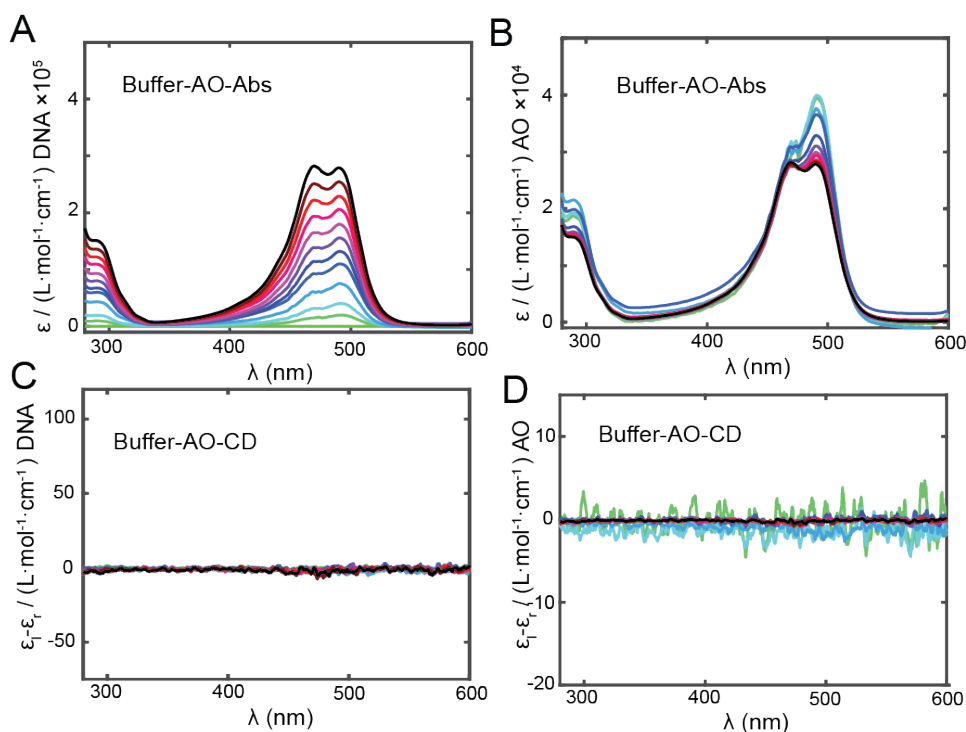
**Figure D.13:** Comparison between experimental and simulated CD spectra of the ACMA tethered CZC/GGY construct as a function of ACMA rotation angles. **(A)** CD calculated results as a function of  $\theta_r$ . **(B)** CD calculated results as a function of  $\theta_r$ , in the absent of ACMA complementary guanine. **(C)** CD calculated results as a function of  $\theta_r$ , when  $\theta = 180^\circ$ . **(D)** CD calculated results as a function of  $\theta$ , when  $\theta_r = 90^\circ$ . **(E)** Optimal results achieved when  $\theta_r$  ranges from  $100^\circ$  to  $120^\circ$ ,  $\theta_f = 0^\circ$ ,  $\theta = 180^\circ$ . **(F)** Optimal results achieved when  $\theta_r = -90^\circ$ ,  $\theta_f = 0^\circ$ ,  $\theta = 90^\circ$ ,  $120^\circ$ .



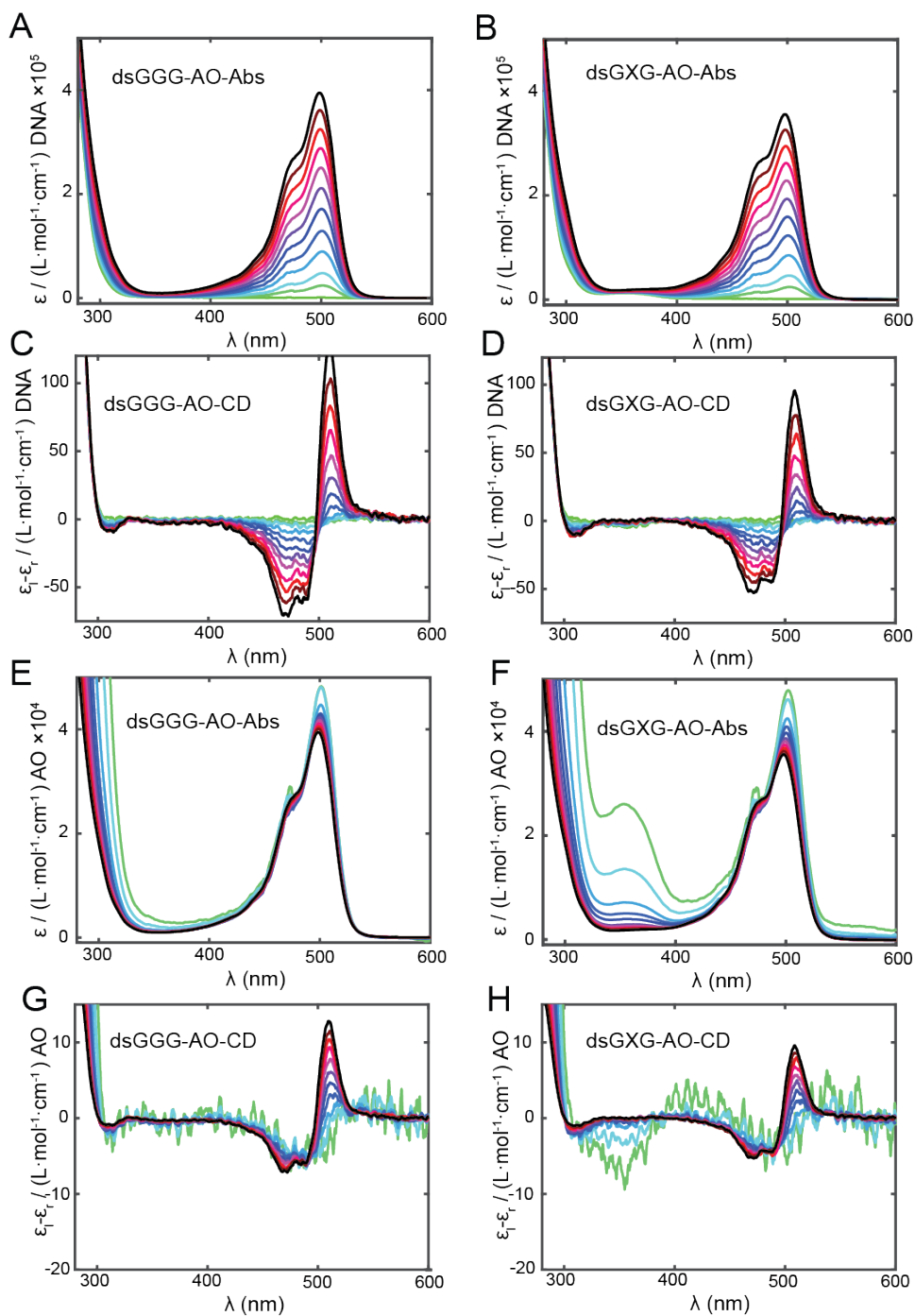
**Figure D.14:** The CZC/GGY conformation that produces optimal CD result, as shown in Figure D.13E, where  $\theta_r = 110^\circ$  ( $-70^\circ$ ),  $\theta_f = 0^\circ$ ,  $\theta = 180^\circ$ . **(A)** Side view of CZC/GGY **(B)** Top view of CZC/GGY from 3' to 5', only nearest neighbors of ACMA were shown.

