The photophysical characterization of the guanine analogue
6-methylisoxanthopterin in DNA oligomers
and its application to probing DNA-Protein interactions

By
Andrew Thomas Moreno
Faculty Advisors: Dr. Joseph Knee and Dr. Ishita Mukerji

A Dissertation submitted to the Faculty of Wesleyan University in partial fulfillment of the requirements for the degree of Doctor of Philosophy
Middletown, Connecticut May 2014
To my family
Acknowledgements

I would like to thank my research advisors Dr. Joseph Knee and Dr. Ishita Mukerji for teaching me valuable scientific skills and excellent mentors as well. I would also like to thank my thesis committee, Dr. Rex Pratt and Dr. David Beveridge for their advice and support during my research project.

I would also like to thank Dr. Roger McMacken for the overexpressing HU protein strain, Dr. Chaconas for the overexpressing HU strains, Dr. Steve Goodman for overexpressing IHF strain, Dr. Manju Hingorani for the Msh2-Msh6 protein and the 6-MI labeled oligonucleotide, Dr. Michael Brenowitz for the TBP protein, and Mary E. Hawkins for additional 6-MI labeled oligonucleotides. I would also like to acknowledge members of the Beveridge laboratory for their readiness to answer my numerous questions in regards to molecular dynamic simulations.

I would like to acknowledge support from an NIH Training Grant in Molecular Biophysics (T32GM08271). Additional funding is acknowledged through Dr. Ishita Mukerji and her awards, the National Science Foundation project grants (MCB-0316625; MCB-0843656).

I would like to acknowledge my past and present fellow graduate students and co-workers in the Mukerji and Knee labs, with whom I have shared hours of lab time and friendship.

A special thank you goes out to my parents, sisters, and extended family; completion of this thesis would not be possible without their love, encouragement and support. Finally, I would like to thank my wife Alison (my research partner), for her continuous love, being my sounding board, partner in problem solving, and helping me develop as a scientist.
# Table of Contents

Acknowledgements ........................................................................................................... ii

Table of Contents ........................................................................................................... iii

Table of Figures ............................................................................................................. vi

List of Tables ................................................................................................................ vii

Abstract ......................................................................................................................... ix

1 Introduction – ................................................................................................................ 1

1.1 Fluorescence spectroscopy a method to examine DNA conformational dynamics ................................................................................................................................. 1

1.1.1 Types of fluorescent probes for studying nucleic acids: ............................... 3

1.1.2 A summary of external fluorescent probes ................................................... 5

1.1.3 A summary of internal fluorescent probes ..................................................... 8

1.1.4 The photophysical properties of 6-methylisoxanthopterin (6-MI) ............ 21

1.1.5 Proteins associated with the bacterial chromosome ...................................... 26

1.2 The nucleoid associated protein HU (Histone-like U93 protein) ................. 28

1.2.1 The Cellular Functions of HU ........................................................................ 29

1.2.2 The structure-function relationship for HU and IHF ................................. 36

1.2.3 HU binds to DNA independent of sequence context: ................................. 47

2 Theory .......................................................................................................................... 64

2.1 Light interacting with macromolecules ......................................................... 64

2.2 Absorption: transition from ground state to excited state ........................... 66

2.3 Depopulating the excited state: ................................................................. 70

2.4 Factors governing fluorescence intensity ..................................................... 73

2.4.1 Vibrational Relaxation .................................................................................. 74

2.4.2 Internal Conversion ...................................................................................... 75

2.4.3 Intersystem crossing and phosphorescence ............................................. 76

2.4.4 Bimolecular quenching of the excited state ............................................. 77

2.4.5 Resonance energy transfer and excited state photochemical reactions .................................................. 80

2.5 Steady-state and time-resolved fluorescence ............................................. 87

2.5.1 Meaning of the multi-exponential decay ................................................... 95

2.6 Fluorescence anisotropy ..................................................................................... 96

2.6.1 Time-Dependent Anisotropy Decay ......................................................... 104
5.2.2 Local DNA structural deformations effect 6-MI photophysics ........................................... 210
5.2.3 Molecular dynamics simulations indicate the conformation of 6-MI is dependent on sequence.................................................................................................................. 230
5.3 Discussion ....................................................................................................................................... 246
  5.3.1 The observed changes to fluorescence reflects the position of 6-MI relative to the perturbation.............................................................. 247
  5.3.2 Molecular modeling indicates the conformation of 6-MI is sequence dependent .............................................................................................. 254
  5.3.3 Summary ...................................................................................................................................... 256
6 Biophysical characterization of the HU-overhang complex ........................................................................ 257
  6.1 Introduction ...................................................................................................................................... 257
  6.2 Results ............................................................................................................................................ 263
    6.2.1 HU-DNA stoichiometry............................................................................................................. 263
    6.2.2 HU affinity is dependent on DNA structure ........................................................................... 269
    6.2.3 HU influences the excited state of 6-MI incorporated within the overhang.............................. 277
    6.2.4 HU binding alters the dynamics of 6-MI incorporated within the overhang............................. 286
    6.2.5 HU binding perturbs global DNA geometry as shown by Förster resonance energy transfer (FRET) ........................................................................ 292
    6.2.6 FRET mapping the HU-overhang complex .............................................................................. 303
  6.3 Discussion ...................................................................................................................................... 308
    6.3.1 Generating models of the HU-overhang complex .................................................................. 308
    6.3.2 Evaluation of models in terms of FRET distances ................................................................. 315
    6.3.3 The model of the HU-overhang complex is consistent with other experimental data .............. 322
    6.3.4 Insights into the mechanism of HU specificity ......................................................................... 328
  6.4 Conclusion ....................................................................................................................................... 331
7 References .......................................................................................................................................... 332
Table of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>External Fluorescent Dyes</td>
<td>4</td>
</tr>
<tr>
<td>1-2</td>
<td>Internal dyes</td>
<td>9</td>
</tr>
<tr>
<td>1-3</td>
<td>the first electronic transition dipole moments for 6-MI (A) and 2-AP (B)</td>
<td>23</td>
</tr>
<tr>
<td>1-4</td>
<td>Phase inversion recombination</td>
<td>31</td>
</tr>
<tr>
<td>1-5</td>
<td>HU and IHF sequence alignment</td>
<td>38</td>
</tr>
<tr>
<td>1-6</td>
<td>Anabaena HUα2 and crystal structure</td>
<td>40</td>
</tr>
<tr>
<td>1-7</td>
<td>IHF crystal structure</td>
<td>42</td>
</tr>
<tr>
<td>1-8</td>
<td>HU-DNA interactions</td>
<td>44</td>
</tr>
<tr>
<td>1-9</td>
<td>HU GMSA with variable overhang</td>
<td>59</td>
</tr>
<tr>
<td>2-1</td>
<td>Jablonski diagram</td>
<td>65</td>
</tr>
<tr>
<td>2-2</td>
<td>Potential energy diagram</td>
<td>68</td>
</tr>
<tr>
<td>2-3</td>
<td>Transfer Efficiency</td>
<td>84</td>
</tr>
<tr>
<td>2-4</td>
<td>Orientation factor</td>
<td>86</td>
</tr>
<tr>
<td>2-5</td>
<td>Excited state decay</td>
<td>90</td>
</tr>
<tr>
<td>2-6</td>
<td>TCSPC schematic</td>
<td>91</td>
</tr>
<tr>
<td>2-7</td>
<td>Anisotropy schematic</td>
<td>97</td>
</tr>
<tr>
<td>2-8</td>
<td>Fluorophore coordinate system</td>
<td>99</td>
</tr>
<tr>
<td>2-9</td>
<td>Depolarization</td>
<td>102</td>
</tr>
<tr>
<td>2-10</td>
<td>Time-dependent anisotropy</td>
<td>107</td>
</tr>
<tr>
<td>2-11</td>
<td>TR-FRET probability distribution</td>
<td>113</td>
</tr>
<tr>
<td>3-1</td>
<td>HU purification</td>
<td>117</td>
</tr>
<tr>
<td>3-2</td>
<td>Force Field Diagram</td>
<td>144</td>
</tr>
<tr>
<td>3-3</td>
<td>Amber flowchart</td>
<td>149</td>
</tr>
<tr>
<td>4-1</td>
<td>IHF crystal structure</td>
<td>156</td>
</tr>
<tr>
<td>4-2</td>
<td>Duplex Temperature Melt</td>
<td>159</td>
</tr>
<tr>
<td>4-3</td>
<td>6-MI relative Q,Y</td>
<td>161</td>
</tr>
<tr>
<td>4-4</td>
<td>Models of 6-MI conformation</td>
<td>164</td>
</tr>
<tr>
<td>4-5</td>
<td>6-MI monomer KI quenching</td>
<td>165</td>
</tr>
<tr>
<td>4-6</td>
<td>6-MI solvent accessible fraction</td>
<td>166</td>
</tr>
<tr>
<td>4-7</td>
<td>Predicted OH cleavage pattern</td>
<td>168</td>
</tr>
<tr>
<td>4-8</td>
<td>6-MI excited state decay</td>
<td>171</td>
</tr>
<tr>
<td>4-9</td>
<td>6-MI TR-anisotropy</td>
<td>174</td>
</tr>
<tr>
<td>4-10</td>
<td>6-MI Anisotropic equilibrium binding curves</td>
<td>177</td>
</tr>
<tr>
<td>4-11</td>
<td>6-MI S/N vs. Sequence Context</td>
<td>179</td>
</tr>
<tr>
<td>4-12</td>
<td>Equilibrium binding curves using picoMolar [6-MI]</td>
<td>181</td>
</tr>
<tr>
<td>4-13</td>
<td>Protein-induced intensity change vs. 6-MI sequence position</td>
<td>184</td>
</tr>
<tr>
<td>4-14</td>
<td>Msh2-Msh6 crystal structure</td>
<td>186</td>
</tr>
<tr>
<td>4-15</td>
<td>Protein-induced change to lifetime vs. 6-MI sequence position</td>
<td>188</td>
</tr>
<tr>
<td>4-16</td>
<td>Protein-induced change to r(t) vs. 6-MI sequence position</td>
<td>191</td>
</tr>
<tr>
<td>5-1</td>
<td>Relative QY for quenched sequences</td>
<td>203</td>
</tr>
<tr>
<td>5-2</td>
<td>Sequence dependent 6-MI lifetime</td>
<td>205</td>
</tr>
<tr>
<td>5-3</td>
<td>Sequence dependent excited state population</td>
<td>206</td>
</tr>
<tr>
<td>5-4</td>
<td>Sequence dependent TR-anisotropy</td>
<td>208</td>
</tr>
<tr>
<td>5-5</td>
<td>Temperature melts 260 nm</td>
<td>211</td>
</tr>
<tr>
<td>5-6</td>
<td>Temperature melt 340 nm</td>
<td>212</td>
</tr>
<tr>
<td>5-7</td>
<td>Fraction solvent accessible</td>
<td>215</td>
</tr>
<tr>
<td>5-8</td>
<td>Perturbation QY</td>
<td>216</td>
</tr>
</tbody>
</table>
List of Tables
Table 4-1 Sequences of 6-MI-containing and complementary oligomers ........................................... 158
Table 4-2 Parameters derived from fitting of time-resolved fluorescence decay curves ...... 162
Table 4-3 Parameters derived from fitting of time-resolved anisotropy decay curves .......... 173
Table 4-4 Equilibrium Protein-DNA Dissociation constants (Kd) measured with 6-MI containing duplexes .................................................................................................................. 176
Table 5-1 6-MI containing oligomers ................................................................................................ 199
Table 5-2 Parameters derived from fitting of time-resolved fluorescence decay curves for bulge/mismatch ........................................................................................................ 200
Table 5-3 Parameters derived from fitting of time-resolved fluorescence decay curves for 6-MI sequences .................................................................................................................. 201
Table 5-4 Temperature Melts for absorbance at 340nm and 260nm ............................................ 213
Table 5-5a Gaussian fitted center of peak for parameters describing base/6-MI step .......... 233
Table 5-6 Base step parameters for B-form and A-form DNA ......................................................... 234
Table 6-1 HU DNA constructs ........................................................................................................ 264

Figure 5-9 6-MI lifetime base-bulge/mismatch ................................................................. 218
Figure 5-10 Fractional population perturbation ...................................................................... 219
Figure 5-11 6-MI r(t) vs base-bulge A ................................................................................ 221
Figure 5-12 6-MI r(t) vs base-bulge A ................................................................................ 222
Figure 5-13 MD RMSD ........................................................................................................... 231
Figure 5-14 15th position base step .................................................................................... 235
Figure 5-15 Rise Base step ..................................................................................................... 239
Figure 5-16 Slide base step .................................................................................................... 240
Figure 5-17 Slide base step .................................................................................................... 241
Figure 5-18 Twist base step ................................................................................................... 242
Figure 5-19 Tilt base step ...................................................................................................... 243
Figure 5-20 Roll base step ..................................................................................................... 244
Figure 5-21 Twist detail analysis ........................................................................................... 245
Figure 6-1 HU crystal structure ............................................................................................ 259
Figure 6-2 HU-DNA Stoichiometry ....................................................................................... 267
Figure 6-3 HU binding affinity for 5’ and 3’ overhang ........................................................ 270
Figure 6-4 HU binding affinity varied overhangs ................................................................. 272
Figure 6-5 Binding affinity of HU mutants ............................................................................ 274
Figure 6-6 HU structural non-specific affinity ..................................................................... 276
Figure 6-7 6-MI overhang excited state decay ...................................................................... 278
Figure 6-8 $\tau_{\text{mol}}(t)$ for 6-MI in the 3’ overhang ................................................................. 281
Figure 6-9 $\tau_{\text{mol}}(t)$ for 6-MI in the 5’ overhang ................................................................. 282
Figure 6-10 6-MI TR-anisotropy in overhang ................................................................... 287
Figure 6-11 6-MI rotational correlation 3’ overhang ........................................................ 289
Figure 6-12 6-MI rotational correlation 5’ overhang ........................................................ 290
Figure 6-13 Calculate Bend angle........................................................................................ 310
Figure 6-14 GMSA HU overhang ....................................................................................... 314
Figure 6-15 HU-overhang models ....................................................................................... 316
Figure 6-16 FRET vs. model distance ................................................................................. 319
Figure 6-17 correlation plots without Q64C and T70C FRET .................................................. 323
Table 6-2 HU equilibrium association constants for specific and non-specific binding modes

Table 6-3 Mapping HU-induced perturbations to base conformation using the position dependent excited state decay of 6-MI within the overhang construct

Table 6-4 FRET determined distance and calculated bend angle for 3 and 5 prime overhang in the presence and absence of HU

Table 6-5 FRET determined distance and calculated bend angle for nicked DNA in the presence and absence of HU

Table 6-6 FRET determined distance and bend angle for H1 duplex (H34aL5’) in the presence and absence of IHF and HU

Table 6-7 TR-FRET between 6-MI incorporated overhangs and FAM site-specific labeled HU

Table 6-8 Experimental FRET and predicted model distances

Table 6-9 Summary of RMSD between model and FRET distances for each construct
Abstract

The ability to observe structure and dynamics of single residues in an oligomer is critical when investigating the function of DNA-binding proteins. The development of fluorescent nucleoside analogs, has greatly improved the amount of information available from both steady-state and time-resolved fluorescence experiments. The significant reduction in quantum yield observed when probes are incorporated into either an oligomer or a duplex limits their potential application and use. The focus of this thesis is on the photophysical characterization and application of the guanine analogue 6-methylisoxanthopterin (6-MI), specifically in the context of sequence specific fluorescence enhancement. While investigating the photophysical properties of this nucleoside analog; we discovered a pentamer DNA sequence (ATFAA), which exhibits enhanced fluorescence (in either orientation) upon formation of duplex DNA. After extensive characterization of this unique sequence we hypothesized the enhanced fluorescence was the result of a sterically hindered conformation of 6-MI. In this sequence 6-MI has reduced mobility preventing access to the normal non-radiative pathways that an analog experiences when incorporated into a polymer of nucleotides. We exploited the high quantum yield and well characterized pentamer system to examine the interaction between nucleoid associated protein HU and the structural specific overhang construct.

HU is a non-sequence specific DNA/RNA binding protein, which exhibits a strong preference for various distortions in DNA; nicks, gaps, cruciform, and overhangs. The key structural and dynamic elements important in the mechanism of HU structural specificity remain unknown. An important advance is to identify the
solution state geometry of the HU-overhang complex and understand the role of nucleic acid topology in HU specificity. We developed a model of the HU-overhang complex based on using the fluorescent base analog 6-methylisoxanthopterin (6-MI) to examine local base dynamics and Förster resonance energy transfer to investigate the global DNA structure by generating distance constraints for the HU-overhang complex in solution. HU binding induces a ~120° bend to the overhang construct. We hypothesize the high affinity complex, $K_D$ is ~300 pM, forms when the HU $\alpha$ subunit recognizes the unique conformational dynamics of bases at the junction, while ~9bp away the HU $\beta$ subunit interacts with single-stranded DNA.
1 Introduction –

1.1 Fluorescence spectroscopy a method to examine DNA conformational dynamics

There are a number of experimental techniques available to characterize physical interactions between proteins and DNA, but many of the commonly used approaches are subject to limitations that reduce their utility and sensitivity. For example, techniques such as Gel Mobility Shift Assay (GMSA), Isothermal Titration Calorimetry (ITC), Atomic Force Microscopy (AFM), and Magnetic Tweezers (MT) are limited in resolution to the entire DNA polymer and only report on global DNA distortions. In contrast, techniques like Nuclear Magnetic Resonance (NMR) and other high-resolution spectroscopy techniques achieve single-base resolution, but these experiments can be prohibitively difficult, time consuming and expensive to execute in high-throughput. NMR in particular requires both expensive instrumentation that is not readily available to many laboratories as well as labeling of individual bases with N\textsuperscript{15}. Moreover, NMR requires sample concentrations that differ by orders of magnitude from physiological conditions. Similarly, gel mobility shift assays also subject protein-nucleic acid complexes to non-ideal and non-physiological conditions, which affects the fraction bound over the course of the measurement. In a GMSA experiment, the binding equilibrium is established in the pores of a polyacrylamide gel and in the presence of an electric field, where free exchange between unbound and bound protein can be hindered. In contrast to the aforementioned techniques, fluorescence spectroscopy represents an analytical method that is able to overcome many of these challenges.

In fluorescence spectroscopy, photons of a specific wavelength are absorbed by a population of fluorophores, which subsequently transition from the ground state into an
excited state. The excited state has a finite lifetime, and there are many ways that energy can be dissipated from the fluorophore returning it to the ground state. One such mechanism is fluorescence, where the excited fluorophore emits a photon at a wavelength matching the energy difference between the ground and excited states. Since the excited state is sensitive to the microenvironment of the fluorophore, it is an excellent reporter on macromolecular interactions. One additional property that makes fluorescence spectroscopy so useful for studying biological systems is that the lifetime of the excited state exists on the same time scale as certain conformational fluctuations in biomolecules. More specifically, fluorescent lifetimes are typically between 0.01 and 20 ns, which match the time scale for a range of macromolecular conformational dynamics. Therefore, the physical properties of fluorophores make them ideal for studying macromolecular interactions.¹

Although fluorophores are uniquely suited for studying biomolecules, their widespread application to DNA- and RNA-based systems is a recent development. Several technical advances in recent years have facilitated the application of fluorescence spectroscopy to the study of DNA/RNA conformational dynamics, including advances in the synthesis of oligonucleotides, discovery of new fluorescent probes, and new attachment methods to DNA and RNA macromolecules. These developments along with greater interest in protein-DNA complexes has increased the accessibility, type, and number of fluorescent probes for studying DNA/RNA dynamics.²

With the introduction of these new fluorophore technologies, understanding how DNA-binding proteins perturb nucleic acid structure becomes possible at single-base resolution. The sensitivity of the probes combined with the wealth of information available
from their measured properties can be used to model their conformation and dynamics within the DNA/RNA structure. However, generating models from fluorescence data requires a detailed understanding of how measured properties from a fluorescent probe, such as quantum yield, lifetimes, anisotropy decay and spectral shifts, correlate with its microenvironment. The focus of the first part of this thesis is on the relationship between the photophysical properties of a specific fluorophore, 6-methylisoxanthopterin (6-MI), and its conformational dynamics within DNA. The second part of the thesis uses those properties to characterize the way the bacterial protein HU changes the structure and dynamics of DNA upon binding and to create a model of the HU-DNA nucleoprotein complex.