#### D.4 AO-DNA Absorbance and CD Titration

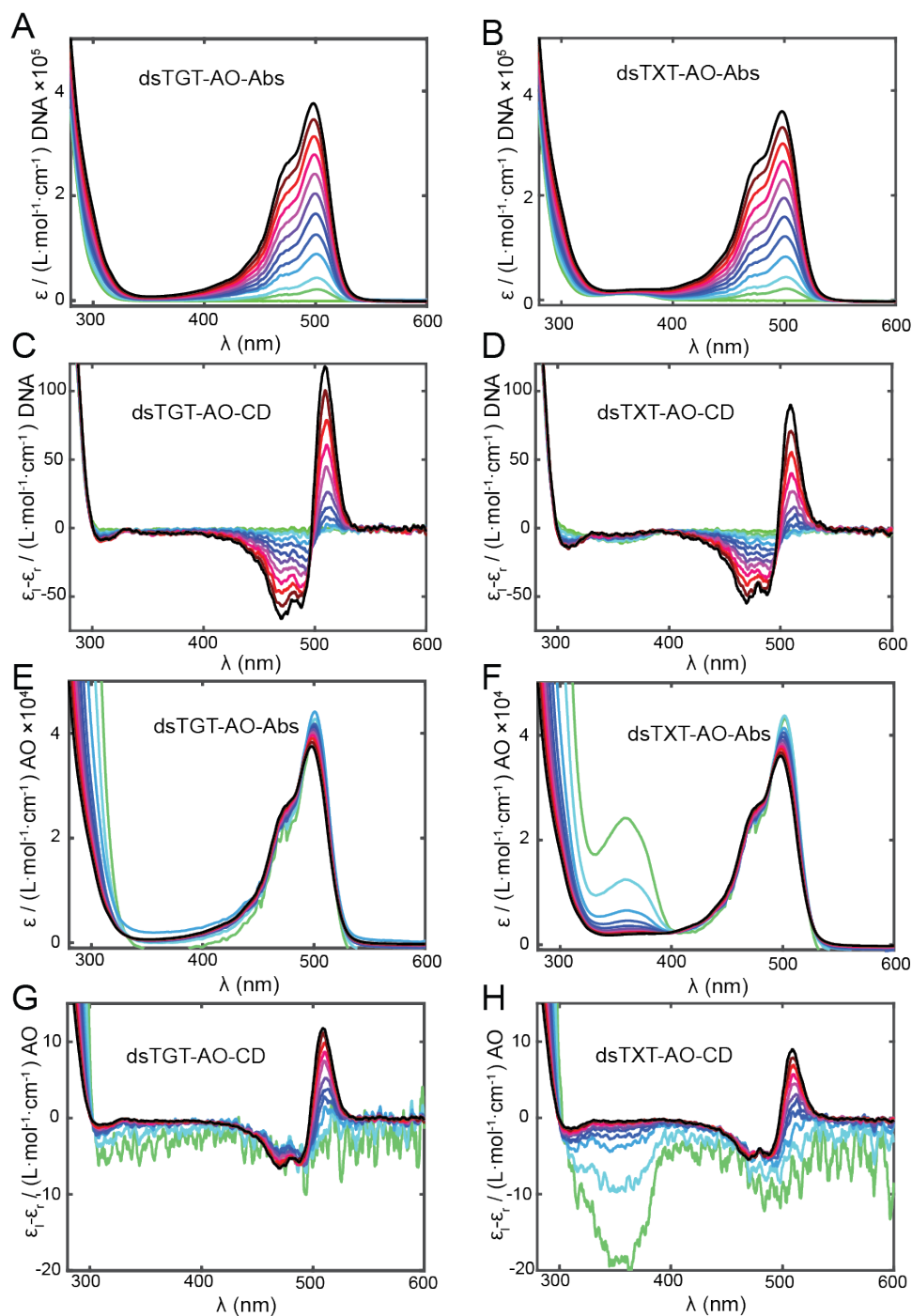
We conducted titration experiments where we add AO to different DNA samples. Concentrations of AO and DNA in solution are indicated by the line color (see Table D.2).



**Figure D.15:** Buffer-AO titration experiments **(A)** Buffer-AO absorbance spectra per M DNA. **(B)** Buffer-AO absorbance spectra per M AO (except when AO concentration was zero). **(C)** Buffer-AO CD spectra per M DNA. **(D)** Buffer-AO CD spectra per M AO (except when AO concentration was zero).















**Figure D.16:** dsGG(X)G-AO titration experiments (A) dsGGG-AO absorbance spectra per M DNA. (B) dsGXG-AO absorbance spectra per M DNA. (C) dsGGG-AO CD spectra per M DNA. (D) dsGXG-AO CD spectra per M DNA. (E) dsGGG-AO absorbance spectra per M AO (except when AO concentration was zero). (F) dsGXG-AO absorbance spectra per M AO (except when AO concentration was zero). (G) dsGGG-AO CD spectra per M AO (except when AO concentration was zero). (H) dsGXG-AO CD spectra per M AO (except when AO concentration was zero).

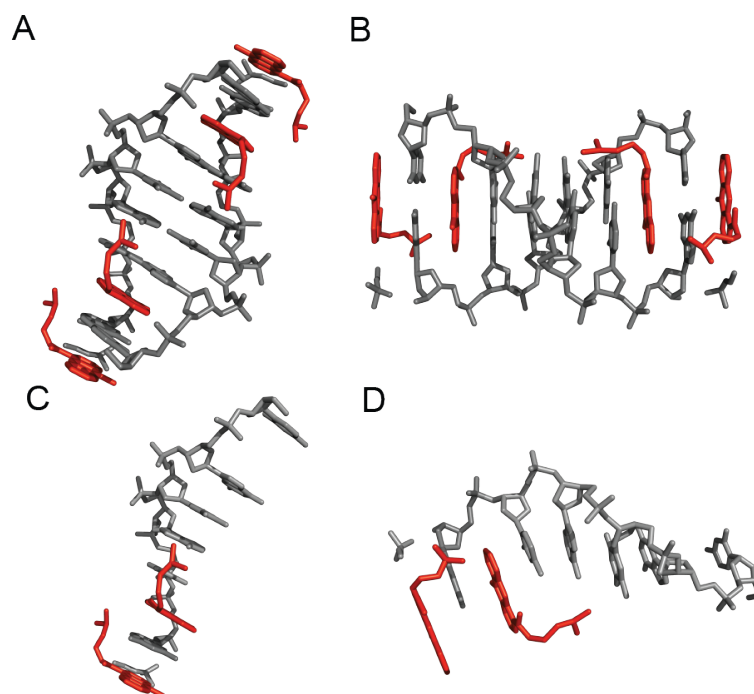


**Figure D.17:** dsTG(X)T-AO titration experiments **(A)** dsTGT-AO absorbance spectra per M DNA. **(B)** dsTXT-AO absorbance spectra per M DNA. **(C)** dsTGT-AO CD spectra per M DNA. **(D)** dsTXT-AO CD spectra per M DNA. **(E)** dsTGT-AO absorbance spectra per M AO (except when AO concentration was zero). **(F)** dsTXT-AO absorbance spectra per M AO (except when AO concentration was zero). **(G)** dsTGT-AO CD spectra per M AO (except when AO concentration was zero). **(H)** dsTXT-AO CD spectra per M AO (except when AO concentration was zero).

**Table D.2.** Concentration of Acridine Orange, dsDNA and the ratio of their concentrations in absorbance and CD titration experiments shown in Figure D.15, Figure D.16 and Figure D.17.

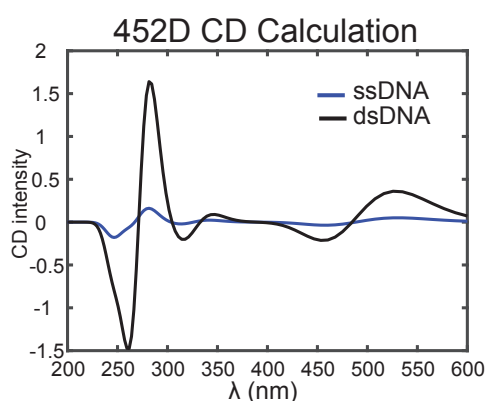
Spectrum Color	[AO] ( $\mu\text{M}$ )	[DNA] ( $\mu\text{M}$ )	[AO]/[DNA]
	0	4.00	0
	1.99	3.97	0.5
	3.95	3.95	1
	7.80	3.90	2
	11.54	3.85	3
	15.19	3.80	4
	18.75	3.75	5
	22.22	3.70	6
	25.61	3.66	7
	28.92	3.61	8
	32.14	3.57	9
	35.29	3.53	10

#### D.5 AO-DNA CD Calculation



**Figure D.18:** X-ray structure of 452D, d (CG (5-BrU) ACG)<sub>2</sub> crystalized with 4 DACA intercalating molecules. DACA molecules are colored red. (A) Duplex side view (B) Duplex oriented view (C) Single stranded side view (D) Single stranded oriented view

Crystal structure of AO-DNA intercalating complex is not available on PDB. We found the DACA X-ray structure complexed to d(CG(5-BrU)ACG)<sub>2</sub> at 1.3-Å resolution<sup>4</sup>. DACA is an Acridine derivative. We calculated CD of this construct, where we substitute DACA with AO, guanine with 6-MI. We are able to get CD calculated results that resemble the experimental result shown in Figure D.16 and D.17. Our preliminary results are shown in Fig D.19. It might be feasible to use CD calculation method to gain AO-DNA complex conformational information.



**Figure D.19:** AO-DNA CD calculation results based on the structure of 452D.

### ***D.6 ACMA Sequence Dependent Fluorescence and CD***

In order to study the interaction between 6-MI and ACMA in duplex DNA we need information about DNA constructs labeled with ACMA alone. Here we examined the fluorescence and CD of tethered ACMA DNA constructs shown in Table D.3. It has been previously claimed that tethered ACMA inserts between bases in ssDNA, primarily based on evidence from fluorescence measurements. Hypochromism and UV line broadening can be larger for tethered ACMA in ssDNA than in dsDNA, which has been interpreted as meaning that tethered ACMA stacks more in single stranded construct<sup>5</sup>. Furthermore, line broadening may indicate that the ACMA moiety has greater

conformational inhomogeneity in ssDNA compared to dsDNA as a result of the greater flexibility of ssDNA.<sup>5</sup>

**Table D.3.** Base sequences and nomenclature for the ACMA-containing ssDNA constructs used in these studies. The letter **Z** indicates the ACMA probe.

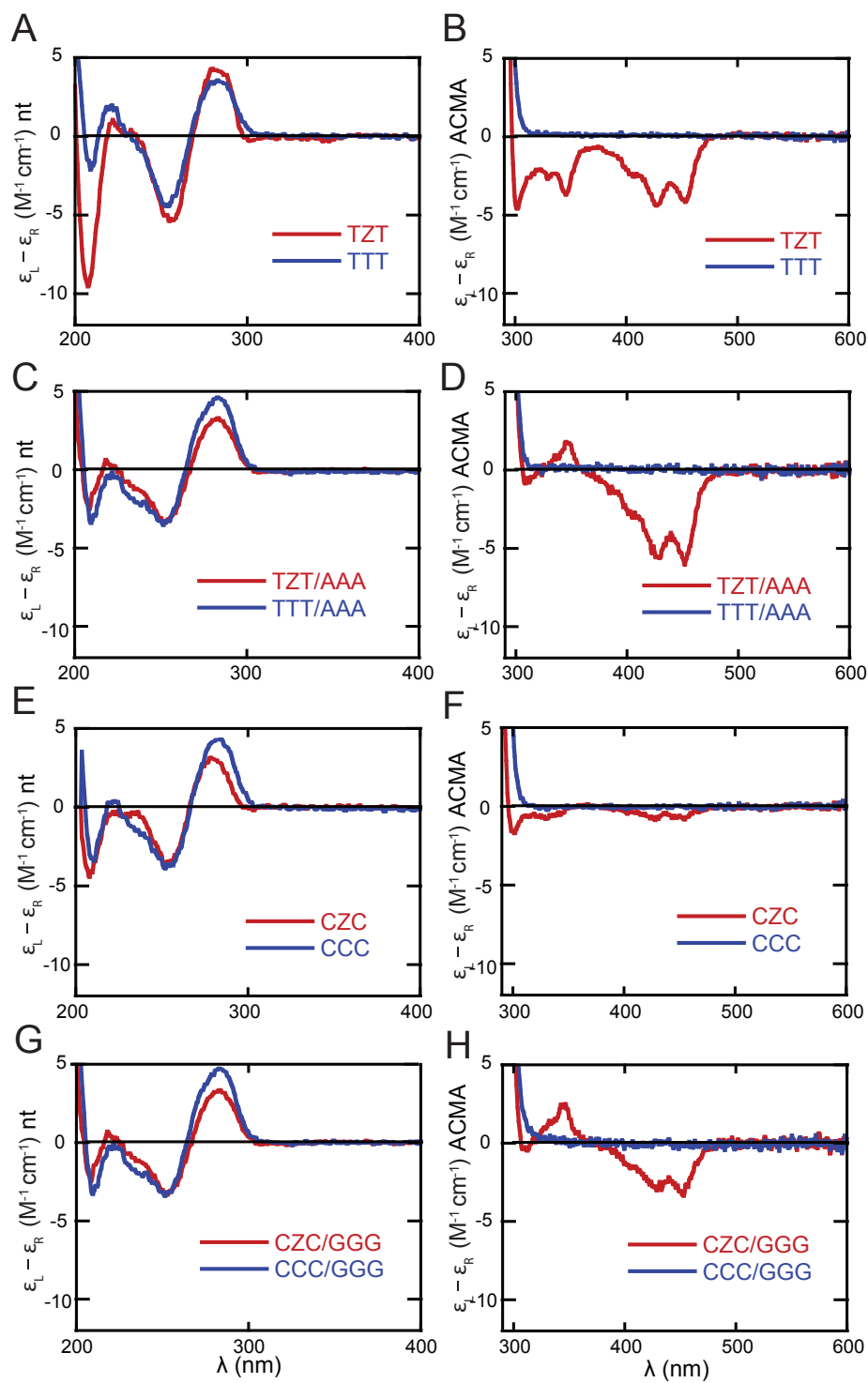
<b>DNA construct</b>	<b>Nucleotide base sequence</b>
<b>TZT</b>	3'-C GCA <b>TZT</b> ACG G-5'
<b>CZC</b>	3'-C GCA <b>CZC</b> ACG G-5'
<b>AZA</b>	3'-C GCA <b>AZA</b> ACG G-5'
<b>GZG</b>	3'-C GCA <b>GZG</b> ACG G-5'

The CD spectra of the ACMA-labeled ssDNA are shown in Figure D.20. These spectra exhibit bands from two electronic transitions at 450 and 350 nm, together with the vibrational progressions of these transitions (see Figure 5.5 D). The planar ACMA molecule is achiral and has no intrinsic optical activity. The CD spectra of the tethered ACMA-ssDNA molecule indicate that the chromophore experiences the asymmetric environment of ssDNA, consistent with the idea of stacking interactions between the chromophore and the DNA bases.

In general, the CD signals from of ACMA are relatively weak; the intensity per ACMA is about 50% the intensity per nucleotide (compare left and right panels of the CD spectra, Figure D.20) with the exception of the TZT spectrum. The TZT oligonucleotide has complementary terminal sequences that could potentially pair with each other to form a stem loop structure and conceivably alter the CD spectrum.

Base stacking interactions stabilize a B form conformation within the persistence length of ssDNA<sup>6</sup>. Base stacking with T is weaker than for other DNA bases. Thus, flanking T residues may allow more flexibility at the ACMA attachment site, leading to a qualitatively different interaction of the TXT oligonucleotide with ACMA in which the

intercalant exhibits a binding mode within the ss molecule with more stacking and a larger CD signal.



**Figure D.20:** Sequence-dependent ACMA CD.



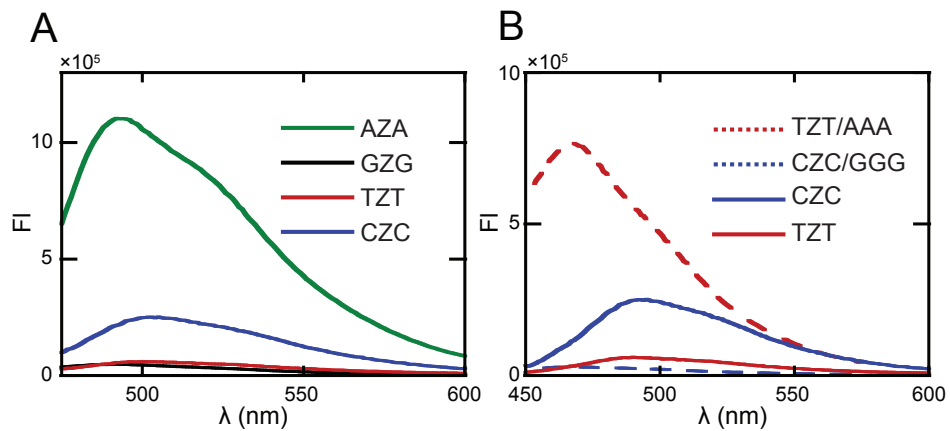
Formation of double stranded complexes TZT/AAA and CZC/GGG increases the intensity of the CD signal near 450 nm in the duplex compared to the ss construct. This increase is much greater for CZC/GGG than for TZT/AAA, in comparison to the CD of their respective ss components. This result is consistent with the idea that ACMA is already more stacked in the ss TZT construct than the ss CZC construct. In addition, the negative CD signal associated with the 350 nm transition of ssDNA changes sign in the duplex molecules.