1.1.1 *Types of fluorescent probes for studying nucleic acids:*

The best probes for fluorescent spectroscopy are highly sensitive to their local environment, have high quantum yields, are very stable even during continued illumination, minimally perturb the system being studies, and exhibit a large Stokes’ shift, which refers to the wavelength difference between the band of maximum absorption and the band of maximum emission. Using these criteria, the best probes that report on local DNA structure have been those which covalently link to the DNA polymer itself. This category of covalent probes can be subdivided into two classes based on the point of attachment to the DNA molecule: external probes are attached by a hydrocarbon linker and internal probes are
Figure 1-1 External Fluorescent Dyes

The external fluorescent dyes: (A) 5-Carboxyfluorescein, succinimidyl ester (FAM) (B) 5 - Carboxytetramethylrhodamine, succinimidyl ester (TAMRA)
attached directly to the ribose sugar of the DNA backbone. There are benefits and
disadvantages to using each class of probe as a fluorescent reporter, and in this thesis both
classes of probes are used in a concerted effort to investigate protein-DNA interactions.
While the external class of probes, such as fluorescein (FAM) and rhodamine (TAMRA)
derivatives (Figure 1-1), were developed first and therefore are well characterized, the newer
internal class of probes can be used to obtain different and complementary information about
protein-DNA interactions (Figure 1-2).

1.1.2 A summary of external fluorescent probes

External probes are small molecules that are covalently linked to a biomolecule.
Their ability to screen local macromolecular interactions limits the utility of external probes.
Most external probes contain aromatic ring systems that exhibit a high degree of
hydrophobicity, which can be problematic when solubilizing the dye or can lead to
aggregation of labeled macromolecule. To address this issue, external dyes containing
charged functional groups have been developed and widely used (Figure 1-1). For example,
two widely used external probes fluorescein and rhodamine contain a carboxylic acid and
protonated imine moiety respectively. However, the spectral properties of fluorophores
containing polar groups can be pH sensitive. Although pH sensitivity can be useful as an
indicator of local polarity, it also represents a complicating factor that must be considered in
experimental design and interpretation. In addition to the pH sensitivity, the polarity of the
fluorophore will affect its interactions with the labeled macromolecule. The negative charge
of fluorescein is known to minimize association with DNA/RNA by repulsing the negatively
charged backbone, while the protonated dye rhodamine makes favorable charge-charge
interactions with the DNA backbone. An additional complicating factor with external
probes is that they can interfere with the formation of protein-DNA complexes, since they range in size from 300 Daltons (Hydroxycoumarin) to 1000 Daltons (Cy5.5). To minimize interactions with proteins longer carbon linkers, 6 to 12 carbon atoms (a 10 to 20 Å), can be used as spacers between the fluorophore and the DNA or protein.

In general, labeling proteins or DNA with external fluorescent dyes is accomplished using either thiol- or amine-reactive chemistry. Unlike nucleic acids, proteins contain naturally occurring thiols (cysteine) and primary amines (lysine), which are ideal for labeling. In cases where a target protein has no reactive groups or more than one reactive group available for modification, site-directed mutagenesis can be used to generate an appropriate substrate. Nucleic acids do not contain naturally reactive thiols and primary amines, thus reactive modifiers are attached to the end of the nucleic acid molecule during synthesis, or an internally modified amine-reactive thymine is incorporated into the DNA oligomer. Thiol- or amine-reactive fluorophores react with the reduced cysteine or deprotonated amine respectively. The resulting thiol-ether or amide linkages are very stable, which makes the fluorophore-biomolecule covalent complex useful for fluorescence spectroscopy. External probes are easily incorporated into DNA and generally cost less than internal probes. However, labeling this way requires excess dye and even under best conditions 100% labeling is hard to achieve. In addition, removal of the excess dye is critical for proper interpretation of experimental results. For example, a sample containing excess dye will exhibit an artificially low anisotropy.

There are several photophysical benefits to using external dyes. First, even when conjugated to proteins or DNA, external dyes can maintain large molar absorptivities.
Coupled with high quantum yields, the bright fluorescent intensity from external dyes provides excellent sensitivity for examining protein-DNA interaction down to the picomolar range. Also, the high quantum yield is beneficial for techniques where signal to noise is an issue such as fluorescence microscopy and single molecule fluorescence.\textsuperscript{2,3} In addition to a high quantum yield, there is good spectral overlap between the emission and the absorption spectra of certain dyes. These dyes that exhibit spectral overlap can be used as FRET pairs to observe the global conformational dynamics of an oligomer. Further, the excited state decay for external dyes is typically single exponential, which is useful for discerning energy transfer in time-resolved FRET (TR-FRET) data. Finally, the motion of certain external dyes (such as fluorescein) in DNA is not restrained, which is beneficial for a couple of reasons. First, the observed anisotropy for DNA labeled with fluorescein is typically < 0.1, such that the binding of large proteins to the labeled DNA is readily observed as a large change in anisotropy. Second, a freely rotating probe is necessary for approximating a random distribution for the orientation of dipole moments for FRET dye pairs. The photophysical properties outlined above underscore the advantages for using external dyes. However these benefits are tempered by certain experimental disadvantages.

Most of the disadvantages for using external probes arise from the molecular properties of the probe, including their chemical nature, size, charge, and method of attachment.\textsuperscript{2,3} For example, the attachment to DNA through a 6-carbon linker increasing the conformational space the probe can occupy and introducing greater uncertainty in its location. This limits the ability of the probe to report on DNA local structure and dynamics and therefore the probe is a poor reporter of local protein-induced structural perturbations.\textsuperscript{2} In addition, the formation of a protein-DNA complex may leave the excited state decay
unchanged for a number of external dyes and while this is a beneficial property for TR-FRET experiments, analyzing only the excited state decay would suggest little change in DNA conformation upon protein binding. Finally, as stated above the size and/or charge of the fluorophore may inhibit formation of the protein DNA complex. In spite of these disadvantages the beneficial physical properties of external dyes ensure an important role in biophysical characterization of protein-DNA interactions.

1.1.3 A summary of internal fluorescent probes

Internal probes, which include fluorescent nucleoside analogs, have greatly increased the amount of information available regarding local DNA structure and dynamics. These probes can be viewed as a complementary system to external probes and used in combination to obtain greater information with regard to DNA conformational dynamics. An internal fluorophore refers to a dye that is covalently attached to a backbone sugar moiety in such a way that it replaces the nucleobase inside the DNA. This group of fluorophores includes planar aromatic molecules as well as fluorescent base analogues (FBAs), all of which are attached to the ribose of the DNA backbone (Figure 1-2). Internal attachment is an important feature of internal probes enabling labeling of the DNA at specific positions, thereby increasing the information available on local DNA structure and dynamics.

The planar aromatics differ in shape and size from FBAs and do not have the requisite structure for the formation of hydrogen bonds with the complementary base. The planar aromatics include molecules such as pyrene, phenanthrene, and stilbene. The molecular structure of these aromatic molecules enables stacking with adjacent bases, and a number of planar aromatics have been shown to minimally distort the DNA structure. In fact,
Internal fluorescent base analogs: (A) 2-Aminopurine (2-AP) (B) pyrrolo-dCytosine (pyrrolo-dC) (C) Tricyclic Cytosine (tC) (D) 6-methylisoxanthopterin (6-MI) (E) 3-methylisoxanthopterin (3-MI) (F) 4-amino-6-methyl-7(8H)-pteridone (6MAP) (G) 4-amino-2,6-methyl-7(8H)-pteridone (DMAP)
there is an example of DNA polymerase incorporating planar aromatic molecules into a DNA strand.\textsuperscript{2} While planar aromatics do stack with the adjacent DNA bases, FBA more closely resemble native bases in structure and are able to hydrogen bond with the requisite complementary base in a DNA duplex. Their ability to participate in molecular interactions that mimic native DNA make FBAs uniquely good at reporting on local DNA structure.

There are several experimental advantages to using FBAs as probes of DNA structure and dynamics. Incorporation of FBAs is minimally disruptive to the DNA structure, which allows them to stack with adjacent bases.\textsuperscript{7,8} Moreover, the structural similarity of internal probes to nucleic acid bases and their ability to hydrogen bond with their cognate base makes these probes ideal for examining protein-DNA interactions.\textsuperscript{9-12} A number of factors such as hydrogen bonding, single- or double-stranded environment, the identity of neighboring bases, and base stacking govern the photophysical properties of a FBA within DNA. Thus, the fluorescent properties of FBAs are highly sensitive to local changes in DNA conformational dynamics, useful for monitoring local protein-induced perturbations.\textsuperscript{13-15} In addition, when coupled with time-resolved fluorescence measurements the internal probe can provide dynamic resolution of DNA structure on the single base scale.\textsuperscript{6}

A deterrent to using internal probes is the fact that base stacking interactions present in single-stranded DNA polymers often lead to a significant reduction in the quantum yield, which is either maintained or further decreased upon formation of duplex DNA.\textsuperscript{2,5} This reduction in quantum yield lowers sensitivity and introduces a concentration limit for using internally labeled oligomers. Many DNA-binding proteins exhibit dissociation constants in the low nanomolar range, which is below the signal to noise threshold for many fluorescent
nucleoside analogs. Thus, the study of these interactions using internal fluorescent probes has remained a challenging problem. As an added complication in TR-FRET applications, the excited state decay of FBAs in DNA is multiple exponentials, which makes modeling TR-FRET data more complex.

Correctly interpreting the biophysical meaning of observed changes in a probe’s fluorescent properties requires a detailed understanding of both the probe alone in solution and when incorporated into a DNA or RNA macromolecule. The fluorescence observed from nucleoside analogs is directly influenced by the structure and dynamics of its local environment. Changes in local environment or molecular motion are reflected in the fluorescent emission spectrum, excited state lifetimes, and anisotropic properties of a fluorophore. However, correlating these measurements with the underlying physical processes remains a challenge, since many of these measurements report on several competing mechanisms. For example, excited state lifetime and fluorescence intensity can be difficult to interpret due to the number of competing non-radiative pathways that also de-populate the excited state (\(k_{\text{ic}}\) internal conversion, \(k_{\text{et}}\) energy transfer, \(k_{\text{Q}}\) quenching, \(k_{\text{isc}}\) intersystem crossing, \(k_p\) photochemistry, \(k_{\text{po}}\) phosphorescence). In the simplest case, the excited state lifetime (and correspondingly, the quantum yield) of a fluorophore are inversely proportional to the rate of radiative decay. In the absence of non-radiative pathways, the radiative rate is proportional to the strength of absorption (dipole strength), which means a stronger absorber has a shorter excited state lifetime and a higher quantum yield. Since non-radiative decay pathways are almost always available to de-populate the excited state in solution, the observed radiative lifetimes are rarely as long as theoretical lifetimes and quantum yields are lower than would be expected by the dipole strength. Understanding
which non-radiative components are competing with the excited state decay and to what extent can be challenging. However, without this information it is difficult to interpret how measurements of fluorescent properties relate to changes in system components, which is particularly important if the probe is used as a physical reporter. A considerable amount of effort has been put into the characterization of several nucleoside analogs’ excited states, which has been used to model potential mechanisms for nucleoside analog fluorescence.6,19-24

Ideally, there would be a FBA available for all four of the native nucleobases that are well characterized with photophysical properties independent of sequence context. However, this is generally not true. FBA photophysical properties are frequently dependent on sequence context and need to be well characterized for each new system. There are a number of FBAs that have been developed for use in biological systems and are commercially available. There are advantages and disadvantages to each FBA. Some of the more commonly used FBA are discussed below.

2-Aminopurine (2-AP) an adenine analogue

2-aminopurine (2-AP), an adenine analog, was discovered in 1969 by Ward et al. and is currently the FBA in greatest use (Figure 1-2).25 2-AP is capable of forming stable base pairs with thymine, uracil, and to a lesser extent cytosine.26 The position of the lowest energy absorption band for 2-AP is centered at 305 nm with an extinction coefficient of 6000 M⁻¹ cm⁻¹, red-shifted relative to native bases allowing for selective excitation when inserted into DNA.
2-aminopurine is a highly fluorescent monomer with a quantum yield of (Φ) ≈ 0.68 ($\lambda_{abs}$=305 nm, $\lambda_{em}$=370 nm). The photophysical properties of 2-AP are sensitive to base-stacking interactions, which reduce its quantum yield in single-stranded or double-stranded oligonucleotides by as much as 100-fold. The extreme sensitivity of 2-AP fluorescence to its microenvironment has been exploited in several types of studies including interrogating nucleic acid structure and dynamics, measuring the strength of protein-nucleic acid interactions, and even quantifying electron transfer within DNA. Examples include EcoRI DNA methyltransferase interaction with DNA, the real-time monitoring of ribozyme folding, and the interaction between uracil DNA glycosylase and DNA. Also, due to 2-AP being the first discovered FBA there is greater volume of work characterizing its photophysics.

As a monomer in solution the excited state decay for 2-AP is single exponential, with a lifetime of 10 ns. 2-AP is dynamically quenched when incorporated into single-stranded or double-stranded DNA. When 2-AP is inserted into DNA, the fluorescence decay is multi-exponential and modeled with as many as four components ($\tau_1 < 100 \text{ ps}$, $\tau_2 < 1 \text{ ns}$, $\tau_3 < 3 \text{ ns}$, $\tau_4 < 10 \text{ ns}$), an indicator of an increased number of competing non-radiative decay pathways. The complex multi-exponential corresponds to several structural populations of 2-AP with varying degrees of interbase stacking. Interestingly, the degree of fluorescence quenching is dependent on the identity of the neighboring bases. Several theoretical quenching models based on semi-empirical methods have been proposed. A recent study sought to address a general problem with these models, which is that the calculations are based on a static geometry and do not address excited state dynamics along a potential energy surface (PES). Based on this work two different competing mechanisms describing fluorescence quenching were proposed: i) when a base is 5-prime adjacent to 2-AP
the excitation is to the $\pi\pi^*$ bright state followed by rapid relaxation to $n\pi^*$ dark state and ii) when a base is 3-prime adjacent to 2-AP a conical intersection (CI) is possible leading to non-radiative decay pathway involving charge-transfer (CT) with the adjacent base. The mechanisms are proposed to compete with each other, depending on the identity and geometry of the adjacent bases. Several studies using femtosecond time-resolved fluorescence indicate charge transfer (CT) is the predominant quenching pathway for 2-AP complexed with other nucleobases. The term CT is used to define two quenching regimes: i. the full transfer of an electron between a donor and acceptor resulting in a pair of radical ions, called photoinduced electron transfer (PET) or ii. “partial” charge transfer, which happens when the excited-state donor forms an excited-state charge transfer complex (CTC) with the acceptor. The process of CT is dependent on the redox potentials of the reactants. Since the excited FBA (FBA*) has a redox potential similar to nucleotide monophosphates, the bases that surround the FBA can be either reduced during PET (called nucleobase reduction or NBR) or oxidized (called nucleobase oxidation or NBO). Whether NBR or NBO occurs is dictated by the free energy change of each process. For 2-AP the PET-quenching mechanism has been investigated using a combination of Stern-Volmer quenching with nucleotide monophosphates, to obtain the quenching rates, and cyclic voltometry (CV) to obtain ground state redox potential, which is then extrapolated for the excited state redox potentials. The results indicate the order in terms of the rate for quenching the excited state 2-AP* is GMP NBO > TMP NBR > AMP NBO > CMP which can occur by either nucleobase oxidation or reduction based on the near equivalence of $\Delta G^{\circ}$NBO and $\Delta G^{\circ}$NBR. Importantly, when 2-AP is incorporated into DNA the geometry and sequence will determine the most favorable non-radiative pathway.
In summary, the advantages for using 2-AP in studying protein-DNA interactions include: i) 2-AP photophysics is highly sensitive to the microenvironment making it an ideal reporter of changes to DNA conformational dynamics, ii) the photophysical properties of 2-AP are well characterized, necessary for correct data interpretation, and iii) 2-AP minimally distorts the DNA structure. The disadvantages include the large reduction in quantum yield when incorporated into ss- and dsDNA, an issue when working with low sample concentrations. Also, compared to the native base adenine, 2-AP is less specific and less efficient at base-pairing to thymine, resulting in increased dynamics within the DNA. The benefits and disadvantages of 2-AP must be evaluated before being used in an experiment.

_Pyrrolo-dC a cytosine analogue_

The fluorescent behaviour of cytosine nucleobase analog pyrrolo-dF*, a derivative of pyrrolo-dC, was first reported in the late 1980s (Figure 1-2). In addition to the 6-membered ring of cytosine, there is either an additional five or six membered ring for the pyrrolo-dC ‘family’. In spite of this additional ring structure, pyrrolo-dF* base-pairs with guanine with similar stability as the Watson-Crick C-G base-pair. During development of pyrrolo family a fluorescent furano-derivative, furano-dT was discovered, it was of little use because the Watson-Crick base-pairing face is not complementary with any of the four native nucleobases. However, treating the furano-dT with concentrated ammonia resulted in the conversion to pyrrolo-cytosine (pyrrolo-dC), which is a methyl derivative of dF* (Figure 1-2). The nucleoside analogue pyrrolo-dC is capable of hydrogen bonding with native bases and has been shown to selectively hybridize with guanine. Further, the stability of duplex DNA containing pyrrolo-dC is virtually the same as duplex DNA containing the native dC base.
The pyrrolo-dC base analogue can be selectively excited amongst native nucleobases because its low-energy absorption band is at 350nm, red-shifted far from the wavelengths where both nucleic acid and protein absorb. Similar to other FBAs the extinction coefficient is small $\varepsilon_{350} = 5900 \text{ M}^{-1} \text{ cm}^{-1}$ compared to external fluorescent probes. There is a large Stokes’ shift for the emission of pyrrolo-dC monomer $\text{em} = 460 \text{ nm}$ and when incorporated into DNA the emission is shifted further to 473 nm. The quantum yield of the pyrrolo-dC monomer has been reported as 0.2, significantly reduced compared to external dyes. Similar to 2-AP, when incorporated into single-stranded DNA there is a substantial drop in quantum yield, which is further reduced by hybridization with the complementary strand to values ranging from 0.03 to 0.05. There is little information pertaining to the relationship between neighboring bases and pyrrolo-dC photophysics. However, limited quantum yield data suggests neighboring bases do not influence pyrrolo-dC photophysics.

Pyrrolo-dC is commercially available and has been used in biochemical systems to characterize protein-DNA interactions. In one study the cytosine analogue was used to characterize the transcription bubble within the elongation complexes of the T7 RNA polymerase. The high degree of quenching upon hybridization has been used to examine mismatch and selective hybridization within a DNA oligomer. In another study, the cytosine analogue pyrrolo-dC was used to examine the kinetics for the human DNA repair enzyme alkyltransferase. Further, hybridizing a DNA strand containing pyrrolo-dC with an RNA strand to form a DNA/RNA hybrid was utilized to investigate the HIV-1 polypurine tract, which is thought to have an abnormal geometry. The fluorescent probe has been used in several other studies including investigation of hairpin stability, utilized in conjunction with
2-AP to form a molecular beacon, and screening DNA sequences for high-affinity and highly-selective triplex forming oligonucleotides.\textsuperscript{46-48} The above research demonstrates the utility of pyrrolo-dC for certain experiments, however the reduced quantum yield and limited photophysical characterization have restricted the broader application of pyrrolo-dC to biological systems.

\textit{Tricyclic cytosine an additional cytosine analogue}

The tricyclic cytosine analogue tC is one of the larger FBAs that is commercially available (Figure 1-2). It was first developed for use in antisense RNA hybridization experiments as it enhances the thermostability of double-stranded RNA and preferentially base-pairs with guanine.\textsuperscript{49} This FBA is highly fluorescent, there is extensive photophysical characterization, and has several homologues.\textsuperscript{50} The family includes the oxo-homologue of tC named tC\textsubscript{O} and the nitro containing homologue tC\textsubscript{nitro}.\textsuperscript{2}

Similar to other FBAs the low-energy absorption band for tC (centered at 395nm) and tC\textsubscript{O} (centered at 365nm) is red-shifted far from DNA and protein absorption wavelengths. In addition, there is a large Stokes’ shift for tC and tC\textsubscript{O}, the fluorescent emission maximum is 505 nm and 455 nm respectively. The FBA tC differs from other nucleoside analogs in several important ways. First, the quantum yield of tC in single or double-stranded DNA is not affected by surrounding bases, ranging from 0.17 to 0.24, considerably higher than 2-AP and pyrrolo-cytosine when incorporated into DNA.\textsuperscript{51} Furthermore, when tC is incorporated into single or double-stranded DNA its fluorescent lifetime is a single exponential decay ($\tau$\textsubscript{ss}) = 5.7 ns and $\tau$\textsubscript{ds} = 6.3 ns), indicating tC adopts a single conformation in DNA. The time-resolved anisotropy decay for double-stranded DNA containing tC reveals minimal local
motion, supporting the hypothesis of a single conformational population. In addition, the high molar absorptivity for the low-energy transition of tC\textsuperscript O (\(\varepsilon = 9000 \text{ M}^{-1} \text{ cm}^{-1}\)) is larger than most other FBAs. Lastly, tC\textsuperscript O is reported as one of the brightest nucleoside analog, as it can be up to 10-50 times brighter than other fluorescent base analogues such as 2-AP, 3-methylisoanthopterin (3-MI) and 4-amino-6-methyl-8-(2'-deoxy-\(\beta\)-d-ribofuranosyl)-7(8H)-pteridone (6MAP) in duplex DNA. Not only are its photophysical properties suitable for performing fluorescence spectroscopy experiments, the cytosine analogues are also well characterized.

NMR and circular dichroism (CD) data indicate that double-stranded DNA containing the analogue tC adopts a canonical B-form conformation and minimal local distortions occur in the immediate proximity of tC\textsuperscript 2. The NMR data also indicates that its base pairing and a base-flipping rates in double-stranded DNA are similar to the native cytosine base. Interestingly, the cytosine analogue tC increases the stability of a 10-mer duplex by approximately 3 °C, possible due to the large planar molecular geometry (Figure 1-2).\textsuperscript 52 In addition to the structural data, the cytosine analogue tC has been used to characterize biological systems.

The high quantum yield and single-exponential decay for tC is ideal for FRET experiments. The cytosine analogue tC has been utilized as the FRET donor paired with TAMRA to examine distances in a peptide-nucleic acid (PNA)-DNA hybrid and paired with ALexa-555 to study the conformational dynamics of DNA polymerase.\textsuperscript 2,53 In another polymerase study, the 5'-triphosphate of the cytosine analogue tC was efficiently incorporated into DNA by the Klenow fragment indicating the large size did not inhibit
enzyme activity. Interestingly, in the same type of experiment tC was misincorporated opposite adenine, possible due to the shape of the FBA. Aside from the above reported experiments there are few other studies using the cytosine analogue in biological systems. Further biological experiments are necessary to determine if the size and shape of this FBA inhibits protein interactions with labeled DNA in other systems.

*The pteridines a class of guanine and adenine analogues*

The pteridines are a class of fluorescent base analogues developed mainly by Hawkins and Pfleiderer (Figure 1-2). There are several analogues in this family which have shown great promise including the adenine analogues 6MAP and 4-amino-2,6-dimethyl-8-(2'-deoxy-β-D-ribofuranosyl)-7(8H)-pteridone (DMAP), and the guanine analogues 6-methylisoxanthopterin (6-MI) and 3-MI (Figure 1-2). The family is characterized by the pteridine molecule, which is a chemical compound composed of fused pyrimidine and pyrazine rings. The members are distinguished by a variety of substitutions on the pteridine ring. Similar to 2-AP and pyrrolo-dC, the photophysical properties of pteridine analogues are highly sensitive to the microenvironment. The quantum yield of pteridines ranges from <0.01 to 0.5 depending on the identity of the surrounding base pairs and whether the DNA is single-stranded or double-stranded. Interestingly, in addition to the base context, the absorption and emission spectra of 6-MI is sensitive to pH. The lowest energy absorption band for the pteridines ranges from 310 to 350 nm, again well resolved from native bases. The molar absorptivity for this family of FBA has not been extensively investigated. However, there are several reported values that range from 5000 to 15000 M⁻¹ cm⁻¹ depending on solvent and analogue identity. For example, the molar absorptivity of 6-MAP in methanol is ε₃₅₁ = 8500 M⁻¹ cm⁻¹, while 3-MI in water is ε₃₅₁ = 13000 M⁻¹ cm⁻¹. One disadvantage to this family
of FBAs is, with the exception of 6-MI and 6MAP, there is a reduction in the thermostability of double-stranded DNA containing a pteridine.\textsuperscript{2,6} Examination of 3-MI and DMAP structure reveals the methyl group at position 3 and 2 respectively hinders proper hydrogen bonding with their respective complementary base (Figure 1-2).

The use of the pteridine fluorescent nucleoside analogs in biological systems is not as prevalent as 2-AP but is greater than tC and pyrrolo-dC. There a number of studies which demonstrate their general utility. For example, 3-MI has been employed in a real-time assay for O\textsuperscript{6}-alkylguanine-DNA alkyltransferase, the study of HU non-specific binding to double-stranded DNA, and as a probe investigating hybridization specificity.\textsuperscript{58-60} In addition 3-MI has been evaluated as a possible probe for use in single-molecule experiments.\textsuperscript{57} The adenine analogue 6-MAP has also been used in several studies including characterization of the premelting transition in DNA A-tracts and examination of DNA photolyase induced base flipping within DNA.

6-MI has also been used in a number of studies of protein-nucleic acid interactions including investigations of the cleavage activity of HIV-1 integrase\textsuperscript{61} and probing the binding and destabilization of G tetrad structures by unwinding protein (UP1).\textsuperscript{62} Also, 6-MI was used to examine the dynamics of single-stranded DNA.\textsuperscript{62} In another study, the authors incorporated 6-MI into an oligomer to monitor the formation of a nucleoprotein filament between RecA and DNA.\textsuperscript{11} In RNA, 6-MI has proven a valuable guanine analog for probing base dynamics in the Tetrahymena group I ribozyme.\textsuperscript{63} 6-MI has several advantages over other nucleoside analogues including the thermostability of 6-MI in duplex DNA, its absorptive and emission properties, and its demonstrated utility in a number of biological
systems. One main drawback to using 6-MI is the 5 to 10 fold reduction in quantum yield when incorporated into DNA. After evaluating its advantages and disadvantages, 6-MI was chosen as a reporter molecule for the purpose of characterizing the biophysical properties of several DNA binding proteins. Before 6-MI could be utilized in a specific system, it was necessary to characterize 6-MI photophysical properties within a DNA construct in the absence of binding partners.

1.1.4 The photophysical properties of 6-methylisoxanthopterin (6-MI)

The first part of this thesis focuses on the characterization of 6-MI within a specific DNA sequence that displays several surprising fluorescent properties.Outlined below are the previously reported biochemical and photophysical properties of 6-MI. 6-MI is a fluorescent guanine analog that is able to hydrogen bond to cytosine in duplex DNA making it an excellent internal reporter of DNA dynamics.\(^6\)\(^4\) The low-energy absorption of 6MI (\(\text{ex}_{\text{max}}\) =340 nm, \(\varepsilon = 14125 \text{ M}^{-1} \text{ cm}^{-1}\))\(^42\) is shifted from the absorption wavelength of natural nucleic acid bases (260 nm), such that it is spectrally distinct from the rest of the DNA oligomer.\(^6\)\(^4\) Like 2-AP, the 6-MI monomer exhibits a high quantum yield \(\Phi = 0.7\).\(^6\)\(^4\) Incorporation of 6-MI into ssDNA leads to quenching of fluorescence, which in most cases is maintained or increased upon duplex formation. The degree of fluorescence quenching is dependent on the identity of the bases adjacent to the 6-MI, and fluorescence quenching is greatest when 6-MI is inserted between purines.\(^6\)\(^5\) Compared to 2-AP, the mechanism of 6-MI fluorescence quenching has been investigated to a lesser extent.

Theoretical calculations of 6-MI not incorporated into DNA (6-MI monomer) electronic states suggest a reversal of excited states (similar to 2-AP) as a cause for its high
quantum yield. For guanine (the closest structural homolog of 6-MI) the lowest excited state is of nπ* character, while for 6-MI it is of ππ* character.\textsuperscript{66} Discerning the mechanism for quenching of 6-MI fluorescence by adjacent bases is an area of active research, with several competing theories exist. The proposed mechanisms include but are not limited to re-ordering of excited states, intersystem crossing, exciplex formation, quenching through a conical intersection, or PET. In one study, it was proposed that interactions between 6-MI and neighboring bases within the DNA causes a re-ordering of excited state energies, such that the lowest excited state exhibits nπ* character.\textsuperscript{66} Another study proposed that PET is the main fluorescence quenching mechanism for pteridones (DMAP, 6MAP) and pteridines (3-MI, 6-MI) free in solution, noting that PET is energetically favorable for 6MAP*, DMAP*, and 3MI*, acting as excited state acceptors, quenching preferentially with purines.\textsuperscript{42} The study indicated that 6-MI had the slowest NMP quenching rate, which is in agreement with its calculated excited-state potential. While the excited-state reduction potentials of 2-AP and 6-MI are comparable, 2-AP* can undergo electron transfer with pyrimidines by oxidation, however, for 6-MI this pathway is thermodynamically unfavorable limiting PET quenching with cytosine and thymine.\textsuperscript{42} It is worth noting, these studies were conducted with free 6-MI in solution (6-MI monomer), and it’s not beyond reason for different behavior when incorporated into single-stranded or double-stranded DNA. It is possible that the geometric arrangement of 6-MI with adjacent bases will influence fluorescence quenching, particularly since all nucleobases exhibit a vector property, called electronic transition dipole moment (ETDM), whose relative orientation with respect to adjacent nucleobases is an important factor in fluorescence quenching. The vector defining the ETDM for 6-MI has been calculated, and the lowest energy transition (υ=29687 cm\textsuperscript{-1}) is in the plane of the FBA and lies roughly perpendicular to the glycosidic bond (Figure 1-3).\textsuperscript{67}
Figure 1-3 the first electronic transition dipole moments for 6-MI (A) and 2-AP (B)

*Adapted from Widom et al.$^{67}$

**Adapted from Holém et al.$^{68}$.
Interestingly, the orientation of the lowest energy ETDM for 6-MI is very similar to that for 2-AP, indicating the orientation of the ETDM within DNA is similar for both FBAs in the absence of steric constraints on geometry (Figure 1-3).\textsuperscript{67,68}

The excited state lifetime of the 6-MI monomer is \(~6.5\text{ns}\) and decays as a single exponential, but when incorporated into either ss- or dsDNA the decay is multi-exponential.\textsuperscript{65} This multi-exponential fluorescence decay has been modeled as three lifetime components of \(~500\text{ps}\), \(~2\text{ns}\), and \(~6.5\text{ns}\).\textsuperscript{65} Each lifetime component is ascribed to a subpopulation of structures with different quenching characteristics that do not interconvert rapidly on the time scale of the excited state lifetime. The fractional subpopulation with a \(~500\text{ps}\) lifetime is suggested to arise from 6-MI stacking with adjacent bases, the smallest fractional subpopulation (6.5ns lifetime) is suggested to correspond to an extrahelical conformation of 6-MI, and the \(~2\text{ns}\) population is an intermediate conformation between the two states.\textsuperscript{65} Previous work probing DNA and RNA dynamics has demonstrated that duplex DNA is minimally distorted upon introduction of 6-MI and that the fluorescent properties of 6-MI are sensitive to local environment.\textsuperscript{64}

As stated earlier, an experimental problem encountered when using FBAs as fluorescent reporters in duplex DNA is the large reduction in quantum yield relative to the free monomer. The first part of this thesis focuses on the discovery and characterization of sequence dependent enhancement of 6-MI fluorescence upon formation of dsDNA.\textsuperscript{55} The enhancement of fluorescence is demonstrated to be a consequence of the specific sequence ATFAA (either 5’ or 3’ ordered), where F represents 6-MI. Importantly, enhanced fluorescence is maintained when the pentamer ATFAA is inserted into any longer sequence,
even into non-canonical DNA structures such as cruciform DNA. Based on time-resolved fluorescence and KI quenching experiments discussed in the next section, it was postulated a single solvent inaccessible conformation of 6-MI is responsible for the enhanced fluorescence. Accordingly, the time dependent anisotropy of 6-MI reveals reduced flexibility in the double-stranded pentamer sequence compared to sequences where 6-MI fluorescence was quenched. In addition, Knutson and co-workers have identified another sequence where duplex DNA formation results in enhanced 6-MI fluorescence.\textsuperscript{69} Knutson et al. proposed that the enhancement of fluorescence was the result of thymines flanking the 6-MI (i.e. GTFTG), which conformationally constrained the 6-MI. Similar to the ATFAA pentamer sequence, the flanking thymines prevent dynamic quenching with the adjacent bases. Interestingly, changing the n+2 guanines in GTFTG to adenine (ATFTA) significantly increases the quantum yield and alters the excited state decay to resemble the decay of 6-MI in the pentamer ATFAA. Thus, the adenine bases at the n+2 position are critical for enhancement of fluorescence.

In contrast to earlier work on the photophysical properties of 2-AP and 6-MI where the analog was fluorescently quenched in single-strand and duplex DNA, the pentamer (ATFAA/AFTA/ATFTA) represents a unique system where duplex formation leads to different and interesting photophysical properties. The first part of this thesis focuses on characterizing the unique photophysics of 6-MI, where small changes in sequence context cause large differences in its fluorescent properties. This information is subsequently used to provide insight into the mechanisms for enhancement and quenching of fluorescence in duplex DNA. In addition, this section highlights the relationship between 6-MI conformational flexibility and dynamic quenching of fluorescence, since it was postulated
that 6-MI enhanced fluorescence was the result of constraining 6-MI motion. To increase the conformational flexibility of 6-MI it was necessary to systematically introduce a base-bulge or DNA mismatch opposite the 6-MI and examined the effect on its photophysical properties, both in the enhanced pentamer sequence ATFAA/AAFTA/ATFTA and in the quenched sequence CAFTC. Photophysical differences between the two classes of sequences report on the physical basis of enhanced or quenched fluorescence within duplex DNA. Additionally, the role of sequence context on 6-MI fluorescence was investigated by measuring the fluorescent properties of the AAFX sequence, where X is one of the following bases: cytosine, guanine, thymine, along with several other sequences. Finally, molecular dynamics simulations were used to compare a molecular model of 6-MI within an enhanced versus a quenched sequence. The information generated by molecular dynamics simulations was combined with experimental data to create a model for how local DNA structure leads to enhanced nucleoside fluorescence in double-stranded DNA.

The second part of this thesis uses the improved model of 6-MI photophysical properties within DNA to examine a specific protein-DNA interaction. With an improved model available, it was possible to ask how protein binding changes the local conformation of the DNA oligomer and influences individual base dynamics. The sections below provide background information on the protein-DNA system.

1.1.5 Proteins associated with the bacterial chromosome

In all prokaryotic and eukaryotic cells, DNA polymers house genetic information that encodes for cellular functions. DNA polymers present a unique challenge for cells, since the size of the chromosome is frequently larger than the space it needs to occupy. For most
bacteria, the chromosome is ~1000-fold longer than the length of a single cell. To overcome this problem, the bacterial chromosome is organized into a highly compacted structure composed of relaxed and supercoiled DNA, proteins, and RNA. Early electron microscopy images of the bacterial chromosome indicated an unbroken rosette conformation with a compacted central core and 10 to 80 plectonemic loops protruding outward (reviewed in Toro and Shapiro). The need for DNA compaction opposes normal cellular processes that occur on DNA like replication, transcription, and DNA repair. Consistent with this idea plectonemic loops, which are less compact than the rest of the bacterial chromosome, contain transcriptionally active regions of the genome.

Nucleoid compaction in bacteria is achieved by a combination of DNA supercoiling and the association of a number of nucleoid proteins that bind DNA, such as HU, H-NS, IHF, FIS. In addition to their role in DNA compaction, nucleoid proteins also influence the expression of a large number of genes that are critical to the survival and growth of the cell, such as those involved in protein synthesis and metabolism. Additionally, nucleoid proteins play a role in the regulation of gene expression in response to environmental stimuli, such as temperature, pH, osmolarity, and virulence of the bacteria. Aside from their role in genome compaction and transcription, nucleoid proteins are also known to be accessory factors in DNA replication, recombination, repair, and transcription regulation. In some cases these proteins are also known to interact with RNA. Interestingly, in spite of their abundance and broad involvement in many cellular processes, HU, H-NS, IHF, among other DNA-associated nucleoid proteins are all nonessential in E. coli, which highlights their functional redundancy in vivo.
1.2 The nucleoid associated protein HU (Histone-like U93 protein)

In *E. coli*, the most abundant nucleoid protein by mass is called HU, which was initially named Factor U after being isolated from the strain U93. Factor U was renamed to HU upon the discovery that its physical properties resembled those of histones, which bind DNA and are essential for compacting chromosomes in eukaryotes. HU makes up 2 to 5 ng per μg total protein in *E. coli* with little variation over the life of the cell. Based on the approximate volume of a cell, HU exists at a cellular concentration of ~20 uM. This concentration is considerably greater than the reported dissociation constant of HU for structural nonspecific binding to duplex DNA. The steadiness in HU cellular concentration suggests an autoregulatory feedback mechanism where an excess accumulation of HU inhibits its own transcription. There are no known cofactors that specifically affect the binding of HU to DNA, therefore there are no known mechanisms for post-translational regulation of its activity. In general, HU has been shown to be evenly distributed throughout the bacterial nucleoid and is necessary for normal compaction of the bacterial chromosome.

There are certain characteristics of HU that distinguish it from other nucleoid associated proteins. First, HU is the most abundant of the nucleoid associated proteins. Second, it is phylogenetically conserved. Accordingly, HU and its homologs are found in every branch of the eubacterial kingdom and are present in a number of archaeabacteria, bacteriophages, and in an animal virus. There are even HU homologs that are located in organelles like chloroplasts, which also seem to organize DNA structurally.
1.2.1 The Cellular Functions of HU

HU was initially isolated from *E. coli* extracts based on its ability to bind to DNA-cellulose columns. Purified HU exhibited characteristics similar to eukaryotic histones and consistent with an architectural role in genome organization, including a high proportion of basic amino acids and the ability to stimulate the transcription of lambda bacteriophage DNA by *E. coli* RNA polymerase in vitro. Following its identification as a highly abundant nucleoid-associated protein in bacteria, new roles for HU related to the maintenance of genomic compaction and transcription regulation continued to emerge. For example, HU is required for transposition of the bacteriophage Mu in vitro. Specifically, HU facilitates the early stages of assembly for the stable synaptic complex (SCC), which is composed of a tetramer of Mu tranposase bound adjacent to the ends of the Mu gene. In the context of Mu transposition, HU performs a structural role that stabilized the SCC by binding to DNA between the Mu tranposase recognition sequences. Consistent with its architectural role in Mu transposition, HU has also been shown to function as an accessory protein in the bacterial nucleoid.

Many nucleoid associated proteins like HU modulate the topology of DNA. For example, HU can maintain negative supercoiling of plasmid DNA even when relaxed with topoisomerase I. Other data supports its ability to bind to and bend DNA, which is a relatively stiff polymer with a persistence length of >100 bp. In the absence of HU, DNA fragments less than 500 bp in length cannot be circularized by DNA ligase. However, HU readily facilitates circularization by DNA ligase when incubated with small 60-100 bp fragments of DNA. More directly, images from electron microscopy of HU-DNA
complexes demonstrate that HU contorts the DNA structure.\textsuperscript{96} Together, this work indicates that HU is able to bind and to stabilize distortions in DNA structure.

Certain cellular processes in bacteria benefit from the ability of HU to bind distorted DNA structures, including DNA repair and recombination. The role of HU in DNA repair is supported by its high affinity (K\textsubscript{D} = 1 nM) for Holliday Junction DNA, which is considered one of the key DNA intermediates in DNA repair, homologous recombination, and viral integration.\textsuperscript{101} HU can displace LexA, a repressor of SOS response genes, from its binding site in the bacterial genome and mutations to HU in vivo confer increased sensitivity to UV and γ irradiation, which underscore the importance of HU in DNA repair.\textsuperscript{102-104} In the context of DNA recombination, HU was shown to be an important accessory factor for the phase variation recombination system in S. typhimurium.\textsuperscript{105} As shown in Figure 1-4, for phase variation recombination the invertasome complex forms over two hix consensus sites (hix\textsubscript{R} – blue and hix\textsubscript{L} – green) that are separated by 100bp on the bacterial chromosome. In order to assemble the invertasome, the two hix sites must be brought together, which is promoted by the DNA-bending activity of HU. Interestingly, the requirement for HU decreases when the length between the hix sites increases, pointing to a direct functional role for HU assisting in the looping of the DNA (Figure 1).\textsuperscript{105} The “protein swap” experiments of Johnson and colleagues provided further proof of HU’s accessory role in invertasome formation. Johnson and colleagues replaced HU with other DNA-binding proteins unrelated in primary, secondary, or tertiary structure, such as eukaryotic proteins HMG1 and HMG2 and showed in vitro enhancement of invertasome formation.\textsuperscript{100,106}
A diagram highlighting HU’s role as an accessory factor in phase variation recombination. In this model the recombination is initiated with the binding of Hin to the \textit{HixR} and \textit{HixL} sites and Fis binding to enhancer. While no direct interaction between HU and Hin or Fis is observed it is necessary for the formation and stabilization of high-order protein-nucleic acid complex to complete the phase variation recombination.