Intercalation of ACMA adjacent to an Adenine can increase its fluorescence by up to a factor of 4 compared to other base; this is the physical basis of selective staining of A-T rich chromosomes by quinacrine. Fluorescence emission of tethered ACMA is also enhanced in DNA consisting of only A and T bases.<sup>5,7</sup> QY of tethered ACMA in an A rich environment is 0.4 – 0.6 (compared to a QY of 0.8 for free ACMA)<sup>5</sup>.

In contrast ACMA is quenched by stacking with Guanine<sup>9-11</sup>. The excited state lifetime of ACMA decreases considerably in the presence of guanine. Quenching is thought to occur through electron transfer (ET).<sup>10,12</sup> Guanine is the most oxidizable base and ACMA can act as an electron donor. Each successive A-T placed between tethered ACMA and G decreases ET by a factor of  $10^3$  (G, A and ACMA are on the same strand of DNA in a duplex molecule).<sup>11</sup>

We have also observed this effect of A and G on the fluorescence emission of tethered ACMA in ssDNA; fluorescence emission from ACMA in single stranded AZA is greater than the other sequence contexts while signal from GZG is lowest (Figure D.21 A). In addition, we observed similar effects of proximal A and G bases in the complementary strand; the fluorescence of TZT/AAA increases compared to TZT while

the fluorescence of ACMA decreased in the CZC/GGG duplex compared to CZC (Figure D.21 B).



**Figure D.21:** Sequence dependent ACMA fluorescence intensity. (A) Fluorescence emission spectra of single-stranded AZA, GZG, TZT and CZC excited at 450 nm. (B) Fluorescence emission spectra of single-stranded CZC, TZT and double-stranded CZC/GGG and TZT/AAA excited at 450 nm.

## REFERENCES CITED

### CHAPTER I

- <sup>1</sup> P.H. von Hippel, N.P. Johnson, and A.H. Marcus, "Fifty years of DNA 'Breathing': Reflections on old and new approaches," *Biopolymers* **99**, 923 (2013).
- <sup>2</sup> D. Jose, K. Datta, N.P. Johnson, and P.H. von Hippel, "Spectroscopic studies of position-specific DNA 'breathing' fluctuations at replication forks and primer-template junctions.," *Proc. Natl. Acad. Sci. U. S. A.* **106**, 4231 (2009).
- <sup>3</sup> D. Jose, S.E. Weitzel, and P.H. von Hippel, "Breathing fluctuations in position-specific DNA base pairs are involved in regulating helicase movement into the replication fork," *Proc. Natl. Acad. Sci.* **109**, 14428 (2012).
- <sup>4</sup> C. Phelps, W. Lee, D. Jose, P.H. von Hippel, and A.H. Marcus, "Single-molecule FRET and linear dichroism studies of DNA breathing and helicase binding at replication fork junctions," *Proc. Natl. Acad. Sci.* **110**, 17320 (2013).
- <sup>5</sup> K. Datta, N.P. Johnson, G. Villani, A.H. Marcus, and P.H. Von Hippel, "Characterization of the 6-methyl isoxanthopterin (6-MI) base analog dimer, a spectroscopic probe for monitoring guanine base conformations at specific sites in nucleic acids," *Nucleic Acids Res.* **40**, 1191 (2012).
- <sup>6</sup> K. Datta, N.P. Johnson, and P.H. von Hippel, "DNA conformational changes at the primer-template junction regulate the fidelity of replication by DNA polymerase.," *Proc. Natl. Acad. Sci. U. S. A.* **107**, 17980 (2010).
- <sup>7</sup> N.P. Johnson, H. Ji, T.H. Steinberg, P.H. von Hippel, and A.H. Marcus, "Sequence-Dependent Conformational Heterogeneity and Proton-Transfer Reactivity of the Fluorescent Guanine Analogue 6-Methyl Isoxanthopterin (6-MI) in DNA," *J. Phys. Chem. B* **119**, 12798 (2015).
- <sup>8</sup> J.H. Han, N. Chitrapriya, H.S. Lee, Y.-A. Lee, S.K. Kim, and M.-J. Jung, "Behavior of the Guanine Base in G-quadruplexes Probed by the Fluorescent Guanine Analog, 6-Methyl Isoxanthopterin," *Bull. Korean Chem. Soc.* **38**, 183 (2017).
- <sup>9</sup> E.B. Nielsen and J.A. Schellman, "Rotatory properties of molecules containing two peptide groups: experimental and theoretical results," *Biopolymers* **10**, 1559 (1971).
- <sup>10</sup> V. Madison and J. Schellman, "Optical activity of polypeptides and proteins," *Biopolymers* **11**, 1041 (1972).
- <sup>11</sup> V. Rizzo and J.A. Schellman, "Matrix-method calculation of linear and circular dichroism spectra of nucleic acids and polynucleotides," *Biopolymers* **23**, 435 (1984).

<sup>12</sup> W.C. Johnson and I. Tinoco, "Circular dichroism of polynucleotides: A simple theory," *Biopolymers* **7**, 727 (1969).

<sup>13</sup> A.L. Williams Jr, C. Cheong, I. Tinico Jr, and L.B. Clark, "Vacuum ultraviolet circular dichroism as an indicator of helical handedness in nucleic acids," *Nucleic Acids Res.* **14**, 6649 (1986).

<sup>14</sup> B. Nordén, A. Rodger and T. Dafforn, *Linear Dichroism and Circular Dichroism*, (Royal Society of Chemistry, 2010).

## CHAPTER II

<sup>1</sup> G. Altan-Bonnet, A. Libchaber, and O. Krichevsky, "Bubble Dynamics in Double-Stranded DNA," *Phys. Rev. Lett.* **90**, 4 (2003).

<sup>2</sup> W. Lee, D. Jose, C. Phelps, A.H. Marcus, and P.H. Von Hippel, "A single-molecule view of the assembly pathway, subunit stoichiometry, and unwinding activity of the bacteriophage T4 primosome (helicase-primase) complex," *Biochemistry* **52**, 3157 (2013).

<sup>3</sup> C. Phelps, W. Lee, D. Jose, P.H. von Hippel, and A.H. Marcus, "Single-molecule FRET and linear dichroism studies of DNA breathing and helicase binding at replication fork junctions," *Proc. Natl. Acad. Sci.* **110**, 17320 (2013).

<sup>4</sup> N.P. Johnson, W.A. Baase, and P.H. Von Hippel, "Low-energy circular dichroism of 2-aminopurine dinucleotide as a probe of local conformation of DNA and RNA," **101**, (2004).

<sup>5</sup> N.P. Johnson, W.A. Baase, and P.H. Von Hippel, "Investigating local conformations of double-stranded DNA by low-energy circular dichroism of pyrrolo-cytosine," **1** (2005).

<sup>6</sup> K. Datta, N.P. Johnson, and P.H. Von Hippel, "Mapping the Conformation of the Nucleic Acid Framework of the T7 RNA Polymerase Elongation Complex in Solution Using Low-energy CD and Fluorescence Spectroscopy," **800** (2006).

<sup>7</sup> K. Datta and P.H. von Hippel, "Direct spectroscopic study of reconstituted transcription complexes reveals that intrinsic termination is driven primarily by thermodynamic destabilization of the nucleic acid framework," *J. Biol. Chem.* **283**, 3537 (2008).

<sup>8</sup> D. Jose, K. Datta, N.P. Johnson, and P.H. von Hippel, "Spectroscopic studies of position-specific DNA 'breathing' fluctuations at replication forks and primer-template junctions," *Proc. Natl. Acad. Sci.* **106**, 4231 (2009).

- <sup>9</sup> K. Datta, N.P. Johnson, and P.H. von Hippel, "DNA conformational changes at the primer-template junction regulate the fidelity of replication by DNA polymerase.," Proc. Natl. Acad. Sci. U. S. A. **107**, 17980 (2010).
- <sup>10</sup> D. Jose, S.E. Weitzel, D. Jing, and P.H. Von Hippel, "Assembly and subunit stoichiometry of the functional of bacteriophage T4," Pnas (2012).
- <sup>11</sup> D. Jose, S.E. Weitzel, and P.H. von Hippel, "Breathing fluctuations in position-specific DNA base pairs are involved in regulating helicase movement into the replication fork," Proc. Natl. Acad. Sci. **109**, 14428 (2012).
- <sup>12</sup> A.P. Demchenko, "The red-edge effects: 30 years of exploration," Luminescence **17**, 19 (2002).
- <sup>13</sup> S. Mukamel, *Principles of Nonlinear Optical Spectroscopy* (Oxford university press New York, 1995).
- <sup>14</sup> A. Nitzan, *Chemical Dynamics in Condensed Phases: Relaxation, Transfer and Reactions in Condensed Molecular Systems (Oxford Graduate Texts)* (Oxford University Press USA, 2006).
- <sup>15</sup> A.D. Stein and M.D. Fayer, "Nanosecond timescale optical inhomogeneous broadening of dye molecules in liquids at and near room temperature," Chem. Phys. Lett. **176**, 159 (1991).
- <sup>16</sup> H. Baessler and B. Schweitzer, "Site-selective fluorescence spectroscopy of conjugated polymers and oligomers," Acc. Chem. Res. **32**, 173 (1999).
- <sup>17</sup> W.C. Galley and R.M. Purkey, "Role of heterogeneity of the solvation site in electronic spectra in solution.," Proc. Natl. Acad. Sci. U. S. A. **67**, 1116 (1970).
- <sup>18</sup> G. Weber and M. Shinitzky, "Failure of energy transfer between identical aromatic molecules on excitation at the long wave edge of the absorption spectrum," Proc. Natl. Acad. Sci. **65**, 823 (1970).
- <sup>19</sup> A. Chattopadhyay and S. Mukherjee, "Red edge excitation shift of a deeply embedded membrane probe: implications in water penetration in the bilayer," J. Phys. Chem. B **103**, 8180 (1999).
- <sup>20</sup> J.R. Lakowicz and S. Keating-Nakamoto, "Red-edge excitation of fluorescence and dynamic properties of proteins and membranes.," Biochemistry **23**, 3013 (1984).
- <sup>21</sup> A.P. Demchenko, "On the nanosecond mobility in proteins: edge excitation fluorescence red shift of protein-bound 2-(p-toluidinylnaphthalene)-6-sulfonate," Biophys. Chem. **15**, 101 (1982).
- <sup>22</sup> M.E. Hawkins, W. Pfleiderer, F.M. Balis, D. Porter, and J.R. Knutson, "Fluorescence properties of pteridine nucleoside analogs as monomers and incorporated into oligonucleotides.," Anal. Biochem. **244**, 86 (1997).

- <sup>23</sup> M.E. Hawkins, "Fluorescent pteridine nucleoside analogs," *Cell Biochem. Biophys.* **34**, 257 (2001).
- <sup>24</sup> K. Wojtuszewski Poulin, A. V Smirnov, M.E. Hawkins, F.M. Balis, and J.R. Knutson, "Conformational heterogeneity and quasi-static self-quenching in DNA containing a fluorescent guanine analogue, 3MI or 6MI," *Biochemistry* **48**, 8861 (2009).
- <sup>25</sup> L.M. Wilhelmsson, "Fluorescent nucleic acid base analogues.," *Q. Rev. Biophys.* **43**, 159 (2010).
- <sup>26</sup> E. Seibert, A.S. Chin, W. Pfliederer, M.E. Hawkins, W.R. Laws, R. Osman, and J.B.A. Ross, "pH-Dependent Spectroscopy and Electronic Structure of the Guanine Analogue 6,8-Dimethylisoxanthopterin †," *J. Phys. Chem. A* **107**, 178 (2003).
- <sup>27</sup> K. Datta, N.P. Johnson, G. Villani, A.H. Marcus, and P.H. Von Hippel, "Characterization of the 6-methyl isoxanthopterin (6-MI) base analog dimer, a spectroscopic probe for monitoring guanine base conformations at specific sites in nucleic acids," *Nucleic Acids Res.* **40**, 1191 (2012).
- <sup>28</sup> J.R. Widom, D. Rappoport, A. Perdomo-Ortiz, H. Thomsen, N.P. Johnson, P.H. von Hippel, A. Aspuru-Guzik, and A.H. Marcus, "Electronic transition moments of 6-methyl isoxanthopterin--a fluorescent analogue of the nucleic acid base guanine.," *Nucleic Acids Res.* **41**, 995 (2013).
- <sup>29</sup> N.J. Turro, V. Ramamurthy, V. Ramamurthy, and J.C. Scaiano, *Principles of Molecular Photochemistry: An Introduction* (University science books, 2009).
- <sup>30</sup> J.R. Lakowicz, *Principles of Fluorescence Spectroscopy*, (Springer, 2006).
- <sup>31</sup> V. Bloomfield and D.M. Crothers, *Nucleic Acids: Structures, Properties and Functions* (2000).
- <sup>32</sup> R.A. Tinsley and N.G. Walter, "Pyrrolo-C as a fluorescent probe for monitoring RNA secondary structure formation," *Rna* **12**, 522 (2006).
- <sup>33</sup> Z. Wu, A. Ono, M. Kainosho, and A. Bax, "H... N hydrogen bond lengths in double stranded DNA from internucleotide dipolar couplings," *J. Biomol. NMR* **19**, 361 (2001).
- <sup>34</sup> M. Nadal-Ferret, R. Gelabert, M. Moreno, and J.M. Lluch, "Are there really low-barrier hydrogen bonds in proteins? The case of photoactive yellow protein," *J. Am. Chem. Soc.* **136**, 3542 (2014).
- <sup>35</sup> A. Warshel and A. Papazyan, "Energy considerations show that low-barrier hydrogen bonds do not offer a catalytic advantage over ordinary hydrogen bonds," *Proc. Natl. Acad. Sci.* **93**, 13665 (1996).
- <sup>36</sup> W.W. Cleland, "Low-barrier hydrogen bonds and low fractionation factor bases in enzymic reactions," *Biochemistry* **31**, 317 (1992).

- <sup>37</sup> S.J. Grabowski, "Theoretical studies of strong hydrogen bonds," *Annu. Reports Sect. C* (Physical Chem. **102**, 131 (2006)).
- <sup>38</sup> K. Saito and H. Ishikita, "Energetics of short hydrogen bonds in photoactive yellow protein," *Proc. Natl. Acad. Sci.* **109**, 167 (2012).
- <sup>39</sup> E. Arunan, G.R. Desiraju, R.A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D.C. Clary, R.H. Crabtree, J.J. Dannenberg, and P. Hobza, "Defining the hydrogen bond: An account (IUPAC Technical Report)," *Pure Appl. Chem.* **83**, 1619 (2011).
- <sup>40</sup> H.A. Sober and R.A. Harte, *CRC Handbook of Biochemistry*, (The Chemical Rubber Co., Cleveland. 1970).

### CHAPTER III

- <sup>1</sup> D. Jose, S.E. Weitzel, W.A. Baase, and P.H. Von Hippel, "Mapping the interactions of the single-stranded DNA binding protein of bacteriophage T4 (gp32) with DNA lattices at single nucleotide resolution: gp32 monomer binding," *Nucleic Acids Res.* **43**, 9276 (2015).
- <sup>2</sup> E.B. Nielsen and J.A. Schellman, "Rotatory properties of molecules containing two peptide groups: experimental and theoretical results," *Biopolymers* **10**, 1559 (1971).
- <sup>3</sup> V. Madison and J. Schellman, "Optical activity of polypeptides and proteins," *Biopolymers* **11**, 1041 (1972).
- <sup>4</sup> V. Rizzo and J.A. Schellman, "Matrix-method calculation of linear and circular dichroism spectra of nucleic acids and polynucleotides," *Biopolymers* **23**, 435 (1984).
- <sup>5</sup> W.C. Johnson and I. Tinoco, "Circular dichroism of polynucleotides: A simple theory," *Biopolymers* **7**, 727 (1969).
- <sup>6</sup> A.L. Williams Jr, C. Cheong, I. Tinico Jr, and L.B. Clark, "Vacuum ultraviolet circular dichroism as an indicator of helical handedness in nucleic acids," *Nucleic Acids Res.* **14**, 6649 (1986).
- <sup>7</sup> B. Nordén, A. Rodger and T. Dafforn, *Linear Dichroism and Circular Dichroism*, (Royal Society of Chemistry, 2010).
- <sup>8</sup> R. Lyng, A. Rodger, and B. Norden, "The CD of ligand-DNA systems. 2. Poly (dA-dT) B-DNA," *Biopolymers* **32**, 1201 (1992).
- <sup>9</sup> R. Lyng, A. Rodger, and B. Nordén, "The CD of ligand-DNA systems. I. Poly (dG-dC) B-DNA," *Biopolymers* **31**, 1709 (1991).