\textit{Adapted from Dhar et al.}^{107}
HU also functions in regulating gene expression at the level of transcription and RNA stability. Depleting HU from *E. coli* prevents the transcriptional up-regulation of the *proVMX* operon in response to hyperosmolar environments; it also inhibits transcriptional repression of the *gal* operon by participating in the formation of the Gal repressosome.\textsuperscript{108-110} By using recombinant HU that had been chemically converted to a nuclease, Aki and Adhya demonstrated that HU is also required for Gal repression in vitro, and specifically bound between its two operator sequences O\textsubscript{E}–O\textsubscript{I}, but only in the presence of the Gal repressor (GalR). Based on this data, they proposed that HU bound DNA cooperatively with the Gal repressor, and that HU binding led to negatively supercoiling between the two Gal operator sites.\textsuperscript{111} More generally, HU promotes many protein-DNA interactions in the cell that require bending of the DNA; by selecting for and stabilizing distorted DNA structures, HU lowers the intracellular concentration of protein required to bind to consensus sites, a paradigm that has been observed for the Lac repressor, CRP protein, and IHF.\textsuperscript{112,113} In addition to its roles in transcriptional regulation, HU also controls gene expression by binding to RNA. In particular, the presence of HU affects translation of the stationary phase sigma factor RpoS, further supported by its ability to bind to RNA directly.\textsuperscript{80,114} HU affects genome organization and the regulation of gene expression in bacteria through its ability to interact non-specifically with nucleic acids.

*Double mutant hupAB (HU\textsuperscript{−}) alters cellular phenotype*

A single deletion of either of the HU-encoding genes *hupA* and *hupB* has very little impact on cell growth, presumably since each of the subunits can substitute for the other.\textsuperscript{88,115,116} The double mutant *hupAB* (HU\textsuperscript{−}) grows very slowly, impairing cellular functions such as DNA replication initiation, transposition, is sensitive to UV light, and other
biochemical processes. \(^{88,117-124}\) Interestingly, the growth defect seen in HU\(^{-}\) strains can be rescued by a null mutation to the gene gyrB encoding for Gyrase, a type II topoisomerase, consistent with a role for HU in DNA superhelicity. \(^{125}\) It is interesting to note that the HU\(^{-}\) mutation is lethal in *Bacillus subtilis* and other gram positive bacteria due to a lack of other nucleoid-associated proteins. \(^{126-128}\)

One clear effect of *hupA/B* knock-out is the relaxation of supercoiled DNA in vivo, where the level of overall relaxation increases in the following order WT < *hupB* < *hupA* < *hupA/B*. Accordingly, the *hupA/B* knock-out causes genetic instability in bacteria, which allows for a rapid accumulation of suppressor mutations. \(^{88}\) In fact, microarray studies indicate that HU controls the transcription of 353 genes in *E. coli*, composing 229 independent operons over 8% of the *E. coli* genome, many related to stress response and metabolic regulation. \(^{79}\) In addition to the deleterious null phenotypes exhibited by HU knock-out strains, substitution of two amino acids in the *hupA* gene, E38K and V42L, causes a large number of expressed genes to be down-regulated and a number of genes that are silenced to be up-regulated. \(^{129}\) The mutant HU\(\alpha^{3842}\) is considered a gain of function mutation and causes visible phenotypic changes to cells including a morphological change from a rod to coccid shape and increased condensation of the bacterial chromosome. \(^{129}\) Studies of the affected genes implicate misregulation at the transcriptional level. \(^{130}\) Using data from a recent crystal structure of multimeric HU\(\alpha\beta\) (PDB ID = 2097) a model was proposed for the changes in gene expression. The crystal structure indicated four HU dimers associate to form an octameric complex, which can form higher order complexes characterized as both left and right-handed spiral filaments. The spiral filaments are proposed to enhance negative supercoiling of the bacterial chromosome. \(^{131}\) Based on this structure, the mutant HU\(\alpha^{3842}\) is
hypothesized to affect interactions between HU dimers shifting the equilibrium between left and right-handed spiral filaments, thereby promoting positive supercoiling in the DNA.\textsuperscript{116} Furthermore, activation of different subsets of protomers for WT HU relative to mutant HU\textsubscript{\alpha}\textsuperscript{3842}, consistent with negative or positive supercoiling facilitating different transcriptional activation.\textsuperscript{132}

The number of roles and diversity of functions that HU plays in maintaining the bacterial nucleoid are interesting given the small size of the protein and simplicity of its structure. In summary, HU is a nucleoid associated protein with a number of cellular functions related to the general organization of the bacterial chromosome. First, HU is known to introduce or to stabilize DNA negative supercoiling, which is critical for compaction of the chromosome. HU also functions as an architectural element in DNA replication, repair, and transcription, and in doing so regulates gene expression. Finally, HU acts as an architectural element in protein-DNA assemblies. The stiffness of DNA prevents protein binding sites that are separated by 30-300 bp from coming into close proximity to one another.\textsuperscript{98} By deforming the duplex, HU binding shortens the through-space distance between protein binding sites and facilitates protein complex assembly on consensus sequences.

\textit{Integration Host Factor (IHF) is a homolog of HU with similar cellular functions}

While HU binds to flexible structural motifs in DNA without preference for a specific sequence, IHF represents a second DNA-associated nucleoid protein in bacteria that binds with low nanomolar affinity to a specific DNA sequence that ranges in size from 28-30bp.\textsuperscript{133-137} The recognition sequence for IHF contains the consensus WATCARXXXXXTTR (W is A
or T; X is A, T, C, or G; R is A or G).\textsuperscript{138} and the specificity ratio for sequence-specific compared to non-specific IHF binding is roughly 1,000 to 10,000.\textsuperscript{139,140} The interaction between IHF and DNA is an example of an indirect readout mechanism, where a protein does not make base-specific contacts, but rather recognizes the sequence-dependent conformation of the DNA. IHF resembles HU in function, in that it acts as an architectural element that bends DNA allowing formation of higher-order nucleoprotein complexes.

Similar to HU, IHF facilitates DNA recombination events in bacteria as well as the assembly of protein complexes on DNA. For example, IHF plays an essential role in lambda phage site-specific integration recombination.\textsuperscript{141} It functions in the lytic cycle for bacteriophage MU, where mutation of IHF hinders early events in its lifecycle.\textsuperscript{142} Specifically, an IHF binding site is flanked by sequence determinants for a Mu phage repressor; IHF binding bends the DNA enhancing the stability of the bound repressor.\textsuperscript{143,144} IHF is also required for recombination by lambda phage, where it performs an architectural function by bending DNA to permit Integrase (Int) assembly, a catalytic protein involved in lambda site-specific recombination, at its consensus site.\textsuperscript{145-147} IHF is also required for viral integration by promoting the interaction between Int and the \textit{attP} site, the phage attachment site for viral integration, which spans 250bp and contains three IHF specific binding sites.\textsuperscript{148,149} In the absence of IHF, Int cannot assemble at its consensus site and initiate site-specific recombination. However, replacing the IHF consensus site with DNA modules that confer flexibility restores Int catalytic activity, consistent with an architectural role for IHF in recombination that enhances Int binding. Like HU, IHF performs important accessory functions in the context of DNA recombination.
Interestingly, HU can functionally substitute for IHF as an accessory protein for a number of cellular processes, pointing to a structural nonspecific architectural role for both proteins in vivo. For example, HU can replace IHF in DNA replication initiation of the oriC minichromosome and the transposon Tn10. In vitro, HU can substitute for IHF at the attL site for site-specific lambda phage recombination. In many cases a higher concentration of HU is required to recapitulate the activity of IHF due to its weaker affinity for duplex DNA. There are also cases where HU cannot replace IHF and vice versa; this is true for IHF binding at the attP intasome and for HU’s role in the Hin-promoted recombination. One key difference between IHF and HU is that the former recognizes a consensus sequence whereas HU binds tightly to flexible DNA motifs independent of sequence context. As such, HU acts as a general architectural element in many different cellular systems by bending the DNA to facilitate the formation of higher-order nucleo-protein complexes, while IHF functions in specific sequence contexts.

1.2.2 The structure-function relationship for HU and IHF

Purification of endogenously-coded HU protein from E. coli revealed that the protein exists as a dimer of ~9 kDa subunits. Since there are two isoforms of HU in E. coli, there are three different forms of the HU complex in cells: the two homodimers HUα2 and HUβ2, and a heterodimer HUαβ. The relative expression levels of each dimer varies with growth phase. It is worth noting, the three HU dimers (HUα2, HUβ2, and HUαβ) differ in their thermodynamic properties, and more importantly there is a reduction structural nonspecific double-stranded, and structure specific nicked/gapped DNA binding affinity for HUβ2. Furthermore, the homodimer HUβ2 lacks the ability to produce negative supercoiling in vitro to DNA relaxed in the presence of topoisomerase I. Like HU, IHF is
a dimeric protein encoded by two subunits, but unlike HU it is only found as a heterodimer since homo-dimerization is very weak. This differential ability to homo- and heterodimerize may play a role in their different binding specificities for DNA, as IHF binds a particular consensus sequence whereas HU binds DNA without sequence preference. Dimer formation in the HU/IHF family does not appear to be promiscuous. For example, there is no evidence for dimers involving one IHF subunit and one HU subunit nor have mixed dimers involving subunits from different bacterial species been reported.

Sequence alignment of the *E. coli* HU and IHF subunits reveals high sequence conservation between the proteins (Figure 1-5). Specifically, the primary sequence of the subunits of IHF and HU can be aligned without gaps and reveal 45% identity between sequences, including certain residues critical for DNA binding. The homodimeric structure of HU from *Bacillus stearothermophilus* (*B. st.* ) solved by NMR and X-ray crystallography indicates that each subunit of the highly basic protein is composed of three α-helices and two β-sheets. The N-terminus of the protein forms two helices named α₁ and α₂ that are joined by a broad turn creating a V-shaped structure. The C-terminus of the protein is composed mainly of three anti-parallel β-sheets that span the top of the V formed by the two N-terminal helices. A sharp turn links β-strands 1 and 2; this turn is highly conserved among HU/IHF homologs and is thought to be important for DNA binding. An extended arm region connects the β-strands 2 and 3, and the monomer terminates after β strand 3 in the short α₃-helix. The dimerization of monomer subunits is due to hydrophobic interactions between the pair of N-terminal V-shaped α-helices, and the dimer interface is composed of eight hydrophobic residues with perfect conservation among HU/IHF family members. The carboxyl terminal residues make up the other part of the alpha helical core. It is worth noting
Figure 1-5 HU and IHF sequence alignment

A comparison of the amino acid sequences from bacterial proteins subunits, in the “histone-like” family E. coli. HU, B. steaordermophilus (B. str.) HU, and Integration Host Factor (IHF) E. coli.72,159 The single letters represent the amino acid sequence. The bold letters indicate conserved amino acids. The rectangles and arrows represent the regions that form α-helices and β-sheets, while the lines signify turns. It is worth noting, the α-helices make up the core of these proteins, while the extended β-sheets compose the DNA binding region, which lies along the minor groove of the DNA and contains the DNA intercalating proline residues.
that HU crystallized in the absence of DNA is highly disordered at the ends of the β-ribbon arms allowing for conformational flexibility and possibly an induced fit mechanism for DNA binding.\textsuperscript{162}

The HUα\textsubscript{2} homodimer from cyanobacterium Anabaena (PDB ID = 1P71) has been crystallized with a pseudo self-complementary duplex containing three T:T mismatches and four unpaired T’s (Figure 1-6). Although such a sequence is highly unlikely to exist in a cell, the crystal structure reveals key features that drive the high affinity HU-DNA interaction. In the crystal structure, highly conserved prolines stack against unpaired thymines spaced 9 bp apart in the DNA duplex. This spacing is mirrored by the binding preference of HU for DNA containing mismatches spaced 9 bp apart.\textsuperscript{138,163-165} Even though the crystal contains an unnatural DNA sequence, the complementarity between solution binding affinities and crystallographic observations suggests that the crystal structure may represent the structural specific high-affinity binding mode for HU. Whichever binding mode the structure reflects, low-affinity structural non-specific versus high-affinity distortion-specific, the crystal demonstrates general features of the HU binding interaction that likely apply to both states.

HU exists as a heterodimer in E. coli, particularly during stationary phase. Unlike HUα\textsubscript{2}, which has been crystallized as a dimer bound to DNA, HUαβ has only been crystallized as a multimer. However, alignment between the HUα\textsubscript{2} protein backbone and one dimer of HUαβ (PDB ID = 2097) is very good, with an RMSD of only 0.56Å (Figure 1-6). This suggests that the DNA-bound structure of the HUα\textsubscript{2} homodimer is a good approximation for HUαβ, even though the HUαβ structure lacks density for the β-ribbon arms that bind DNA. It is worth noting of the differences between the crystal structures. For example, the
Figure 1-6 *Anabaena* HUα2 and crystal structure

(A) The crystal structure for homodimer HUαα from cyanobacterium *Anabaena* (PDB ID = 1P71) bound to a pseudo-self-complementary double-stranded DNA molecule. The crystal structure reveals the two intercalating prolines (dark blue and green) stack with unpaired thymines (purple) (B) Alignment of the protein backbones for the DNA bound HUαα homodimer (PDB = 1P71) (red) and one dimer of HUαβ (PDB ID = 2O97) (cyan) reveals a high degree of structural similarity; the RMSD for the alignment of the Cα backbone between the two proteins is 0.56Å.
intersubunit interactions for HUαβ dimer in the multiprotein complex are asymmetric in contrast to the symmetric intersubunit interactions for HUα2 homodimer.\textsuperscript{166} Based on crystal packing, the authors suggest that the asymmetric dimer interactions are critical for forming higher-ordered nucleoprotein complexes.

In the DNA-bound state, HU (PDB ID = 1P71) and IHF (PDB ID = 1IHF) exhibit remarkable structural similarity. The RMSD between Cα backbones for the two structures is only 1.35 Å (Figure 1-7).\textsuperscript{164,165} In further agreement, the crystal structures for HU and IHF both contain the protein interacting with two adjacent DNA molecules such that continuous chains are formed throughout the crystal. There are few hydrophobic or electrostatic interactions with the actual bases in the crystal structures for IHF and HU, consistent with an indirect readout mechanism for binding.\textsuperscript{164,165} The stabilizing interactions between HU/IHF and the DNA can be separated into several key components. The next few paragraphs compare and contrast the distinguishing features of HU- and IHF- DNA bound crystal structures that are important for stabilizing protein-DNA interactions.

\textit{B-ribbon prolines intercalate and stack with a nucleobase}

In the IHF/HU crystal structures, the β-ribbon arms wrap around the DNA minor groove (Figure 1-8). Each arm contains a single conserved proline residue per monomer (HU P63, P63/IHF P64, P65), which intercalates between base pairs, creating and/or stabilizing two kinks in DNA upon binding.\textsuperscript{165} The NMR structure for \textit{B. st.} HU\textit{αα} in the unbound state indicates that the prolines in the β-sheet arm are flipped out toward solvent, making them
The crystal structure for IHF bound to DNA (PDB ID = 1IHF) reveals a conformation similar to HU. Two prolines (magenta) intercalate into the DNA. The IHF consensus sequence is shown in yellow. The RMSD between Cα backbones for the two structures IHF (red) and HU (cyan) is only 1.303 Å.
readily accessible for intercalation into DNA.\textsuperscript{160} In crystal structures for both IHF and HU, the two prolines intercalate into the DNA and in the HU structure they stack with the unpaired thymines, causing kinks in the DNA to form at bases adjacent to the intercalation sites (Figure 1-8). In order to bend DNA sharply, the proteins must balance the free energy penalty for kinking the DNA by making favorable electrostatic interactions with the DNA.\textsuperscript{164,165}

\textit{Electrostatic interactions stabilize the nucleoprotein complex}

The electrostatic surface of HU and IHF are both positively charged at the DNA-binding site, which extends entirely down the sides of each protein (Figure 1-8).\textsuperscript{165} The β-ribbon arms contain a total of 8 arginine and 8 lysine residues, which form electrostatic interactions with the DNA backbone. The positively charged surface of the body is composed of the N-terminus of all three helices and numerous positively charged side chains. In the IHF crystal structure, the body of the protein interacts with the DNA phosphate backbone mainly through positively charged R-groups and the N-terminus of helices 1 and 3.\textsuperscript{164} Unlike IHF, crystal packing and the length of the DNA in the HU crystal structure limits the number of observed electrostatic interactions made with the sides of the protein. Another difference between the crystal structures is the bonding partner for Lys63, a widely conserved residue located on the N-terminus of helix 1. Lys63 forms a hydrogen bond with Thr26 in the HU structure rather than a salt bridge with the phosphate backbone, which may explain its role in determining binding site size for HU. Importantly, a longer length of DNA may have facilitated this interaction in the IHF crystal.
Figure 1-8 HU-DNA interactions

(A) In the HU crystal structure two prolines stack with the unpaired thymines, causing kinks in the DNA to form at bases adjacent to the intercalation sites. (B) The electrostatic surface of HU and IHF are both positively charged at the DNA-binding site, which extends entirely down the sides of each protein.
**HU induced DNA bend is smaller compared to IHF**

Since HU has been co-crystallized with three similar DNA constructs, bend angle comparisons can be drawn not only between HU and IHF, but also between the different crystal forms of HU. The DNA used in the *Anabaena* HU co-crystal structures share most primary sequence elements but differ in the number of mismatched base pairs and the presence of a 3’ overhang. In the absence of overhangs, HU-bound DNA adopts a bend angle of 105-106°, regardless of the presence of a mismatched thymine. In contrast, a co-crystal structure of HU with DNA containing a 3’ overhang crystallized in a different space group and adopted a larger bend angle of 139°. By comparison IHF was bound to a single construct and exhibited a larger single bend angle of 163° (Figure 1-6 and 1-7). In addition, the HU-DNA crystal structures were asymmetric with regards to the magnitude of the bend on each side of the protein. In fact, the bend magnitude varied not only within a single complex but also between crystal forms, which may reflect the flexibility of this complex in solution. This spread in bend angle supports a general role for HU in defining DNA architecture. These distinct DNA-bound features observed in crystal structures – conformational heterogeneity of HU and the large bend angle stabilized by IHF – are consistent with measurements obtained in solution. In the case of IHF, the large 160° bend angle is supported by bulk FRET measurements. Additionally, IHF has been shown to distort the shape of duplex DNA by 150° based on GMSA and electron microscopy. The relationship between HU DNA bend angle and solution structure is not as straightforward, and the results show a greater distribution of values. The experimentally measured HU bend angle, discussed in greater detail below, ranges from 25-150°, and depends on the DNA
construct and HU concentration.\textsuperscript{173-177} The flexibility in DNA bending may allow HU to participate in a wide variety of roles \textit{in vivo}.

\textit{HU untwists DNA to a greater degree than IHF}

Based on its crystal structure bound to DNA, HU introduces both undertwisting and negative writhe into the DNA structure. The average twist angle for relaxed B-DNA is \(\sim 34.3^\circ/\text{bp}\), while the twist for the IHF crystal is \(34^\circ/\text{bp}\) and the twist angle for HU crystal is \(31^\circ/\text{bp}\).\textsuperscript{165} The difference in undertwisting between IHF and HU is due to variation in DNA contacts with the sides of the proteins, possibly related to the different binding mechanism employed by the two proteins. In both structures the undertwisting that occurs at the kinks in the DNA is partially compensated by overtwisting between the kinks in the DNA.\textsuperscript{165} In addition, for IHF further overtwisting occurs at sites where the protein body is bound to the minor groove of the DNA. It is worth noting, that shortening the DNA in the IHF crystal structure to the length in the HU structure reduces the overall twist angle to \(31^\circ/\text{bp}\), which matches the measurement from the HU structure.\textsuperscript{165} Importantly, undertwisting the DNA is important for stabilization of negatively supercoiled DNA, an important function of HU \textit{in vivo}. In a contrast to HU, IHF has been shown to have little effect on the chromosomal DNA topology with similar affinity for binding sites in B-DNA or supercoiled DNA.\textsuperscript{178} It is postulated that distortion to the DNA writhe and twist induced by IHF is largely a self-canceling effect, therefore no untwisting of DNA is observed.