- <sup>10</sup> T.P. Orlando and T. System, “Two-Level System with Static and Dynamic Coupling,” (2004).
- <sup>11</sup> C.L. Cech, W. Hug, and I. Tinoco, “Polynucleotide circular dichroism calculations: Use of an all-order classical coupled oscillator polarizability theory,” *Biopolymers* **15**, 131 (1976).
- <sup>12</sup> H. DeVoe and I. Tinoco, “The stability of helical polynucleotides: Base contributions,” *J. Mol. Biol.* **4**, 500 (1962).
- <sup>13</sup> A. Holmén, A. Broo, B. Albinsson, and B. Nordén, “Assignment of electronic transition moment directions of adenine from linear dichroism measurements,” *J. Am. Chem. Soc.* **119**, 12240 (1997).
- <sup>14</sup> F. žaloudek, J.S. Novros, and L.B. Clark, “The Electronic Spectrum of Cytosine,” *J. Am. Chem. Soc.* **107**, 7344 (1985).
- <sup>15</sup> J.R. Widom, D. Rappoport, A. Perdomo-Ortiz, H. Thomsen, N.P. Johnson, P.H. von Hippel, A. Aspuru-Guzik, and A.H. Marcus, “Electronic transition moments of 6-methyl isoxanthopterin--a fluorescent analogue of the nucleic acid base guanine.,” *Nucleic Acids Res.* **41**, 995 (2013).
- <sup>16</sup> G. Kodali, M. Narayanan, and R.J. Stanley, “The Excited State Electronic Properties of 6-Methylisoxanthopterin (6-MI): An Experimental and Theoretical Study,” *J. Phys. Chem. B* (2012).
- <sup>17</sup> K. Datta, N.P. Johnson, G. Villani, A.H. Marcus, and P.H. Von Hippel, “Characterization of the 6-methyl isoxanthopterin (6-MI) base analog dimer, a spectroscopic probe for monitoring guanine base conformations at specific sites in nucleic acids,” *Nucleic Acids Res.* **40**, 1191 (2012).
- <sup>18</sup> W.D. Wilson and I.G. Lopp, “Analysis of cooperativity and ion effects in the interaction of quinacrine with DNA,” *Biopolymers* **18**, 3025 (1979).
- <sup>19</sup> U. Asseline, M. Delarue, G. Lancelot, F. Toulmé, N.T. Thuong, T. Montenay-Garestier, and C. Hélène, “Nucleic acid-binding molecules with high affinity and base sequence specificity: intercalating agents covalently linked to oligodeoxynucleotides.,” *Proc. Natl. Acad. Sci. U. S. A.* **81**, 3297 (1984).
- <sup>20</sup> M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, and E. Lindah, “Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers,” *SoftwareX* **1–2**, 19 (2015).
- <sup>21</sup> T.E. Cheatham III and P.A. Kollman, “Molecular dynamics simulation of nucleic acids,” *Annu. Rev. Phys. Chem.* **51**, 435 (2000).
- <sup>22</sup> A. Pérez, I. Marchán, D. Svozil, J. Sponer, T.E. Cheatham, C.A. Laughton, and M. Orozco, “Refinement of the AMBER force field for nucleic acids: Improving the description of  $\alpha/\gamma$  conformers,” *Biophys. J.* **92**, 3817 (2007).



- <sup>23</sup> W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.* **79**, 926 (1983).
- <sup>24</sup> R. Galindo-Murillo, C. Bergonzo, and T.E. Cheatham, "Molecular modeling of nucleic acid structure: Setup and analysis," *Curr. Protoc. Nucleic Acid Chem.* **56**, 7 (2014).
- <sup>25</sup> G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *J. Chem. Phys.* **126**, 014101 (2007).
- <sup>26</sup> M. Parrinello, A. Rahman, and R. a. Parrinello M, "Crystal Structure and Pair Potentials: A Molecular\_Dynamics Study," *Phys. Rev. Lett.* **45**, 1196 (1980).
- <sup>27</sup> C.R. Cantor, M.M. Warshaw, and H. Shapiro, "Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleolides," *Biopolymers* **9**, 1059 (1970).

#### CHAPTER IV

- <sup>1</sup> D. Jose, S.E. Weitzel, W.A. Baase, and P.H. Von Hippel, "Mapping the interactions of the single-stranded DNA binding protein of bacteriophage T4 (gp32) with DNA lattices at single nucleotide resolution: gp32 monomer binding," *Nucleic Acids Res.* **43**, 9276 (2015).
- <sup>2</sup> J.R. Widom, N.P. Johnson, P.H. Von Hippel, and A.H. Marcus, "Solution conformation of 2-aminopurine dinucleotide determined by ultraviolet two-dimensional fluorescence spectroscopy," *New J. Phys.* **15**, (2013).
- <sup>3</sup> M.E. Hawkins, W. Pfliederer, F.M. Balis, D. Porter, and J.R. Knutson, "Fluorescence properties of pteridine nucleoside analogs as monomers and incorporated into oligonucleotides.," *Anal. Biochem.* **244**, 86 (1997).
- <sup>4</sup> K. Datta, N.P. Johnson, G. Villani, A.H. Marcus, and P.H. Von Hippel, "Characterization of the 6-methyl isoxanthopterin (6-MI) base analog dimer, a spectroscopic probe for monitoring guanine base conformations at specific sites in nucleic acids," *Nucleic Acids Res.* **40**, 1191 (2012).
- <sup>5</sup> J.R. Widom, D. Rappoport, A. Perdomo-Ortiz, H. Thomsen, N.P. Johnson, P.H. von Hippel, A. Aspuru-Guzik, and A.H. Marcus, "Electronic transition moments of 6-methyl isoxanthopterin--a fluorescent analogue of the nucleic acid base guanine.," *Nucleic Acids Res.* **41**, 995 (2013).

- <sup>6</sup> N.P. Johnson, H. Ji, T.H. Steinberg, P.H. von Hippel, and A.H. Marcus, “Sequence-Dependent Conformational Heterogeneity and Proton-Transfer Reactivity of the Fluorescent Guanine Analogue 6-Methyl Isoxanthopterin (6-MI) in DNA,” *J. Phys. Chem. B* **119**, 12798 (2015).
- <sup>7</sup> J.H. Han, N. Chitrapriya, H.S. Lee, Y.-A. Lee, S.K. Kim, and M.-J. Jung, “Behavior of the Guanine Base in G-quadruplexes Probed by the Fluorescent Guanine Analog, 6-Methyl Isoxanthopterin,” *Bull. Korean Chem. Soc.* **38**, 183 (2017).
- <sup>8</sup> C.R. Cantor and P.R. Schimmel, “Biophysical Chemistry, Part I: The Conformation of Biological Molecules,” *J. Solid-Phase Biochem.* **5**, (1980).
- <sup>9</sup> et al Olson, Wilma K., “DNA sequence-dependent deformability deduced from protein–DNA crystal complexes.,” **95**, 1 (1998).
- <sup>10</sup> A. Plumridge, S.P. Meisburger, K. Andresen, and L. Pollack, “The impact of base stacking on the conformations and electrostatics of single-stranded DNA,” *Nucleic Acids Res.* **45**, 3932 (2017).
- <sup>11</sup> H. Chen, S.P. Meisburger, S.A. Pabit, J.L. Sutton, W.W. Webb, and L. Pollack, “Ionic strength-dependent persistence lengths of single-stranded RNA and DNA,” *Proc. Natl. Acad. Sci.* **109**, 799 (2012).

## CHAPTER V

- <sup>1</sup> L.S. Lerman, “Structural considerations in the interaction of DNA and acridines.,” *J. Mol. Biol.* **3**, 18 (1961).
- <sup>2</sup> L.S. Lerman, “the Structure of the Dna-Acridine Complex\*,” *Proc. Natl. Acad. Sci. U. S. A.* **49**, 94 (1963).
- <sup>3</sup> R.B. Macgregor Jr, R.M. Clegg, and T.M. Jovin, “Viscosity dependence of ethidium-DNA intercalation kinetics,” *Biochemistry* **26**, 4008 (1987).
- <sup>4</sup> A.S. Biebricher, I. Heller, R.F.H. Roijmans, T.P. Hoekstra, E.J.G. Peterman, and G.J.L. Wuite, “The impact of DNA intercalators on DNA and DNA-processing enzymes elucidated through force-dependent binding kinetics,” *Nat. Commun.* **6**, 7304 (2015).
- <sup>5</sup> R.F. Bazoni, C.H.M. Lima, E.B. Ramos, and M.S. Rocha, “Force-dependent persistence length of DNA–intercalator complexes measured in single molecule stretching experiments,” *Soft Matter* **11**, 4306 (2015).
- <sup>6</sup> A. Mukherjee, R. Lavery, B. Bagchi, and J.T. Hynes, “On the molecular mechanism of drug intercalation into DNA: A simulation study of the intercalation pathway, free energy, and DNA structural changes,” *J. Am. Chem. Soc.* **130**, 9747 (2008).

- <sup>7</sup> M. Wilhelm, A. Mukherjee, B. Bouvier, K. Zakrzewska, J.T. Hynes, and R. Lavery, "Multistep drug intercalation: molecular dynamics and free energy studies of daunomycin binding to DNA.," *J. Am. Chem. Soc.* **134**, 8588 (2012).
- <sup>8</sup> W.D. Sasikala and A. Mukherjee, "Molecular mechanism of direct proflavine-DNA intercalation: Evidence for drug-induced minimum base-stacking penalty pathway," *J. Phys. Chem. B* **116**, 12208 (2012).
- <sup>9</sup> W.D. Sasikala and A. Mukherjee, "Intercalation and de-intercalation pathway of proflavine through the minor and major grooves of DNA: roles of water and entropy.," *Phys. Chem. Chem. Phys.* **15**, 6446 (2013).
- <sup>10</sup> R. Lyng, T. Härd, and B. Norden, "Induced CD of DNA intercalators: Electric dipole allowed transitions," *Biopolymers* **26**, 1327 (1987).
- <sup>11</sup> K. Datta, N.P. Johnson, G. Villani, A.H. Marcus, and P.H. Von Hippel, "Characterization of the 6-methyl isoxanthopterin (6-MI) base analog dimer, a spectroscopic probe for monitoring guanine base conformations at specific sites in nucleic acids," *Nucleic Acids Res.* **40**, 1191 (2012).
- <sup>12</sup> A. Rescifina, C. Zagni, M.G. Varrica, V. Pistarà, and A. Corsaro, "Recent advances in small organic molecules as DNA intercalating agents: Synthesis, activity, and modeling," *Eur. J. Med. Chem.* **74**, 95 (2014).
- <sup>13</sup> B. Zhang, X. Li, B. Li, C. Gao, and Y. Jiang, "Acridine and its derivatives: a patent review (2009 – 2013)," *Expert Opin. Ther. Pat.* **24**, 647 (2014).
- <sup>14</sup> S. Nafisi, A.A. Saboury, N. Keramat, J.-F. Neault, and H.-A. Tajmir-Riahi, "Stability and structural features of DNA intercalation with ethidium bromide, acridine orange and methylene blue," *J. Mol. Struct.* **827**, 35 (2007).
- <sup>15</sup> P. Cieplak, S.N. Rao, C. Hélène, T. Montenay-Garestier, and P.A. Kollman, "Conformations of duplex structures formed by oligodeoxynucleotides covalently linked to the intercalator 2-methoxy-6-chloro-9-aminoacridine," *J. Biomol. Struct. Dyn.* **5**, 361 (1987).
- <sup>16</sup> K. Fukui and K. Tanaka, "The acridine ring selectively intercalated into a DNA helix at various types of abasic sites: Double strand formation and photophysical properties," *Nucleic Acids Res.* **24**, 3962 (1996).
- <sup>17</sup> N.P. Ernsting, S.A. Kovalenko, J.L.P. Lustres, T. Senyushkina, and M.E. Michelbeyerle, "Photophysics of 9-Amino-6-chloro-2-methoxyacridine in aqueous solutions and selectively intercalated in duplex DNA," 2001 (2002).
- <sup>18</sup> N.P. Johnson, H. Ji, T.H. Steinberg, P.H. von Hippel, and A.H. Marcus, "Sequence-Dependent Conformational Heterogeneity and Proton-Transfer Reactivity of the Fluorescent Guanine Analogue 6-Methyl Isoxanthopterin (6-MI) in DNA," *J. Phys. Chem. B* **119**, 12798 (2015).

- <sup>19</sup> S. Hess, W.B. Davis, A.A. Voityuk, N. Rösch, M.E. Michel-Beyerle, N.P. Ernstring, S.A. Kovalenko, and J.L. Pérez Lustres, “Excited-state photophysics of an acridine derivative selectively intercalated in duplex DNA,” *ChemPhysChem* **3**, 452 (2002).
- <sup>20</sup> T. Forster, “Delocalization excitation and excitation transfer,” *Mod. Quantum Chem.* (1965).
- <sup>21</sup> A. Löffler and G. Schwarz, “Thermodynamic Behaviour of Acridine Orange in Solution,” *56* (1972).
- <sup>22</sup> S.K. Obendorf, J. Pickworth, H.M. Berman, H.L. Carrell, and P.R. Hansen, “Aggregation of Acridine Orange : Crystal Structure of Acridine Orange Tetrachlorozincate  $\sim C_{22}H_{14}N_2 \cdot 4HCl \cdot ZnCl_2 \cdot 4H_2O$ ,” **2944**, (1976).
- <sup>23</sup> G. Bunt and F.S. Wouters, “FRET from single to multiplexed signaling events,” *Biophys. Rev.* **9**, 119 (2017).
- <sup>24</sup> L. Olejko and I. Bald, “FRET efficiency and antenna effect in multi-color DNA origami-based light harvesting systems,” *RSC Adv.* **7**, 23924 (2017).
- <sup>25</sup> C. Phelps, W. Lee, D. Jose, P.H. von Hippel, and A.H. Marcus, “Single-molecule FRET and linear dichroism studies of DNA breathing and helicase binding at replication fork junctions,” *Proc. Natl. Acad. Sci. U. S. A.* **110**, 17320 (2013).
- <sup>26</sup> P.H. von Hippel, N.P. Johnson, and A.H. Marcus, “50 years of DNA ‘Breathing’: Reflections on old and new approaches [For special issue of biopolymers on 50 years of nucleic acids research],” *Biopolymers* **99**, n/a (2013).
- <sup>27</sup> M.D. Frank-Kamenetskii and S. Prakash, “Fluctuations in the DNA double helix: A critical review,” *Phys. Life Rev.* **11**, 153 (2014).
- <sup>28</sup> S. Neidle and Z. Abraham, “Structural and sequence-dependent aspects of drug intercalation into nucleic acids,” *CRC Crit. Rev. Biochem.* **17**, 73 (1984).
- <sup>29</sup> H.M. Sobell, T.D. Sakore, S.C. Jain, A. Banerjee, K.K. Bhandary, B.S. Reddy, and E.D. Lozansky, in *Cold Spring Harbor Symposia on Quantitative Biology* (Cold Spring Harbor Laboratory Press, 1983), pp. 293–314.
- <sup>30</sup> S. Li, V.R. Cooper, T. Thonhauser, B.I. Lundqvist, and D.C. Langreth, “Stacking interactions and DNA intercalation,” *J. Phys. Chem. B* **113**, 11166 (2009).

## APPENDIX A

<sup>1</sup> Turro, N. T., V. Ramamurthy, J. C. Scaiano, *Principles of Molecular Photochemistry*; University Science Books: Sausalito, CA, 2009.

<sup>2</sup> Nitzan, A., *Chemical Dynamics in Condensed Phases: Relaxation, Transfer and Reactions in Condensed Molecular Systems*; Oxford University Press: New York, 2006.

<sup>3</sup> Datta, K.; Johnson, N. P.; Villani, G.; Marcus, A. H.; von Hippel, P. H., Characterization of the 6-Methyl Isoxanthopterin (6-MI) Base Analog Dimer, a Spectroscopic Probe for Monitoring Guanine Base Conformations at Specific Sites in Nucleic Acids. *Nucl. Acids Res.* 2012, 40, 1191-202.