In summary, the crystal structures of HU reveal several key features of the protein-DNA binding interactions that allow it to associate non-specifically with duplex DNA, but also to bind to flexible DNA structures. In the crystal structure, two prolines intercalate and
stack with unpaired thymines that are nine base pairs apart. The intercalating proline results in a DNA kink between the unpaired thymine and the adjacent 3 prime nucleic acid base. HU utilizes electrostatic interactions between the DNA backbone and the N-termini of its helical core, positively charged side chains in the β- sheet arms, and the body of the protein to compensate for the free energy penalty for kinking the DNA. HU and IHF recognize different DNA structural motifs, which may be explained by the dissimilarities between the HU and IHF crystal structures. For example, the HU-induced DNA bend angle is smaller, there are fewer electrostatic interactions between the DNA and protein, and there is greater conformational heterogeneity between HU crystal structures. The distributions in HU-induced DNA bend angles (100 to 140°) and general conformational variability may be related to the two different binding modes HU has for DNA, which differ in affinity and structural epitope. Based on the crystal structure heterogeneity, these binding modes may also differ in the magnitude of distortion introduced by HU to the DNA structure.

1.2.3 HU binds to DNA independent of sequence context:

Unlike other nucleoid associated proteins, HU binds to DNA with little preference for sequence, and a specific topology-independent sequence has yet to be identified for HU binding. Its association constant for structural nonspecific binding to duplex DNA has been quantitatively measured to be $K_A = 0.08 \text{ uM}^{-1}$ (under stringent conditions 200 mM NaCl) using gel mobility shift assay (GMSA).\textsuperscript{179} The binding of HU to double-stranded DNA, double-stranded RNA, RNA-DNA hybrids, and single-stranded DNA is considered structural nonspecific because of the strong dependence of affinity on salt concentration and because of its comparatively low affinity for DNA.\textsuperscript{117,180,181} The affinity of HU is 3-fold higher for negatively supercoiled DNA ($K_A = 2.22 \text{ uM}^{-1}$) and increased further, ranging from 100 uM$^{-1}$
to 1000 uM$^{-1}$, for flexible DNA structures such as nicked DNA, junction DNA, bent DNA, single-stranded and double-stranded forks, or DNA overhangs.\textsuperscript{156,102,103,173,179,182,183} To summarize, the HU apparent association constants range from 1000 uM$^{-1}$ for structural specific binding to flexible motifs, to 0.04 uM$^{-1}$ for structural nonspecific binding to bacterial RNA.\textsuperscript{181} Based on these observations HU is thought to have two binding modes: i) structural nonspecific binding to single-strand DNA, double-strand DNA, and RNA with micromolar affinity characterized by strong sensitivity to salt concentration and ii) structural specific binding to flexible DNA conformations containing a single-strand/ double-strand junction such as nicks, gaps, overhangs, forks, spaced mismatches, and cruciform DNA characterized by picomolar to nanomolar affinities that are independent of salt concentration.\textsuperscript{80,114,132}

The ability of HU to maintain the global and local structure of the bacterial chromosome requires both binding modes. For example, as an architectural element HU participates in the global folding of the bacterial chromosome into a compact structure containing distinct topological domains, while at the same time dynamically altering the local structure for various cellular functions such as replication, transcription, recombination, repair, and gene regulation.\textsuperscript{184,185} To understand the mechanism of HU in nucleoid organization, a detailed model for each binding mode is necessary. To characterize HU-DNA interactions a considerable variety and number of experimental approaches have already been utilized which include GMSA, hydroxyl radical footprinting, \textit{Fluorescence Resonance Energy Transfer} (FRET), fluorescence anisotropy, ITC, X-ray crystallography, and single molecule techniques such as \textit{Atomic Force Microscopy} (AFM) and \textit{Force Extension Magnetic Tweezers} (MT).\textsuperscript{60,165,175-177,179} As might be expected based on the conformational variability present in co-crystal structures, the experimentally observed physical properties for HU-
bound DNA, such as DNA compaction and extension, binding site size, binding affinity and HU cooperativity, vary considerably. Also, more experimental work has focused on the characterization of the structural nonspecific binding mode for HU. In an effort to improve the HU binding model, the second part of this thesis describes experimental results characterizing HU structural specific binding. Specifically the relationship between HU-induced local changes in DNA conformational flexibility are correlated with global changes in DNA structure and dynamics.

The structural nonspecific binding mode of HU

The structural nonspecific binding of HU to double-stranded DNA is thought to influence bacterial chromosomal structure. In vitro, HU has no sequence preference in linear duplex DNA, consistent with an overall role in chromosome compaction.\textsuperscript{180} The structural nonspecific affinity of HU to linear duplex DNA was quantitatively measured under stringent conditions (200 mM NaCl) to be weak, $K_A = 0.08 \text{ uM}^{-1}$.\textsuperscript{179} Reducing the salt stringency to 50 mM NaCl increases the affinity of HU for duplex DNA $K_A = 0.71 \text{ uM}^{-1}$.\textsuperscript{60} The dependence of HU structural nonspecific binding affinity on salt concentration indicates electrostatic interactions are important for formation of a stable nucleoprotein complex. HU binding to single-stranded DNA is also strongly salt dependent. For example, GMSA data shows that HU binds a 40 nt single-stranded oligomer under low stringency (10 mM NaCl) with slightly weaker affinity ($K_{A_{app}} = 4.0 \text{ uM}^{-1}$) compared to double-stranded DNA ($K_{A_{app}} = 14.29 \text{ uM}^{-1}$), increasing the stringency to 150 mM NaCl reduces the binding affinity of HU for single-stranded DNA. In both cases the magnitude of the association constant is dependent on salt concentration.\textsuperscript{181} A model that accounts for this observation is that as the salt concentration
increases a greater number of counterions associate with the DNA phosphate backbone and charged amino acid side chains, effectively screening charge-charge interactions between the protein and DNA. An observation that emerges from this model is that the binding site size dictates the number of available charge-charge interactions between the protein and the DNA and therefore might affect the structural nonspecific binding affinity.

The structural nonspecific binding site size varies with DNA length and HU concentration

Determining the minimal binding site length for HU on double-stranded-DNA is difficult since it binds independent of sequence. In contrast, IHF binds to a specific sequence ensuring greater confidence in the predicted binding site size. For example, hydroxyl radical footprinting of the IHF-DNA complex indicates that the IHF binding site size is between 28-30 bp, with one NMR study indicating the minimal site is 19 bp. Also, the IHF-DNA crystal structure showed only a slightly larger binding site size of 35 bp. It is worth noting that IHF binding sites are mainly located in regulatory regions of the genome and are only modestly conserved, as the IHF consensus sequence contains only nine conserved bases that are clustered within one region of the recognition site. This asymmetric clustering indicates greater interactions with one subunit of the protein. Although the minimal binding site size for HU is unknown, structural similarities between IHF and HU may suggest a comparable binding site size between the two proteins. However, binding site size might be an important factor in separating their functions, since the two proteins recognize different DNA epitopes.

Compared to IHF, there is greater variability in experimental estimates of the HU nonspecific binding site size. For example, based on the HU-DNA crystal structure the
estimated binding site size ranges from 14 to 19 bp, whereas GMSA data for the HU/ double-stranded DNA complex estimates the binding site size is between 6 and 10 bp.\cite{165,97,180,188} This variability is not limited to HU homologues from \textit{E. coli} and \textit{Anabaena}; GMSA data for \textit{T. maritima} HU predicts the binding site size to be 37 base-pairs, contrasted with a binding site length of 9bp reported by surface plasmon resonance.\cite{189,190} It is worth noting that variability in the binding site size determined by GMSA may reflect misinterpretation of multiple bands as different HU-DNA stoichiometric ratios instead of multiple conformations of the same stoichiometry. Wojtuszewski and coworkers determined this to be the case when comparing data from GMSA to data from analytical ultracentrifugation and fluorescence anisotropy.\cite{60} They observed that the measured binding site size was correlated with duplex length. Specifically a 6.5 bp binding site was observed with a 13 bp duplex DNA and an 11 bp binding site was observed with a 34 bp duplex. DNA\cite{60} Koh and colleagues used ITC to further investigate the non-specific binding of HU to DNA oligomers of varying lengths, 8, 15, 34, 38, and 160 bp, at salt concentrations varying from 0.06-0.15 M Na$^+$.\cite{176,191} They interpreted the ITC data to reflect several binding modes with different thermodynamic properties. The authors modeled the data to contain three binding modes that required a footprint of 6 bp, 10 bp, or 34 bp of DNA, and proposed that the population involved in each binding mode was dependent on the ratio of the total HU concentration to the total DNA concentration ([HU]/[DNA]) and the length of the DNA fragment. Decreasing [HU]/[DNA] or increasing the DNA length drove the transition from the smaller binding-site size (6 bp or 10 bp) to the larger binding site size, which occluded 34 bp. Similarly, GMSA data for HU binding to single-stranded DNA revealed 24 nucleotides (nt) as the minimum length necessary for HU binding. The variability in the observed binding site size is an indication of HU’s promiscuity in binding non-specifically to DNA.
**HU structural nonspecific binding compacts DNA**

HU is known as an architectural protein that stabilizes bends in DNA. A number of experiments have shown that HU-induced distortion to duplex DNA is dependent on both HU concentration and the length of the DNA oligomer. For example Magnetic Tweezers (MT) experiments at increasing concentrations of HU protein showed mild compaction of the duplex DNA, measured by decreased persistence length at low protein concentration, which became more compact with increasing amounts of HU over concentrations ranging from 40 to 800 nM. At concentrations of HU above 800nM, the persistence length began to increase.\(^\text{177}\) Consistent with the MT data, AFM images display a salt-dependent, highly cooperative, and rigid nucleoprotein filament with a periodicity of 16 nm that appears at high HU concentrations (900 nM of HU or \(~1\) dimer per 1.8bp).\(^\text{177}\) However, the existence and biological relevance of this rigid nucleoprotein filament is controversial. A recent MT experiment measured persistence length of DNA in the presence of different HU and salt concentrations and only observed DNA compaction, at high concentrations of HU.\(^\text{192}\) Also, based on the HU concentration present in the cell there is only \(~1\) dimer per 300 to 400 basepairs, indicating the nucleoprotein filament may be physiologically irrelevant.\(^\text{193}\) Finally, IHF does not participate in DNA extension or form a rigid nucleoprotein filament upon binding DNA, even though it shares extensive sequence and structural homology with HU.\(^\text{177}\)

Although both IHF and HU cause 30% compaction of DNA in MT experiments, the physical properties of the respective protein-DNA complexes differ under mechanical force. HU is able to stably associate with DNA in the presence of high force unlike IHF, where high force leads to elongation of the DNA purportedly caused by protein dissociation. The authors
suggest that this difference is due to the flexibility of HU induced bends, which can accommodate force induced extension. Accordingly, HU binding increases thermal fluctuations of end to end DNA length by 61% rms, which are directly proportional to the derivative of the stiffness of the DNA tether and therefore a direct measure of the increased HU-induced DNA flexibility. HU not only induces DNA compaction but also affects DNA dynamics, effectively increasing the rate of bending. The HU induced compaction of the DNA is thought to arise from local bending of the DNA.

**HU achieves DNA compaction by inducing sharp bends in the polymer**

Measurements of the bend angle formed by HU-bound DNA range from 25 to 150°. Although this certainly reflects experimental differences such as the experimental technique and salt concentration, it also alludes to the inherent flexibility of the HU-DNA interaction and the resulting heterogeneity of the populated complexes. For example, ensemble FRET experiments using end labeled 55-mer duplexes measured a bend angle of ~50°/dimer, which agreed with single molecule FRET results obtained by the same research group. However, other bulk FRET measurements using a shorter duplex reported more complicated binding behavior, which was biphasic and could be modeled with three different bending modes, one corresponding to a bend angle of 143 ± 6° and two others that produced a decrease in energy transfer. AFM has also been used to capture the ensemble of structures created when HU binds non-specifically to DNA. At low concentrations of HU (18 nM) incubated with long DNA fragments of 1000 bp (~1 dimer/92 bp) vanNoort and colleagues observed a wide distribution of DNA bend angles with a large percentage of the distribution located between 100-150°. HU can also bend single-stranded DNA, as titrating the protein into labeled 28 nt and 38 nt single-stranded DNA produces an increase in energy transfer.
current literature indicates that the structural nonspecific binding mode of HU bends DNA, and that there is considerable variability in reported bend angles. The wide spread in measured DNA bend angles likely echoes the intrinsic flexibility of this interaction, which may reflect the diverse roles that HU plays in the cell, each requiring a different degree of DNA bending.

**HU structural nonspecific binding is driven by electrostatic interactions**

The strong dependence of HU binding affinity on salt concentration provides clear evidence that electrostatic interactions are important for its structural nonspecific binding mode. Many research groups have demonstrated this property. First, the dissociation behavior of HU-DNA complexes in single molecule force extension experiments is highly sensitive to salt concentration. At high ionic strength (150 mM NaCl) the dissociation rate is force-dependent; increasing the applied force from 0.1 pN to 0.3 pN results in ~3-fold reduction in nucleoprotein complex lifetime. At lower salt concentrations (40 mM NaCl) a “stalled” complex forms, which does not dissociate from DNA even at high forces.192 Second, ITC has been used to dissect the role of electrostatic interactions in the HU-DNA complex.176,191 Just like in bulk FRET studies, ITC data collected by titrating HU into a solution of DNA duplex can be fit using three independent binding modes, which are all driven by electrostatic interactions. Presumably, HU replaces the salt ions that neutralize the phosphate backbone when it binds to DNA, which are released to bulk solvent. Since this process releases salt ions, all three binding modes are entropically driven. The authors also noted that the binding enthalpy for proteins containing a large amino acid interface with DNA decreases with higher salt concentrations, indicating exclusion of solvent between the protein and DNA in the bound state. Accordingly, the change in enthalpy for IHF binding is linearly
dependent on salt concentration with a positive slope ($\partial \Delta H_{\text{obs}}^o / \partial [\text{KCl}] = 38 \text{ kcal mol}^{-1} \text{ M}^{-1}$).\textsuperscript{195} The enthalpies for the three binding modes of HU are also linearly dependent on salt concentration, but $\partial \Delta H_{\text{obs}}^o / \partial [\text{KCl}]$ slightly increases with decreasing DNA footprint, as the 34 bp binding mode exhibits a slope of 11.2 kcal mol$^{-1}$ M$^{-1}$, the 10 bp binding mode exhibits a slope of 7.6 kcal mol$^{-1}$ M$^{-1}$, and the 6 bp binding mode exhibits a slope of 5.4 kcal mol$^{-1}$ M$^{-1}$.

Since the magnitude of this slope is proportional to the amount of surface area that is buried in the protein-DNA complex, the IHF-DNA complex buries more than double the surface area of the HU-DNA complex. This observation is consistent with MT and bulk FRET experiments, where HU does not wrap the DNA as tightly and forms a heterogeneous population of bend angles in complex with DNA. However, in both cases the electrostatic interactions are critical for stabilizing the nucleoprotein complex.

In summary, the structural nonspecific binding mode between HU and duplex DNA is intrinsically heterogeneous, with considerable variability in the binding site size, bend angle, and thermodynamics that depends on both protein and salt concentrations. Electrostatic interactions drive the association between HU and DNA, representing a possible link between DNA length, HU concentration, and protein binding mode. Another important conclusion is that structural nonspecific binding by HU to double-stranded DNA is accompanied by DNA compaction, which may represent the physiological role of this binding mode.\textsuperscript{88,97,119}

\textit{HU recognizes distorted DNA conformations with high affinity}
While HU binds duplex DNA with weak affinity that depends on salt concentration, it can also populate a structural specific binding mode characterized by tight affinity that is independent of salt concentration. The first indication of an HU structural specific binding mode came while screening *E. coli* extracts for the ability to form complexes with cruciform DNA by GMSA. HU was identified as a cruciform binding partner with equilibrium association constants ranging from 100-1000 uM\(^{-1}\). The affinity of HU for cruciform DNA was several orders of magnitude higher than that for duplex DNA, which the authors hypothesized reflected a difference in substrate flexibility. To test this hypothesis, the authors generated a flexible non-cruciform construct containing six unpaired adenines, and saw a 40-fold increase in association constant over double-stranded DNA. In addition, substituting 5-hydroxymethyluracil for thymine bases in a DNA duplex increase the equilibrium association constant from 0.08 uM\(^{-1}\) for double-stranded DNA to 38.5 uM\(^{-1}\) for the modified duplex. The association constant is further increased to 285 uM\(^{-1}\) when pairs of mismatched bases (T:T/T:T) are separated by 9 bp. HU also has higher affinity for A-tract DNA \(K_A = 6.10\) uM\(^{-1}\) compared to a random duplex of the same length \(K_A = 0.68\) uM\(^{-1}\) under moderately stringent conditions. A-tract DNA is known to exhibit greater curvature than linear DNA and accordingly is a better binding partner for HU. However, the equilibrium dissociation constant for A-tract sequence A\(_4\)T\(_4\) and non-phased sequence T\(_4\)A\(_4\), which in theory is not pre-bent, are within error of each other, requiring further investigation to explain this result. HU affinity is also increased 3-fold for negatively supercoiled DNA \(K_A = 2.22\) uM\(^{-1}\), which is known to be undertwisted. It is now known that two non-cooperative HU dimers bind to each cruciform molecule, binding opposite each other and bending the cruciform by ~60°. HU participates in cellular events that involve cruciform structures *in vivo*. HU promotes the cleavage of four-way junctions by phage T4
endonuclease VII, known as a DNA-junction resolving enzyme. Interestingly, the T4 endonuclease VII is known to bind to bent DNA, which may be stabilized structurally by HU binding at the junction.

Cruciform DNA contains a single-strand double-strand junction, which is a structure found in many double-stranded break repair, transcription, and DNA replication intermediates. HU exhibits similar association constants in the range of 50-1000 uM\(^{-1}\) for DNAs containing structural elements that resemble cruciform DNA like short gaps, forks, single-stranded overhangs, and cruciform DNA itself. This suggests that HU binds these flexible and structurally unusual DNA constructs using the same high affinity binding mode. Interestingly, GMSA data for DNA constructs containing overhangs indicates that the affinity varies based on the position of the overhang. The affinity of HU for a 3’ overhang is \(K_{A_{\text{app}}} = 62.5\ \text{uM}^{-1}\), which drops \(K_{A_{\text{app}}} = 1\ \text{uM}^{-1}\) for a 5’ overhang. This difference in affinity may reflect asymmetry in HU binding at ss/ds junctions, where only the 5’ double-stranded branch can interact with the HU body.

Unlike many other types of DNA binding proteins, HU specifically recognizes structural features in DNA rather than a specific sequence. The affinity between HU and DNA increases as greater flexibility is introduced into the DNA. In fact, affinity can be ordered based on flexibility: double-stranded DNA < negatively supercoiled DNA < A-Tract DNA < 5-hydroxymethyluracil substituted DNA < DNA structural intermediates (Holliday junction, overhang etc.) == double TT:TT mismatch. The majority of constructs that are high affinity binding partners and also exist in vivo contain the same motif: single-stranded/double-stranded junction (ss/ds junction). This junction is hypothesized to exhibit increased
flexibility, critical to HU high affinity. The second part of this thesis focuses on HU’s direct interaction with the ss/ds junction, specifically addressing the propagation of local HU-induced perturbation to conformational dynamics throughout the DNA that lead to global bending of the DNA construct. The size of the HU specific binding-site size, the specific interactions with ss/ds junction, and the global distortion to DNA structure upon HU binding are already known. However, a detailed picture of the binding site dynamics and mechanism of DNA distortion is still lacking. There are several mechanisms that can explain how proteins induce distortion of DNA structures: i) protein binding and bending of DNA construct is concerted, ii) protein binding and bending of DNA is sequential proceeding through a transition state intermediate, and iii) protein binds to pre-bent constructs shifting the equilibrium from straight to pre-bent. An interesting question is which mechanism best describes HU. Discussed below is an overview of the current literature related to the HU high affinity binding mode.