<sup>4</sup> Widom, J. R.; Johnson, N. P.; von Hippel, P. H.; Marcus, A. H., Solution Conformation of 2-Aminopurine (2-AP) Dinucleotide Determined by Ultraviolet 2D Fluorescence Spectroscopy (UV-2DFS). *New J. Phys.* 2013, 15, 025029- 025016.

## APPENDIX B

<sup>1</sup> R.F. Stewart and N. Davidson, "Polarized absorption spectra of purines and pyrimidines," *J. Chem. Phys.* **39**, 255 (1963).

<sup>2</sup> J.S. Novros and L.B. Clark, "On the electronic spectrum of 1-methyluracil," *J. Phys. Chem.* **90**, 5666 (1986).

<sup>3</sup> W.A. Eaton and T.P. Lewis, "Polarized Single-Crystal Absorption Spectrum of 1-Methyluracil," *J. Chem. Phys.* **53**, 2164 (1970).

<sup>4</sup> Y. Matsuoka and B. Norden, "Linear dichroism studies of nucleic acid bases in stretched poly(vinyl alcohol) film. Molecular orientation and electronic transition moment directions," *J. Phys. Chem.* **86**, 1378 (1982).

<sup>5</sup> B.G. Anex, A.F. Fucaloro, and A. Dutta-Ahmed, "Single-crystal quartz-ultraviolet spectra of thymine anhydrate. Further evidence for the degeneracy of the 260-m.mu. band of the uracils," *J. Phys. Chem.* **79**, 2636 (1975).

<sup>6</sup> B. Bouvier, T. Gustavsson, D. Markovitsi, and P. Millié, "Dipolar coupling between electronic transitions of the DNA bases and its relevance to exciton states in double helices," *Chem. Phys.* **275**, 75 (2002).

<sup>7</sup> A.L. Williams Jr, C. Cheong, I. Tinico Jr, and L.B. Clark, "Vacuum ultraviolet circular dichroism as an indicator of helical handedness in nucleic acids," *Nucleic Acids Res.* **14**, 6649 (1986).

<sup>8</sup> L.B. Clark, "Electronic spectrum of the adenine chromophore," *J. Phys. Chem.* **94**, 2873 (1990).

<sup>9</sup> L.B. Clark, "Polarization assignments in the 270-nm band of the adenine chromophore," *J. Phys. Chem.* **93**, 5345 (1989).

- <sup>10</sup> R.F. Stewart and L.H. Jensen, "Crystal Structure of 9-Methyladenine," *J. Chem. Phys.* **40**, 2071 (1964).
- <sup>11</sup> A. Holmén, A. Broo, B. Albinsson, and B. Nordén, "Assignment of electronic transition moment directions of adenine from linear dichroism measurements," *J. Am. Chem. Soc.* **119**, 12240 (1997).
- <sup>12</sup> P.R. Callis and W. Simpson, "Polarization of electronic transitions in cytosine," *J. Am. Chem. Soc.* **92**, 3593 (1970).
- <sup>13</sup> F. žaloudek, J.S. Novros, and L.B. Clark, "The Electronic Spectrum of Cytosine," *J. Am. Chem. Soc.* **107**, 7344 (1985).
- <sup>14</sup> T.P. Lewis and W.A. Eaton, "Polarized Single-Crystal Absorption Spectrum of Cytosine Monohydrate," *J. Am. Chem. Soc.* **93**, 2054 (1971).
- <sup>15</sup> L.B. Clark, "Electronic spectra of crystalline 9-ethylguanine and guanine hydrochloride," *J. Am. Chem. Soc.* **99**, 3934 (1977).
- <sup>16</sup> L.B. Clark, "Electronic Spectra of Crystalline Guanosine: Transition Moment Directions of the Guanine Chromophore," *J. Am. Chem. Soc.* **116**, 5265 (1994).
- <sup>17</sup> Y. Matsuoka and K. Yamaoka, "Film Dichroism. II. Linearly-polarized absorption spectra of acridine dyes in the stretched poly (vinyl alcohol) films," *Bull. Chem. Soc. Jpn.* **52**, 3163 (1979).
- <sup>18</sup> Y. Matsuoka and K. Yamaoka, "Film Dichroism. V. Linear Dichroism Study of Acridine Dyes in Films with Emphasis on the Electronic Transitions Involved in the Long-wavelength Band of the Absorption Spectrum," *Bull. Chem. Soc. Jpn.* **53**, 2146 (1980).
- <sup>19</sup> T. Imae, "Calculated Circular Dichroism of Acridine Orange Bound to Random-Coil Polypeptide," *Polym. J.* **9**, 541 (1977).

## **APPENDIX D**

- <sup>1</sup> A. Moreno, J. Knee, and I. Mukerji, "Applying 6-Methylisoxanthopterin-Enhanced Fluorescence To Examine Protein–DNA Interactions in the Picomolar Range," *Biochemistry* **51**, 6847 (2012).
- <sup>2</sup> J.R. Lakowicz, *Principles of Fluorescence Spectroscopy* (Springer, 2008).
- <sup>3</sup> K.B. Eisenthal and S. Siegel, "Influence of Resonance Transfer on Luminescence Decay," *J. Chem. Phys.* **41**, 652 (1964).

- <sup>4</sup> A.K. Todd, A. Adams, J.H. Thorpe, W.A. Denny, L.P.G. Wakelin, and C.J. Cardin, "Major groove binding and 'DNA-induced' fit in the intercalation of a derivative of the mixed topoisomerase I/II poison N-(2- (dimethylamino)ethyl)acridine-4-carboxamide (DACA) into DNA: X-ray structure complexed to d(CG(5-Bru)ACG)<sub>2</sub> at 1.3-Å resolution [2]," *J. Med. Chem.* **42**, 536 (1999).
- <sup>5</sup> K. Fukui and K. Tanaka, "The acridine ring selectively intercalated into a DNA helix at various types of abasic sites: Double strand formation and photophysical properties," *Nucleic Acids Res.* **24**, 3962 (1996).
- <sup>6</sup> A. Plumridge, S.P. Meisburger, K. Andresen, and L. Pollack, "The impact of base stacking on the conformations and electrostatics of single-stranded DNA," *Nucleic Acids Res.* **45**, 3932 (2017).
- <sup>7</sup> J.S. Sun, J.C. François, T. Montenay-Garestier, T. Saison-Behmoaras, V. Roig, N.T. Thuong, and C. Hélène, "Sequence-specific intercalating agents: intercalation at specific sequences on duplex DNA via major groove recognition by oligonucleotide-intercalator conjugates.," *Proc. Natl. Acad. Sci. U. S. A.* **86**, 9198 (1989).
- <sup>8</sup> S. Hess, W.B. Davis, A.A. Voityuk, N. Rösch, M.E. Michel-Beyerle, N.P. Ernsting, S.A. Kovalenko, and J.L. Pérez Lustres, "Excited-state photophysics of an acridine derivative selectively intercalated in duplex DNA," *ChemPhysChem* **3**, 452 (2002).
- <sup>9</sup> J.L. Mergny, A.S. Boutorine, T. Garestier, F. Belloc, M. Rougee, N. V Bulychev, A.A. Koshkin, J. Bourson, A. V Lebedev, B. Valeur, and et al., "Fluorescence energy transfer as a probe for nucleic acid structures and sequences," *Nucleic Acids Res.* **22**, 920 (1994).
- <sup>10</sup> N. Busto, B. García, J.M. Leal, J.F. Gaspar, C. Martins, A. Boggioni, and F. Secco, "ACMA (9-amino-6-chloro-2-methoxy acridine) forms three complexes in the presence of DNA," *Phys. Chem. Chem. Phys.* **13**, 19534 (2011).
- <sup>11</sup> W.B. Davis, S. Hess, I. Naydenova, R. Haselsberger, A. Ogrodnik, M.D. Newton, and M.E. Michel-Beyerle, "Distance-dependent activation energies for hole injection from protonated 9-amino-6-chloro-2-methoxyacridine into duplex DNA," *J. Am. Chem. Soc.* **124**, 2422 (2002).
- <sup>12</sup> N. Busto, B. García, J.M. Leal, F. Secco, and M. Venturini, "The mode of binding ACMA–DNA relies on the base-pair nature," *Org. Biomol. Chem.* **10**, 2594 (2012).