_HU structural specific binding site size is independent of DNA length_

While the length of the HU structural specific binding site is not precisely known, estimates can be made based on affinity measurements between HU and different DNA constructs. Although HU binds to ss/ds junctions, there is no data available that studies the way affinity varies with the length of the double-stranded portion. In contrast, one group has systematically varied the length of the 3’ single-stranded overhang and correlated single-stranded length with overall affinity. The researchers examined HU binding to an array of overhang constructs containing an invariant 20 bp duplex portion and 3’ single-strand overhang of varying size from 0 nt to 20 nt. HU affinity was similar for 3’ overhangs ranging in length from 8 to 20 nt, while HU affinity was significantly reduced when the
HU-binding DNA 3'-overhangs of different lengths of the ss-part (indicated at the bottom in nt), or DNA containing a nick (indicated as n), were analyzed. HU protein was added to a final concentration of 8 (left) or 25 nM (right) to 0.05 pmol of labeled DNA. Bound and free DNA were gel-separated, and binding constants of HU–DNA complexes were calculated.\(^{173}\)

Adapted from Kamashev et al.\(^{173}\)
single-stranded section was 4 to 6 nt in length, and no binding was detected at shorter lengths (Figure 1-9).\textsuperscript{103} Therefore HU requires a single-strand length of ~8 nt to maintain high affinity. Another study used phenanthroline footprinting to map the HU binding site using nicked constructs as well as cruciform DNA.\textsuperscript{173} They showed that HU protects a region of ~20 bp in both constructs, and that the protection pattern displayed asymmetric binding. A clever experiment to map the footprint of HU on DNA in a different way converted the protein into a nuclease and used cleavage patterns to improve the model of the HU-DNA complex.\textsuperscript{173} Data obtained by nuclease conversion agreed with phenanthroline footprinting data for both nicked and cruciform constructs, suggesting that HUαβ binds asymmetrically with the β-hairpin arms running antiparallel to the broken strand. Also, under stringent salt conditions where the junction is known to be folded, HU protects only two of the four cruciform strands from hydroxyl radicals with greatest protection 2 bp from junction.\textsuperscript{179} These measurements match what is seen in a co-crystal structure between HU and a pseudo-self-complementary construct, where 9 bp separates the two intercalating prolines and the total contact length ranges from 14 to 19 bp.\textsuperscript{165} There are several problems with the footprinting and HU nuclease experiments. First, the protection and cleavage assays required isolating a stable complex from a gel, which may differ from the equilibrium solution structure. Secondly, HU was chemically converted into a nuclease at only three positions, which is not enough to generate a complete model of the nucleoprotein complex. The experimental results do provide information that can be coupled with other types of data to generate a better model.

In summary, the HU binding site size is hypothesized to be similar for all high affinity constructs. The crystal structure indicates a binding site size between 14 and 19 bp,
which agrees with phenanthroline footprinting and nuclease cleavage data. Additionally, HU requires a minimum of 8 nt of ssDNA to maintain high affinity binding with a d 3’ overhang. It is worth noting that the structural specific binding footprint for HU is similar in size to its structural nonspecific binding footprint, and is smaller than the sequence specific binding site size for IHF. This suggests that flexibility and not the total number of contacts is responsible for the high affinity of the structural specific HU binding mode.

*When HU binds distorted DNA it induces larger bend*

Structural specific binding of HU to flexible DNA motifs has the additional consequence of bending the DNA, and these bend angles have been measured in several ways. The mobility of HU bound to 40 bp DNA containing a nick migrates 10% faster than HU bound to 40 bp linear DNA, while the mobility of the unbound DNA is the same.\(^{173}\) The HU-nick DNA complex is thought to be more compact leading to quicker migration in the gel, signifying greater bending of the DNA. In fact, circular permutation assays with nicked DNA show that HU kinks the DNA at the nick by at least 65°.\(^{173}\) It is important to note, the bend angle determined from electrophoresis data and from the crystal structure may not reflect the solution structure.

The bend angle for HU specific binding in solution was investigated with a longer 55 bp DNA construct, which contained a high affinity binding site comprised of two T:T/T:T mismatches spaced 9 bp apart.\(^{163,174}\) Based on ensemble FRET measurements, the authors reported a bend angle of 145° at a 1:1 stoichiometric ratio, which is in good agreement with the calculated co-crystal structure bend angle.\(^{174}\) Additional HU causes a sharp decrease in DNA bending, as the HU begins to populate its non-specific binding mode.\(^{174}\) SmFRET data
for HU bound to the same construct revealed a single population distribution from 110 to 165 ° centered at ~145 °. In further similarity, increasing the HU concentration shifted the population to lower FRET caused by non-specific binding.\textsuperscript{165,174} One concern is the observed bend angle for the T:T mismatch construct may not represent the physiologically relevant high affinity motifs.

In addition HU affects the structure of cruciform DNA. Recent work studying the HU induced cruciform bend angle in solution has shown that HU stabilizes the stacked conformation of the junction without an appreciable increase in DNA bending. The calculated interduplex angle (IDA), referring to the angle between two arms of the junction, is 42.5±1.6° in 200 mM NaCl, which decreases to 37.9±0.9° in the presence of HU.\textsuperscript{196} At 30 mM NaCl where cruciform DNA is a heterogeneous mixture of open and stacked conformations, HU stabilizes the stacked conformation of the junction.\textsuperscript{196}

A-tract DNA is intrinsically flexible and also a high-affinity binding partner for HU. Therefore, it represents a good model system for studying the relationship between bending and specific binding. FRET has been used to examine the global bend angle of HU bound to doubly labeled A-tract DNA.\textsuperscript{175} By FRET measurements, a 20 bp A-tract sequence (A\textsubscript{4}T\textsubscript{4}) is 75.3 Å, whereas a non-A-tract (T\textsubscript{4}A\textsubscript{4}) 20 mer is slightly longer at 80.9 Å. This means that the A-tract sequence is pre-bent by 20-30°. When HU binds to either construct, the final bend angle increases to ~70° whether or not the A-tract is present. This demonstrates that DNA flexibility affects the overall affinity of the protein-DNA interaction, but not the final distortion angle.\textsuperscript{175}
Since HU can participate in two binding modes that exploit different structural features of the DNA polymer, it represents an interesting model system for understanding some of the fundamental principles that underlie protein-DNA interactions. In its structural nonspecific binding mode, the interaction between HU and DNA is relatively weak, heterogeneously populated, and very sensitive to ionic strength and component concentrations. In contrast, HU recognizes flexible motifs in DNA using a high affinity specific binding mode that is insensitive to the concentration of salt. In both binding modes HU only interacts with the DNA backbone and is entirely sequence-independent. Understanding the physical parameters that govern whether HU will participate in the structure specific versus structure nonspecific binding can be used to gain insight into the recognition mechanisms of other DNA-binding proteins. A number of models have been generated to describe the mechanism of HU specificity, but, more information is necessary to discern the correct model and fully describe the two types of HU-DNA interactions. The objective of the second half of this thesis is to improve the model for HU specific binding to flexible DNA motifs. A combination of techniques are used to this end, with an emphasis on using the FBA 6-MI to examine local HU induced perturbation to the DNA structure. In addition, fluorescence resonance energy transfer is employed to examine the global distortion to DNA structure and generate a solution model of HU binding to the overhang DNA motif.
2 Theory

2.1 Light interacting with macromolecules

To understand the interaction between light and biomolecules it is important to clearly define them. Light is generally defined as an electromagnetic field that is rapidly oscillating. Biomolecules can be defined as a nucleus surrounded by electron clouds with charge and spin. Both systems have electrical and magnetic properties, and therefore can interact with one another. During an experiment either a continuous wave or pulsed light is shined onto a sample. The light interacts with biomolecules in a specific state, which alters the state of the system and reciprocally can change fundamental properties of the light. This produces a number of measurable changes or “observables” such as: the wavelength of light that can interact with the biomolecular state, the wavelength or energy of the emitted radiation after biomolecular interaction, and the rate of change between biomolecular states. Discerning how these observables correlate with the physics of the system requires knowledge of biomolecular states.

Light facilitates the transition of a biomolecule from the ground state to an excited state in a process known as absorption. Environmental factors influence eigenstates of a system and therefore the potential energy gap between states, making absorption of light an excellent reporter on a biomolecule’s conformation and dynamics. Although biomolecules are very large and therefore are described by complicated wavefunctions, most biomolecules are composed of smaller components that can selectively absorb light. These components, such as
The Jablonski diagram represents the partial energy pathway of a fluorophore in different electronic states. The purple and blue vertical arrows represent absorption to an excited state. Red wavy arrows represent vibrational relaxation, green arrows represent fluorescence and black arrows represent competing non-radiative pathways.

Figure 2-1 Jablonski diagram
amino acid side chains or DNA nucleobases, are known as chromophores and can be considered molecular systems unto themselves. These are defined by simpler wavefunctions requiring a certain frequency of energy to transition to an excited state. These compounds are largely composed of conjugated pi-bonds whose energy states are sensitive to the local environment.

2.2 Absorption: transition from ground state to excited state

The excited state can be regarded as a new electronic distribution of the ground state with a different geometry. The new electronic geometry enables the excited state molecule to undergo chemical processes relative to the molecules in the ground state. The Jablonski diagram shown in Figure 2-1 illustrates the partial energy pathway for a luminescence system. Key features of this diagram are outlined in the following sections.

The energy of a photon required to transition a molecule from the ground state to a particular excited state is the difference in energy between the two states (depicted as vertical lines in Figure 2-1) given by the following equation:

$$\Delta E = h\nu = E_b - E_a$$

(2.1)

The potential energy diagram illustrates an electronic transition, with the horizontal axis representing internuclear distance and y-axis the energy of the system (Figure 2-2). The potential energy surface for the two lowest singlet electronic states, the ground state and first excited singlet state, are shown in (Figure 2-2). Superimposed on the potential energy surface are a series of vibrational modes $\nu$. In general, the energy gap between the lowest vibrational modes of the two electronic transitions is ~80 kcal mole$^{-1}$ and the energy spacing between modes is on the order of 2-10 kcal mole$^{-1}$, which means at room temperature the
The process of absorption is very rapid occurring on the femtosecond timescale, much faster than nuclear motions. After the transition, the nuclei realign with the excited electronic state through vibrational relaxation. The intensity of particular vibronic transitions is governed by the Frank-Condon principle, which states that during an electronic transition a change from one vibrational energy level to another is more likely for two vibrational wave functions with significant overlap (Figure 2-2). In practice this means that if the two electronic potentials are similar there is a propensity for $\Delta v = 0$ transitions, for instance $v''=0$ to $v'=0$. If the excited state geometry is distorted from the ground state there is a propensity for $\Delta v \neq 0$, i.e. transitions of $\pm 1$, $\pm 2$, $\pm 3$ etc. (Figure 2-2). In addition, while absorption to higher energy singlet states ($S_2$) readily occurs, the transition from the ground state ($S_0$) to the triplet states is symmetry forbidden due to differing electron spin multiplicities.

The absorption intensity is proportional to the concentration of molecules in a certain eigenstate. To appreciate the meaning of absorption it is useful to start with a simple system. Consider a system where a molecule can exist in two states, the ground-state $S_a$ and an excited state $S_b$ separated by an energy gap of $\Delta E = h\nu$. The two states differ in the electronic
Potential energy diagram of an electronic transition with photon coupling along the configuration coordinate, representing the normal modes. The upwards arrows represent absorption transitions and the downwards arrows represent the symmetric emission process.

**Figure 2-2 Potential energy diagram**
structure coupled by the dipole moment operator $\hat{\mu}$. The transition rate probability $dP_b/dt$, defines the rate of transformation from state $S_a$ to state $S_b$ in response to the light:\(^{18}\)

$$ \frac{dP_b}{dt} = \frac{1}{2\hbar^2} |< \Psi_b | \hat{\mu} | \Psi_a > |^2 \ast E_0 |^2 = B_{ab} I(\nu) $$

Where $B_{ab}$ is the transition rate per unit energy density of radiation, $\hbar$ represents Planck’s constant divided by $2\pi$, and $I(\nu)$ is the energy density of light at frequency $\nu$. The intensity of absorption is proportional to light stimulated rate of both the $a \rightarrow b$ and the $b \rightarrow a$ transition. The rate can then be defined in terms of energy density of light given by:

$$ -\frac{dI(\nu)}{dt} = h\nu (N_a B_{ab} - N_b B_{ba}) I(\nu) $$

Where $N_a$ and $N_b$ represent the number of molecules in states $a$ and $b$ per unit area $(cm^3)$ respectively. Equation 2.3 means that absorption of light is directly dependent on the sample concentration (terms $N_a$ and $N_b$). The quantities $B_{ab}$ and $B_{ba}$ are known as the Einstein coefficients representing the stimulated absorption and emission respectively.\(^{198}\) The transition probability $B_{ab}$ is directly proportional to $<\Psi_b | \hat{\mu} | \Psi_a >^2$. Light induces a dipole in the molecule ($\mu_{ind} = \vec{a} \cdot \vec{E}$), dependent on the electric field ($E$) and the polarizability of the molecule ($\vec{a}$). Thus, in equation 2.3 the intensity is proportional to the number of molecules interacting with $h\nu$. To calculate the concentration of absorbing molecules equation 2.3 is converted to represent $dI$ as a function of concentration given by:

$$ -\frac{dI}{l} = Ce dl $$

Where $C$ is the molecular concentration and $\varepsilon$ is called the molar extinction coefficient, a wavelength dependent parameter that describes the absorption spectrum and is independent of concentration for non-interacting molecules. Calculating the extinction coefficient for a given biomolecule requires measuring the absorbance of a sample of known concentration.
Integrating equation 2.4 and converting to log base 10 gives the Beer-Lambert Law for absorption:

\[ A(\lambda) = \log \left( \frac{I_0}{I} \right) = C\varepsilon(\lambda)l \]  

(2.5)

where \( A \) is the absorbance and \( l \) is the sample length. In this thesis the absorption spectrum is used to calculate sample concentration, determine percentage fluorophore labeling, and to characterize ground-state interactions of 6-MI with the surrounding bases and solvent. It is known that the absorption spectrum of 6-MI is sensitive to both the pH and solvent polarity.\(^{23,66}\) In addition the absorption spectrum of 6-MI is altered when located in single and double stranded DNA, which manifest as a reduction in absorption known as hypochromism.

Hypochromism is observed in intact double-stranded oligomers, in which absorption is 10-30% less than the corresponding mixture of complementary single-strands.\(^ {18}\) The origin of hypochromism is the intermolecular interaction between one particular electronic excited state of a given chromophore and different electronic states of the neighboring chromophore.\(^ {18}\) The interaction is described as the coupling between the transition dipoles of one chromophore and those of its neighbors. The transition dipole moment of the absorbing chromophore will interact with the induced dipoles of the neighboring chromophores. The orientation of the dipoles is critical for the hypochromism effect and can be used to examine transitions between single-stranded and double-stranded DNA.

### 2.3 Depopulating the excited state:

Fluorescent emission is an excellent indicator of a fluorophore’s environment. Interpretation of fluorescence requires knowledge of both the molecular excited and ground
states. Compared to absorption, the timescale of fluorescence (~10^{-8} s) is much longer. The slower time scale means a wide range of excited state processes can occur.

In equation 2.3 the transition rate from state $S_a$ to state $S_b$ is dependent on radiation density $I(v)$ and the $B_{ab}$ Einstein coefficient. The reverse process, stimulated emission from the excited state $S_b$ to the ground state $S_a$ can be described by a transition rate. The simplest case defines the transition as a radiation-induced process with the transition rate $B_{ba} = B_{ab}$. However, at equilibrium the two rates would be equal with the same number of molecules in each state ($n_a = n_b$) independent of radiation density, which is incorrect. Without light the majority of molecules will be in the ground state $S_a$ as determined by the Boltzmann distribution:

$$n_a/n_b = \exp^{-(E_a - E_b)/kT} = \exp^{hv/kT}$$  \hspace{1cm} (2.6)

Albert Einstein proposed an additional rate $A_{ba}$ for the spontaneous emission of a photon from state $S_b$. The rate $A_{ba}$ is independent of the radiation intensity $I(v)$. To define $A_{ba}$, the rates of $S_a$ and $S_b$ interconversion are first made equal and equation 2.6 becomes:

$$n_a/n_b = [B_{ba}I(v) + A_{ba}]/B_{ab}I(v) = 1 + A_{ba}/B_{ab}I(v)$$  \hspace{1cm} (2.7)

Equation 2.6 can then be set equal to equation 2.7 and evaluated in terms of $A_{ba}$, which requires defining the radiation density $I(v)$ in terms of a black body at temperature $T$:

$$\frac{n_a}{n_b} = [B_{ba}I(v) + A_{ba}]/B_{ab}I(v) = 1 + A_{ba}(\exp^{hv/kT} - 1)/8\pi hv^3c^{-3}B_{ab}$$  \hspace{1cm} (2.8)

where $h$ is Planck’s constant, $\nu$ is the frequency, $c$ is the speed of light, and $k$ is the Boltzmann constant. Setting equation 2.8 equal to $\exp^{hv/kT}$ from the equation 2.6 and solving for $A_{ba}$ gives:

$$A_{ba} = 8\pi hv^3c^{-3}B_{ab}$$  \hspace{1cm} (2.9)
The spontaneous emission depends on the cube of the frequency, which means as the wavelength becomes shorter spontaneous emission rate increases. The Einstein coefficient \( B_{ab} \) is equal to \((2\pi/3\hbar)D_{ab} \), where \( D_{ab} \) is the dipole strength, substituting into equation 3.9 gives:

\[
A_{ba} = \left(\frac{32\pi^3V^3}{3c^3\hbar}\right)D_{ab}
\]  

(2.10)

The dipole strength \( D_{ab} \) is calculated from the absorption spectrum, \( A_{ba} \), and can be determined without making an emission measurement. In the absence of non-radiative processes and stimulated emission, the rate of de-excitation from the excited state \( S_b \) is directly proportional to \( A_{ab} \):

\[
\frac{dn_b}{dt} = -A_{ba}n_b
\]  

(2.11)

Integrating equation 2.11 gives the solution, \( n_b(t) = n_b(0)e^{A_{ba}t} \) where \( n_b(0) \) is the number of molecules in excited state at time zero. The radiative lifetime of the state \( S_b \) is then given by:

\[
\tau_R = \frac{1}{A_{ba}} = \frac{1}{k_F}
\]  

(2.12)

where \( k_F \) is the fluorescence radiative rate, and \( \tau_R \) is the lifetime of the excited state. The radiative lifetime can be calculated from the absorption spectra, extinction coefficient, and emission spectra of the fluorophore using the following\(^{199,200} \):

\[
\tau_R = 2.88 \cdot 10^9 n^2 \frac{\int F(v)dv}{\int F(v)dv/V^2} \int \frac{\varepsilon(v)}{v}dv
\]  

(2.13)

where \( F(v) \) is the emission spectrum plotted in units of wavenumber (cm\(^{-1}\)), \( \varepsilon(v) \) is the absorption spectrum, and \( n \) is the refractive index of the medium. Based on equation 2.12 and 2.13, for an isolated molecule an increase in absorption rate means a shorter lifetime, which is only valid if the molecule emits radiation from the absorbing excited state. The
actual lifetime of an excited state is rarely as long as the calculated radiative lifetime based on equation 2.12. A molecule can possibly leave the excited state through several other processes besides spontaneous emission. The following sections outline different pathways facilitating the transition between the excited state and ground state. Understanding which pathways are available provides information related to the local environment, conformation, and dynamics of a fluorophore.

2.4 Factors governing fluorescence intensity

A molecule can exit the excited state through a number of nonradiative pathways. Each nonradiative pathway is associated with a rate term such as: internal conversion $k_{ic}$, collisional quenching $k_q$, intersystem crossing $k_{is}$, excited state reaction $k_{pc}$, and energy transfer $k_{et}$ (Figure 2-1). Environmental factors influence available pathways for depopulating the singlet excited state. The ratio of the radiative rate to the sum of nonradiative rates is known as the fluorescence quantum yield:

$$\Phi = \frac{k_F}{k_F + k_{ic} + k_q [Q] + k_{pc} + k_{is} + k_{et}}$$

(2.14)

The fluorescence quantum yield $\Phi$ is equal to the ratio of photons emitted to photons absorbed. Competing nonradiative pathways makes the observed lifetime shorter than $\tau_R$. Therefore, the time dependence of the concentration of excited state molecules $[S_b]$ is given by:

$$\frac{-d[S_b]}{dt} = (k_F + k_{ic} + k_q [Q] + k_{pc} + k_{is} + k_{et})[S_b]$$

(2.15)

When the above equation is integrated the solution is:

$$[S_b](t) = [S_{b0}]exp^{-t/\tau_F}$$

(2.16)
$[S_{00}]$ is the concentration of the excited molecule at time zero, and $\tau_F$ is the observed fluorescence lifetime defined as:

$$\tau_F = \left( k_F + k_{ic} + k_Q [Q] + k_{pc} + k_{is} + k_{et} \right)^{-1}$$

(2.17)

The fluorescent lifetime $\tau_F$ is inversely proportional to the sum of rates that depopulate the excited state, the radiative rate ($k_r$) and nonradiative rate ($k_{nr}$) which is the sum of all individual non-radiative decay rates. Combining the definitions of $\tau_F$, $\tau_R$, and $\Phi$ gives the following equation:

$$\Phi = \frac{\tau_F}{\tau_R}$$

(2.18)

The quantum yield is dependent on the rate of spontaneous emission and rate of depopulating the excited state nonradiatively. Proper characterization of fluorescence requires understanding the available nonradiative pathways. The following sections outline the competing nonradiative pathways for an excited state molecule.

### 2.4.1 Vibrational Relaxation

Immediately following excitation, the molecule has the same geometry and is in the same environment as the ground state. At this point, the molecule can do one of two things: emit a photon from the initial excited vibrational level or change conformation prior to emission. In solution a molecule is likely to decay rapidly to the lowest vibrational level of $S_1$ prior to emission. On the picosecond timescale solvent relaxation occurs, which refers to energy transfer from a vibrationally excited molecule to solvent. A fluorophore in the excited state has a large dipole moment prompting the solvent molecules to reorient or “relax” around
the excited molecule thereby lowering the energy of the system. Vibrational relaxation leads to a shift in the fluorescence spectrum to lower energy (longer wavelength) compared to the absorbance spectrum, known as a Stokes’ shift. In addition, the fluorescent spectrum is independent of excitation wavelength, since photon emission will always occur from the thermally equilibrated lowest vibrational levels of the $S_1$ excited state, known as Kasha’s Rule.

Vibrational relaxation also occurs following radiative decay. Similar to the excited state, the excess vibrational energy is lost through thermal motions. Assuming that the vibronic levels of $S_1$ and $S_0$ are similar, the same transitions that are favorable during excitation are favorable for emission, and the shape of the emission spectrum mirrors the absorption spectrum (Figure 2-2 For example, the transition in absorption $S_0(v=0) \rightarrow S_1(v=0)$ will occur at the same rate as $S_2(v=0) \rightarrow S_1(v=0)$ (Figure 2.2). Factors that influence the fluorescence spectrum include homogeneous and heterogeneous broadening. The lifetime of the excited state is determined by such processes as internal conversion, excited-state reactions, collisional quenching, intersystem crossing, and energy transfer. The following sections outline these alternative pathways leading to depopulation of the excited state.

2.4.2 Internal Conversion

Internal conversion occurs on the picosecond to nanosecond timescale when a molecule in the singlet excited state returns to the ground state without the emission of a photon.\textsuperscript{201,202} The excited state electronic energy is converted to vibrational energy in the ground state which is subsequently transferred to surrounding solvent molecules as heat. \textit{In vivo}, internal conversion is an important process preventing many photochemical reactions.
that are detrimental to biomolecules. For example, when DNA and RNA bases are excited by ultraviolet light, internal conversion is hypothesized as the main pathway for depopulation of the excited state.\textsuperscript{202,203} For highly fluorescent molecules, internal conversion accounts for a small fraction of transitions to the ground state because it is an inefficient process.

Internal conversion is inefficient because it involves the formation of conical intersection (CI) the crossing point between the excited and ground state. CIs are nonadiabatic events where the Born-Oppenheimer (BO) approximation breaks down, which means there is coupling between nuclear motions and electronic motions. Formation of a CI requires degeneracy between the excited state and ground state potential energy surfaces (PES).\textsuperscript{203} In terms of DNA and RNA, internal conversion is efficient because twisting double bonds of ground-state nucleobase enable a three-state CI pathway.\textsuperscript{204,205} Internal conversion is difficult to measure and not completely understood. Experimental data on CI has increased with the development of femtosecond time-resolved spectroscopy pump-probe techniques.\textsuperscript{206}

\subsection*{2.4.3 Intersystem crossing and phosphorescence}

Compared to internal conversion there is often a greater probability of populating the triplet state from the excited state by a process known as intersystem crossing, which occurs on the nanosecond timescale. Intersystem crossing requires a change of the spin state. The Intersystem crossing mechanism involves vibrational coupling between the excited singlet and triplet states. Subsequent to the initial intersystem crossing the triplet state can transition to the singlet ground state either by phosphorescence or by another intersystem crossing event.
Transition from the triplet excited state to the singlet state is spin forbidden and as a result the lifetime of triplet excited states are several orders of magnitude greater than those of fluorescence ($10^{-5}$ s). The long radiative lifetime means collision with quenchers and internal conversion can effectively compete with phosphorescence. Fluorescence quenching by halogens and heavy atoms is promoted by spin-orbit coupling which induces intersystem crossing to the triplet state.\textsuperscript{207-209} In this study intersystem crossing occurs via bi-molecular iodide ion quenching.

### 2.4.4 Bimolecular quenching of the excited state

Bimolecular quenching describes a decrease in fluorescence from interactions between fluorophore and quencher [Q]. Dynamic or collisional quenching refers to collisional encounters between excited state fluorophore and quencher, while static quenching represents formation of a ground-state complex between the quencher and fluorophore. During static quenching, transitions to the excited state are followed by immediate return to the ground state without the emission of a photon. The intrinsic molecular interaction between fluorophore and quencher is ideal for the study of biological systems. Bimolecular quenching can be used to examine accessibility of fluorophore to quencher or the diffusion coefficient of quencher within a biological system.\textsuperscript{1} In addition, bimolecular quenching of fluorophore usually occurs with no permanent change to the molecules.

Quenchers include a wide variety of molecules and atoms, and the mechanism of quenching depends on the specific quencher. For example quenching of anthracence by diethylaniline proceeds through an excited-state charge transfer complex, where the excited state accepts an electron from the amine.\textsuperscript{210} The proposed mechanism for paramagnetic
oxygen, a common quencher, is by facilitating intersystem crossing of the fluorophore to the triplet state.\textsuperscript{1} Similarly, intersystem crossing is the mechanism for quenching with large halogen ions such as bromide and iodide.\textsuperscript{207} In addition, pyrimidines and purines are known to quench fluorophores by a combination of dynamic and static quenching. The mechanism for nucleobase quenching of a fluorophore is thought to be photoinduced electron transfer; the nucleobase can act as a redox donor or acceptor with the fluorophore in the excited singlet state.\textsuperscript{42,44} In all these cases collisional quenching is considered a bimolecular process occurring with a rate $k_q[Q]$ and described by the following mechanism:

$$S_b + Q \rightarrow^k S_a + Q$$  \hspace{1cm} (2.19)

The collisional quenching of fluorescence is described by the Stern-Volmer equation. The Stern-Volmer equation is derived based on the fluorescence intensity in the presence and absence of quencher. The fluorescence intensity over time is defined by:

$$\frac{d[F^*]}{dt} = f(t) - \gamma[F^*]_0 = 0$$  \hspace{1cm} (2.20)

where $[F^*]$ is the concentration of fluorophore in the excited state, $f(t)$ is a function describing constant excitation, and $\gamma = \tau_0^{-1}$ is the fluorescence decay rate in the absence of quencher, which is defined as the sum of $k_R$ radiative rate and $k_{NR}$ nonradiative rates ($\gamma = k_R + k_{NR}$). In the presence of quencher an additional rate $k_q[Q]$ is required to give the Stern-Volmer equation:

$$\frac{F_0}{F} = \frac{\gamma k_q[Q]}{\gamma} = 1 + k_q \tau_0[Q]$$  \hspace{1cm} (2.21)

where $F_0$ and $F$ represent the fluorescence intensity in the absence and presence of quencher respectively, $k_q$ is the bi-molecular quenching constant, $\tau_0$ is the excited-state lifetime of the fluorophore in the absence of the quencher, and $[Q]$ is the concentration of the quencher.
Since dynamic quenching is a collisional rate process depopulating the excited state, the lifetime of the fluorophore in the absence \((\tau_0)\) and presence \((\tau)\) of quencher is defined by:

\[
\tau_0 = \gamma^{-1} \quad (2.22)
\]

\[
\tau = (\gamma + k_q [Q])^{-1} \quad (2.23)
\]

Substituting equations 2.22 and 2.23 into the Stern-Volmer equation 2.21 gives the excited state lifetime as a function of quencher concentration:

\[
\frac{F_0}{F} = \frac{\tau_0}{\tau} = 1 + \tau_0 k_q [Q] \quad (2.24)
\]

The Stern-Volmer quenching constant is equal to \(k_q \tau_0\) and if the quenching is known to be dynamic the constant is represented by \(K_d\). Static and dynamic quenching can be distinguished by their differing dependence on temperature, viscosity, and/or whether the excited state lifetime changes in the presence of quencher. The best method to determine if there is static quenching is to examine the excited state lifetime of a fluorophore in the absence and presence of quencher. Since collisional quenching is a rate process that depopulates the excited state, the lifetime in the presence of quencher will be shorter. In contrast, static quenchers do not emit a photon therefore the excited state lifetime will correspond to uncomplexed fluorophores. Fluorescent quenching is linearly dependent on quencher concentration. Deviations from linearity indicate the presence of multiple fluorescent populations with different accessibility to quencher. The Stern-Volmer equation can be modified to resolve the fraction of fluorophores accessible and inaccessible to quencher. The total fluorescence in the absence of quencher \((F_0)\) is given by the following equation:

\[
F_0 = F_{0a} + F_{0b} \quad (2.25)
\]
where the subscript \(a\) represents fluorescence from fluorophores accessible to quencher, the subscript \(b\) represents the fluorescence from fluorophores buried from quencher, and the \(0\) subscript refers to fluorescence in the absence of quencher. In the presence of quencher only the fluorescence intensity for the fraction of accessible fluorophores \(f_a\) will decrease. The observed fluorescent intensity is then given by:

\[
F = \frac{F_{0a}}{1 + K_a[Q]} + F_{0b}
\]

(2.26)

where \(K_a\) is the Stern-Volmer quenching constant for the fraction of fluorophores accessible to quencher and \([Q]\) represents the concentration of quencher. The fraction of fluorophore accessible to quencher can be used to infer exposure to solvent. In this thesis iodide quenching is used to examine how sequence context affects the accessibility of 6-MI to solvent. The fraction accessible is used to interpret whether 6-MI is intrahelical or extrahelical in certain sequences.

2.4.5 **Resonance energy transfer and excited state photochemical reactions**

The depopulation of the excited state by photochemical reaction occurs with a rate constant \(k_{pc}\). The term photochemical reaction refers to molecular processes that change the structure of the excited state fluorophore. Light provides the activation energy necessary for a chemical reaction to occur. The chemical reactions can be uni-molecular or bi-molecular. Biological examples of photochemical reactions include photosynthesis, the formation of vitamin D, isomerization of rhodopsin, and damage to DNA by ultraviolet irradiation.\(^1\) For molecular systems under consideration herein depopulation of a fluorescent molecule in the
excited state by the photochemical reaction pathway is insignificant and will not be discussed further.

A relevant process to the work in this thesis is Förster resonance energy transfer (FRET). FRET is a spectroscopic process by which energy is passed nonradiatively between molecules. The “donor” excited state fluorophore transfers energy nonradiatively to an “acceptor” molecule. During FRET there is no intermediate photon, rather the excited-state donor and ground-state acceptor molecule are coupled by dipole-dipole interaction. The extent of energy transfer is dependent on the donor quantum yield, the distance between donor and acceptor molecules, the relative orientation of the donor and acceptor transition dipoles, and the spectral overlap between donor emission and acceptor absorption. Typically FRET can occur over a donor-acceptor distance range of 10-100 Å, which has led to the development of FRET as a molecular ruler.

Providing the distance or orientation between donor and acceptor pairs does not change during the donor excited state lifetime, the distance between donor and acceptor can be determined from efficiency of energy transfer. The rate for energy transfer $k_{et}$ for the simplest case of a single donor-acceptor pair separated by a single distance $r$ is given by:

$$k_{et}(r) = \frac{Q_D k^2}{\tau_D \pi^6} \left( \frac{9000(ln10)}{128\pi^5 Nn^4} \right) \int_0^\infty F_D(\lambda)\epsilon_A(\lambda)\lambda^4 d\lambda$$

where $Q_D$ represents the donor quantum yield in the absence of acceptor, $\tau_D$ is the excited-state lifetime of the donor in absence of the acceptor, $N$ is Avogadro’s number, and $n$ is the refractive index of the medium normally assumed to be 1.4 for biological macromolecules in aqueous solution. For efficient transfer of energy, the donor and acceptor must be in resonance where
the fluorescent emission energy level of the donor must overlap with the acceptor absorption
energy levels as defined by the overlap integral $J(\lambda)$ in equation 2.37 where $F_D(\lambda)$ is the
fluorescence intensity of the donor in the wavelength range $\lambda$ to $\lambda + \Delta\lambda$ normalized to unity,
$\varepsilon_A(\lambda)$ is the extinction coefficient of the acceptor at $\lambda$ with units of M$^{-1}$ cm$^{-1}$, and $k^2$ is the
orientation factor of the relative transition dipoles of the donor and acceptor assumed to be
equal to 2/3. The $k^2$ value of 2/3 is only appropriate for a system where there is dynamic
random averaging of donor and acceptor conformation

In biological systems it is useful to simplify the equation for rate of energy transfer $k_{et}$
using the Föster distance $R_0$. The Föster distance corresponds to the distance at which $k_{et}$ is
equal to the donor excited state decay rate ($1/\tau_D$) in the absence of acceptor. When donor and
acceptor are separated by this distance one half of donor molecules decay by energy transfer,
while the other half decay by the pathways available in the absence of acceptor (radiative and
nonradiative). The rate of energy transfer can then is then defined as:

$$k_{et} = \frac{1}{\tau_D} \left( \frac{R_0}{\lambda} \right)^6 \quad (2.28)$$

The Föster distance $R_0$ can then be calculated by setting equation 2.28 equal to equation 2.27
and solving for $R_0$ to give the following:

$$R_0 = \left[ \frac{9000(\ln 10)Q_D k^2}{128\pi^5 N n^4} \right]^{1/6} \int_0^\infty F_D(\lambda) \varepsilon_A(\lambda) d\lambda \quad (2.29)$$

From the above expression it is clear the Föster distance is based on spectral properties and
the donor quantum yield. The above expression can be simplified by combing constants and
converting wavelength to units of centimeters ($J(\lambda)$ is then in units of cm$^3$ M$^{-1}$) and the $R_0$ in
Å is given by:

$$R_0 = 9.78 \times 10^3 \left[ k^2 n^{-4} Q_D J(\lambda) \right]^{1/6} \quad (2.30)$$
Importantly, efficient transfer will only occur if the rate of transfer is much faster than the decay rate (1/τ_D). If the transfer rate is slower than the decay rate, there will be little transfer during the excited state lifetime. The efficiency of energy transfer E is defined as the fraction of photons that are transferred to the acceptor relative to the total photons absorbed by the donor given by:

\[ E = \frac{k_{et}}{\tau_D^{-1} + k_{et}} \]  

Substituting equation 2.28 for the rate of energy transfer, the efficiency of energy can be defined in the terms of R_0 and r given as:

\[ E = \frac{R_0^6}{R_0^6 + r^6} \]  

Form equation 2.32 the efficiency is strongly dependent on the distance between the donor and acceptor molecules. When the Föster distance is known for a donor and acceptor pair a plot can be made for the efficiency of energy transfer versus distance using equation 2.32 (Figure 2-3). From the plot, transfer efficiency between 0.2 and 0.8 is linearly dependent on the distance (Figure 2-3). Föster’s theory of dipole-dipole energy transfer was validated using well-defined systems. One method was to separate donor and acceptor by a known distance with rigid system. A rigid molecular system was formed using two steroids separating dyes by 20 Å, 50% transfer efficiency was observed, consistent with the predicted value. The use of energy transfer as molecular ruler was further demonstrated by Stryer and Haugland using oligomers of poly-L-proline (n= 1-12) to separate donor naphthalene molecule from the dansyl acceptor molecule by fixed distances ranging from 12 to 46 Å. The authors observed transfer efficiency ranged from 100% to 16% dependent on inverse the distance between chromophores to the power of 5.9 ± 0.3, consistent with the r^6 dependence.
The plot displays the efficiency of energy transfer as a function of distance between donor and acceptor. The above plot highlights the strong dependence of RET transfer efficiency on distance between donor and acceptor. The transfer efficiency between 0.2 and 0.8 is linearly dependent on the distance between donor and acceptor.

**Figure 2-3 Transfer Efficiency**

The plot shows the efficiency of energy transfer as a function of distance between donor and acceptor. The efficiency decreases sharply from 1.0 to 0.2 as the distance increases from 0 to 60 Å, and then levels off at a lower efficiency as the distance further increases to 100 Å.
predicted by Foster’s theory. In addition, an $r^6$ dependence was observed using cyanine dyes separated by multilayers of fatty acids of known dimensions. There are several methods to measure the transfer efficiency. One of the most common methods is to measure the fluorescence intensity of the donor in the absence ($F_D$) or presence ($F_{DA}$) of the acceptor. The transfer efficiency can also be determined based on the excited state lifetime of the donor in the absence ($\tau_D$) and presence ($\tau_{DA}$) of the acceptor given by the following:

$$E = 1 - \frac{F_{DA}}{F_D} = 1 - \frac{\tau_{DA}}{\tau_D}$$ (2.33)

When using equation 2.33 several assumptions are necessary. Firstly, the donor and acceptor assumed to be separated by a fixed distance, which is correct for certain systems such as labeled proteins, but may not be the correct model for macromolecule-ligand interactions or a macromolecule that can adopt multiple conformations. In addition, the excited state decay is assumed to be single-exponential, which is rare for dyes attached to biomolecules. If the intensity decay is multi-exponential it is important to use the average decay times, given by the sum of the $\alpha_i \tau_i$ products, which are proportional to the steady state intensities. While a single distance is presumed, there are more complex models for donor-acceptor distance distribution or multiple distance distributions.

As stated above if the donor and acceptor molecules are able to freely rotate on a timescale shorter than the lifetime of the donor, the $k^2$ average value of two thirds is correct. The orientation factor $k^2$ represents the angular dependence between donor and acceptor molecules on efficiency of energy transfer and is given by:

$$k^2 = (\cos \theta_T - 3 \cos \theta_D \cos \theta_A)^2 = (\sin \theta_D \sin \theta_A \cos \phi - 2 \cos \theta_D \cos \theta_A)^2$$ (2.44)
Model defining the angles for orientation factor $k^2$. Vector D and A represent the dipole moment of excited state donor and ground-state acceptor. $\theta_T$ represents the angle between the emission transition dipole of the donor and the absorption transition dipole of the acceptor, $\theta_D$ and $\theta_A$ define the angles between the respective dipoles and the vector joining the center of mass for donor and acceptor, and finally $\phi$ is the angle between the planes defining the donor and acceptor.
Figure 2-4 illustrates the angles used to define the orientation factor where $\theta_T$ represents the angle between the emission transition dipole of the donor and the absorption transition dipole of the acceptor, $\theta_D$ and $\theta_A$ define the angles between the respective dipoles and the vector joining the center of mass for donor and acceptor, and finally $\phi$ is the angle between the planes defining the donor and acceptor. Iterating over all angular combinations reveals $k^2$ can range from 0 and 4, leading to a large uncertainty in certain cases. When the transition dipoles are collinear $k^2 = 4$, while for parallel dipoles $k^2 = 1$. Given that calculating the distance uses $k^2$ to the sixth power, variations in $k^2$ from 1 to 4 produces a 26% change in $r$, which means when using $k^2 = 2/3$ the maximum error in calculated distance is 35%. Importantly, the error in the calculated distance is excessively large if the dipoles are oriented perpendicular to one another ($k^2 = 0$).

A problem arises on how to treat $k^2$ when the rotation of donor and acceptor is restrained. Although, currently there is no experimental method to directly measure $k^2$, it is possible to estimate the error in $k^2$. The best option is to determine the distribution in $k^2$ establishing limits on the range of possible distances. One method to estimate the limits for $k^2$ is based on the anisotropies of the donor and acceptor, reflecting how well $k^2 = 2/3$ defines the conformational dynamics of the system.

### 2.5 Steady-state and time-resolved fluorescence

Fluorescence measurements can be grouped into two types called steady-state and time-resolved. The most common type is steady-state measurements, which use constant illumination of the sample. During steady-state experiments the sample is illuminated with a
continuous beam of light and the sample intensity, emission spectrum, or anisotropy is recorded. Fluorescence occurs on the nanosecond timescale, meaning when a sample is first exposed to the continuous wave of light steady-state emission is reached almost immediately.

To measure the excited state decay of a sample requires one of the following methods, time-domain or frequency-domain. The time domain method uses a pulse of light to excite the sample. The time width of the pulse is made as short as possible and at a minimum shorter than the decay time of the fluorophore. Following the excitation pulse the excited-state decay is recorded. The lifetime of a fluorophore is defined as the time necessary for the initial intensity at \( t=0 \) to decrease by \( e^{-1} \). The concentration of excited state molecules is given:

\[
\frac{d[S_b]}{dt} = -(k_r + k_{nr})[S_b] 
\]

where \([S_b]\) represents the number of molecules in the excited state at time \( t \) following excitation, \( k_r \) is the radiative rate, and \( k_{nr} \) is the nonradiative rate. The emission of a photon is a random event with each excited state molecule having equal probability of radiative decay which can be expressed as an exponential decay:

\[
[S_b](t) = [S_{b0}]e^{\frac{-t}{\tau}}
\]

The number of excited state molecules is not observed during experiments but rather the fluorescence intensity is recorded as a function of time \( I(t) \). The fluorescence intensity is directly proportional to the number of molecules in the excited state and defined as:

\[
I(t) = I_0 e^{\frac{-t}{\tau}}
\]
Time-resolved fluorescence measurements provide additional molecular information that is lost in the time-averaging process of steady-state fluorescence. The lost information is best understood by first considering how the steady-state intensity $I_{ss}$ for a homogenous sample is related to the excited state decay given by:

$$I_{ss} = \int_0^\infty I_0 \exp^{-t/\tau} dt = I_0 \tau$$  \hspace{1cm} (2.48)

where $I_0$ represents the concentration of excited state molecules and $\tau$ is the lifetime of the fluorophore. For a fluorophore in a homogenous environment the steady-state intensity is proportional to the lifetime (Figure 2-5). When a fluorescent probe is attached to biological macromolecule normally there exists more than one conformation. The steady-state fluorescence measurement will represent the ensemble of heterogeneous fluorophores as a single emission band. In contrast, the excited-state lifetime is a multiple exponential decay dependent on the fractional population of fluorophores in different environments (Figure 2-5).

Modeling of a fluorophores’ excited state decay is straightforward when it is single exponential. As stated above this is not the case for most fluorophores attached biological macromolecule. The fluorophore can exist in multiple conformations and a range of dynamics, which leads to a complex exponential decay. Modeling the complex lifetime requires a multi-exponential decay equation given by:

$$I(t) = \sum_i \alpha_i \exp(-t/\tau_i)$$  \hspace{1cm} (2.39)

where the sum of $\alpha_i$ is normalized to unity. Determining the precise values of $\alpha_i$ and $\tau_i$ for a multiple exponential decay is challenging because of correlation between the values. This requires a large data set $>10000$ counts, careful experimentation, and conservative interpretation of the data. Time-resolved fluorescence is measured using a technique known
The above plot displays the binned intensity counts as a function of time for fluorescent base analogue (FBA) 6-MI in three different environments: black = 6-MI free in solution, red = 6-MI incorporated into single-strand DNA, blue = 6-MI incorporated in duplex DNA, and gray is the instrument response function. Free in solution 6-MI has an excited state decay that is mono-exponential consistent with a homogenous environment. Incorporation of 6-MI into an oligomer results in a complex exponential decay reflecting the heterogeneity of its conformation. Formation of duplex DNA increases dynamic quenching with adjacent bases corresponding to a faster decay. Decay data adapted from Chapter 4
Time-correlated single-photon counting (TCSPC) set-up. Laser diode is used to excite sample with pulsed wavelength of light. Detected photon is used to trigger voltage ramp and delayed excitation pulse is used to trigger stop.

Figure 2-6 TCSPC schematic
as time correlated single photon counting (TCSPC). Examination of the instrument schematic is useful for understanding how TCSPC works (Figure 2-6). TCSPC is a digital method, binning photons based on time-correlation with respect to the excitation pulse. The time-to-amplitude converter (TAC) is the component that takes a voltage amplitude and converts it to time. In the forward configuration, the experiment starts with an excitation pulse, exciting the sample and starting the time measurement. In this thesis, data collection is done in the “reverse” configuration, where the emitted photon is used as the start and the excitation pulse is used as the stop.

The sample is repetitively excited using a pulsed laser LED. Each pulse from the laser LED is optically monitored by a high-speed photodiode. When the first photon from the sample is detected a voltage ramp in the TAC is triggered. Detection of the delayed excitation pulse stops the voltage ramp and the TAC sends an output pulse whose voltage is proportional to the time between the photon start and end pulse. An analog to digital converter (ADC) is used to convert the voltage to time and the time is binned using a multichannel analyzer (MCA). The summing of many pulses is used to build up a probability histogram of counts versus time channels. The experiment is continued until the counts in the peak channel reach a specified maximum. To eliminate pulse pileup and ensure detection of one photon at a time, the photon counting rate is limited to 1 photon per 100 laser pulses. Typically there are three curves associated with an intensity decay experiment: (i) the measured decay $N(t_k)$ (ii) the instrument response function $L(t_k)$, and (iii) the calculated decay curve $N_c(t_k)$. For all three curves $t_k$ refers to the discrete time channels with widths ($\Delta t$). The instrument response function (IRF) $L(t_k)$ represents the shortest time profile that can be
measured by the instrument, in this case ~200 ps. The model decay function $N_c(t_k)$ represents the convolution of the IRF with the intensity decay law. The decay time is the value of $\tau$ which is the best match between the measured data $N(t_k)$ and the calculated time-dependent intensities $N_c(t_k)$.

To understand how the measured intensity reflects a convolution of the actual sample intensity with the IRF, it is first necessary to define a function impulse function $I(t)$, which represents what would be experimentally observed if the IRF was a Dirac delta function ($\delta$-function). The excitation by the $\delta$-function is assumed to excite an impulse response at $t_k$ and is given by:

$$I_k(t) = L(t_k)I(t - t_k)\Delta t \quad (t > t_k)$$  \hspace{1cm} (2.40)

where $(t-t_k)$ is due to the impulse staring at $t = t_k$ with no emission occurring before excitation $(t < t_k)$. The measured decay $N(t_k)$ is then defined as the sum of the impulse response function that result from all the individual $\delta$-function excitation pulses occurring until $t_k$ and is given by:

$$N(t_k) = \sum_{t=0}^{t=t_k} L(t_k)I(t - t_k)\Delta t$$  \hspace{1cm} (2.41)

if the value of $\Delta t$ is very small, equation 2.41 can be expressed as an integral:

$$N(t) = \int_{0}^{t} L(t')I(t - t')dt'$$  \hspace{1cm} (2.42)

In equation 2.42 the experimentally measured intensity at a given time $t$ is defined as the sum of the intensities for all $\delta$-function excitation pulses occurring until time $t$, as long as the intensity for $L(t_k)$ is nonzero new intensity decays are created in the sample. The convolution integral is simplified by replace $t'$ with $\mu-t$ giving:

$$N(t) = \int_{0}^{t} L(t - \mu)I(\mu)d\mu$$  \hspace{1cm} (2.43)
The objective of convolution is to define an impulse response function that best represents the experimental data. The data that is collected is digital (counts per a channel) and for this reason equation 2.41 is the best form for modeling the data.

There are a number of methods that have been developed for analysis of TCSPC data. The methods include the maximum entropy method, the phase plane, Laplace transformation, nonlinear least squares analysis, and several others.\(^1\) One of the most established methods is the nonlinear least-squares analysis, which determines values with the highest probability for a specified model. The nonlinear least squares method of analysis requires satisfying a set of assumptions. The assumptions include: all uncertainty is in the dependent variable, the uncertainty is Gaussian distributed, there are no systematic errors, the system is correctly described by the mathematical fitting function, data points are independent observations, and there are sufficient data points so that the parameters are not overdetermined.

Nonlinear least squares analysis requires selecting a model thought to describe the data. As stated above, the goal is to obtain values for parameters in the model providing the best match between the data \(N(t_k)\) and the calculated decay \(N_c(t_k)\). To accomplish this goal the goodness-of-fit parameter \(\chi^2\) is minimized, the \(\chi^2\) is given by:

\[
\chi^2 = \sum_{k=1}^{n} \frac{1}{\sigma_k^2} \left[ N(t_k) - N_c(t_k) \right]^2 = \sum_{k=1}^{n} \frac{\left[ N(t_k) - N_c(t_k) \right]^2}{N(t_k)}
\]  

(2.44)

The sum extends over the number of data points \(n\) or the number of channels used during an experiment and \(\sigma_k\) is the standard deviation of each data point. The standard deviation is defined as the square root of the number of photon counts, \(\sigma_k = [N(t_k)]^{1/2}\) because TCSPC data follows a Poisson distribution, therefore the relative uncertainty decreases as the
number of points increases. The value of \( \chi^2 \) is proportional to the sum of the squared deviations between the measured values \( N(t_k) \) and the model values \( N_c(t_k) \), each divided by the squared deviation of the number of detected photons. To determine when \( \chi^2 \) is minimized the floating variables, such as \( \alpha_i \) and \( \tau_i \) are varied until \( N(t_k) \) and \( N_c(t_k) \) are closely matched. It is hard to interpret the values of \( \chi^2 \) because it depends on the number of data points. An easier value to understand is the reduced \( \chi^2_R \) given by:

\[
\chi^2_R = \frac{\chi^2}{n-p} = \frac{\chi^2}{\nu}
\]

where \( n \) represents the number of data points, \( p \) is the number of parameters that are floating, and \( \nu \) represents the number of degrees of freedom (calculated as which is \( n-p \)). If the errors contributing to \( \chi^2_R \) are random then the value will be close to unity. The \( \chi^2_R \) is close to one because the average \( \chi^2 \) per a data point is close to one, and normally the number of data points is much larger than the number of parameters. When the model used is not the best representation of the system the values of the \( \chi^2 \) and \( \chi^2_R \) are both larger than the values if the error were random.

### 2.5.1 Meaning of the multi-exponential decay

If the excited state decay requires a multi-exponential model, the intensity is thought to decay as the sum of individual single-exponential decays and is given by the expression:

\[
I(t) = \sum_{i=1}^{n} \alpha_i \exp\left(-\frac{t}{\tau_i}\right)
\]

where \( \tau_i \) is the decay time for the ith exponential, \( \alpha_i \) represents the amplitude of the ith component at \( t = 0 \), and \( n \) is the number of single-exponential decays. This expression can represent a number of systems. The multi-exponential can be applied to a mixture of fluorophores each with their own excited state lifetime \( \tau_i \). The multi-exponential is also used
to model another case where a single fluorophore displays a complex exponential. In this case the fluorophore is proposed to exist in multiple conformations each with their individual decay times. The multi-exponential decay is interpreted as each single-exponential representing a conformational distribution.

The meaning of the pre-exponential factor $\alpha_i$ depends on the system. One interpretation is that similar conformations of a fluorophore have the same radiative decay rate, meaning the pre-exponential factor $\alpha_i$ represents the fraction of molecules in a certain conformation at $t = 0$. The fractional distribution is thought to represent the ground state equilibrium. The pre-exponential factor $\alpha_i$ and excited state decay $\tau_i$ can then be used to determine the fractional contribution $f_i$ of each single-exponential decay to the time-averaged intensity given by:

$$f_i = \frac{\alpha_i \tau_i}{\sum_j \alpha_j \tau_j}$$ (2.47)

where $\alpha_i \tau_i$ corresponds to the area under the multi-exponential decay curve for a given single-exponential decay component. The time-averaged fluorescence measurement detects the emission of all photons regardless of the time the photon is emitted; therefore for shorter lifetimes the intensity is usually weaker.

2.6 Fluorescence anisotropy

When a fluorescent system is excited with plane polarized light, the emission will contain linearly polarized components. The polarized emission from a sample is related to the shape, size, and flexibility of the macromolecule. A typical experimental set-up is shown
in Figure 2-7. The excitation light propagates along the x-axis and the emission is detected along the y-axis. A polarizer is placed between the excitation source and the sample to orient.

Figure 2-7 Anisotropy schematic

Schematic for measurement of fluorescence anisotropy. Light polarized along the z-axis is used to excite the sample. A polarizer is placed before the detector and the emission polarized parallel or perpendicular to the z-axis is collected.
the electric vector $E$ of the excitation light parallel to the $z$-axis. The sample emission intensity is then measured through a polarizer. There are two components of the emitted light that are measured: $I_\parallel$ when the emission polarizer is parallel with the $z$-axis and $I_\perp$ when the emission polarizer is perpendicular with the $z$-axis.

One common process for depolarization of a fluorophore is rotational diffusion. The rate of rotational diffusion is dependent on the solvent viscosity, the size, and the shape of the fluorophore. Rotational diffusion is approximated by measuring a sample’s anisotropy, which describes the angular displacement of the fluorophore during the excited-state. The anisotropy $r$ is calculated using the following equation:

$$r = \frac{I_\parallel - I_\perp}{I_\parallel + 2I_\perp}$$  \hspace{1cm} (2.48)

where $I_\parallel$ is the emission through a polarizer oriented parallel with the electric vector and $I_\perp$ is the emission through a polarizer oriented perpendicular with the electric vector. Anisotropy is best understood using spherical coordinates with angle $\theta$ to the $z$-axis and $\phi$ to the $y$-axis to describe molecular orientation (Figure 2-8). If a sample contains rigid molecules in a random orientation with parallel absorption and emission transition, then the intensity from the sample is proportional to the square of its dipole moment vector projected onto the axis of observation $(E\mu)^2$ defined as:

$$I_\parallel(\theta,\phi) = \cos^2\theta$$ \hspace{1cm} (2.49)

$$I_\perp(\theta,\phi) = \sin^2\theta\sin^2\phi$$ \hspace{1cm} (2.50)

When exposed to polarized light, molecules with their transition moment aligned with the electric vector have the highest probability of absorption. This does not exclude all other molecules, and a distribution of molecules will absorb proportional to $\cos^2 \theta$, the angle
Figure 2-8 Fluorophore coordinate system

(A) Diagram for a fluorophore in terms of spherical coordinates. (B) Bottom diagram reveals how only molecules aligned with the polarized light will be excited (photoselection). Top represents the excited state distribution for a population of immobile fluorophores, the distribution of fluorophores produce an $r_0=0.4$
between absorption dipole moment and the electric vector. When direction of propagation for excitation is along the x-axis, molecules having an angle of $\phi$ with respect to the z-axis are excited with equal probability in a process known as photoselection (Figure 2-8). Since the majority of the photoexcited fluorophore’s transition moments are aligned with the z-axis, the dependence on $\phi$ can be eliminated replaced by the average value of $\sin^2 \phi$ given by:

$$\langle \sin^2 \phi \rangle = \frac{\int_0^{2\pi} \sin^2 \phi d\phi}{\int_0^{2\pi} d\phi} = \frac{1}{2}$$  \hspace{1cm} (2.51)

The solution is then substituted into equations 2.49 and 2.50 to give:

$$I_{\parallel}(\theta) = \cos^2 \theta$$  \hspace{1cm} (2.52)

$$I_{\perp}(\theta) = \frac{1}{2} \sin^2 \theta$$  \hspace{1cm} (2.53)

The probability of observing fluorophores oriented with respect to the z-axis is then be defined as $f(\theta)$. The intensity measured for each polarization is defined as follows:

$$I_{\parallel} = \int_0^{\pi/2} f(\theta) \cos^2 \theta d\theta = k \langle \cos^2 \theta \rangle$$  \hspace{1cm} (2.54)

$$I_{\perp} = \frac{1}{2} \int_0^{\pi/2} f(\theta) \sin^2 \theta d\theta = \frac{k}{2} \langle \sin^2 \theta \rangle$$  \hspace{1cm} (2.55)

where the function $f(\theta)$ describes the probability that a fluorophore is oriented between the angles $\theta$ and $\theta + d\theta$ and $k$ represents an instrumental constant. When the identity $\sin^2 \theta = 1 - \cos^2 \theta$ is substituted into equations 2.54-2.55 and combined with equation 2.48 the result for $r$ is:

$$r = \frac{3\langle \cos^2 \theta \rangle - 1}{2}$$  \hspace{1cm} (2.56)

In equation 2.56 the anisotropy is determined by the average value of $\cos^2 \theta$, the angle of the emission dipole relative to the z-axis. The observed parallel and perpendicular intensities are proportional to the square of the projection of the transition moments onto the
y and z-axis. When the fluorophore is aligned with the z-axis and the transition is collinear $\theta = 0$ and $r = 1.0$. A sample is never a homogenous population of perfectly aligned molecules, which means anisotropies are lower than 1.0. In contrast, the complete loss of anisotropy corresponds to an angle of 54.7°, the average value of $\cos^2 \theta = 1/3$. The above derivation is for collinear dipoles, so a more complex expression is needed for most fluorophores. In solution the molecules are in a random distribution. The number of molecules oriented with their transition dipole at an angle between $\theta$ and $d\theta$ is proportional to the area of a sphere described by $\sin \theta d\theta$ (Figure 2-8). The probability distribution of molecules excited by vertically polarized light is then given by:

$$f(\theta) d\theta = \cos^2 \theta \sin \theta d\theta$$  \hspace{1cm} (2.57)

when collinear absorption and emission occurs the maximum value of $<\cos^2 \theta>$ is given by:

$$<\cos^2 \theta> = \frac{\int_0^{\pi/2} \cos^2 \theta f(\theta) d\theta}{\int_0^{\pi/2} f(\theta) d\theta}$$  \hspace{1cm} (2.58)

substituting equation 2.57 into equation 2.58 gives $<\cos^2 \theta> = 3/5$ and the maximum anisotropy $r_0$ becomes 0.4, which requires collinear absorption and emission dipoles and no depolarization processes. When the excited-state population is preferentially aligned with the z-axis $I_\perp$ is 1/3 the value of $I_\parallel$. There are other processes that have higher anisotropies closer to one such as light scattering, and these processes interfere with fluorescence anisotropy measurements by skewing results to higher anisotropy.

In most cases the absorption and emission dipoles are not collinear and the value of $r_0$ is less than 0.4. When the transition moment is displaced by an angle $\beta$ between absorption and emission the $r_0$ is reduced (Figure 2-9). In addition, when the observed anisotropy for a
Figure 2-9 Depolarization

Diagram of the absorption and emission transition dipole moments. The angle $\theta$ is defined as the angle between the two dipole vectors. As the angle between becomes larger the limiting anisotropy $r_0$ decreases.
solution of fluorophores is decreased due to photoselection (decrease by 2/5) the fundamental anisotropy is then given by:

\[ r_0 = \frac{2}{5} \left( \frac{3 \cos^2 \beta - 1}{2} \right) \]  

(2.59)

The term \( r_0 \) is used to describe the anisotropy that is observed in the absence of other depolarizing processes such as rotational diffusion or energy transfer. When a fluorophore is attached to a macromolecule, there are several factors that determine rotational diffusion, such as size and shape of the macromolecule. In the simplest case for a spherical rotor the depolarization is described by the Perrin equation.

\[ \frac{r_0}{r} = 1 + \frac{\tau}{\theta} = 1 + 6D\tau \]  

(2.60)

where \( \tau \) represent the excited-state lifetime, \( \theta \) is the rotational correlation time, and \( D \) is the rotational diffusion coefficient (Figure 2.9). When the correlation time is much larger than the excited-state lifetime, the measured anisotropy \( r \) is equal to the fundamental anisotropy \( r_0 \). When the correlation time is much shorter than the lifetime the measured anisotropy approaches zero. It is easy to see the relationship between rotational diffusion and anisotropy by examining the model for the single-exponential time-resolved anisotropy decay for a spherical molecule:

\[ r(t) = r_0 e^{-t/\theta} = r_0 e^{-6Dt} \]  

(2.61)

in equation 2.61 the rotational correlation time for the fluorophore is given by:

\[ \theta = \frac{1}{6D} = \frac{\eta V}{RT} \]  

(2.62)

where \( \eta \) is the solvent viscosity, \( T \) is the temperature in kelvins, \( R \) is the gas constant, and \( V \) is the volume of the rotating unit. In equation 2.61 the macromolecular volume is an important factor, dictating the rotational correlation time of a molecule. Only spherical molecules
display a single-exponential anisotropy decay and more complex expression are necessary for nonsymmetrical molecules. The multi-exponential anisotropy decay for nonsymmetrical biomolecules reflects both rotation about the molecular axes and the segmental motions of a macromolecule.

Steady-state anisotropy measurements are an established method for characterizing macromolecules. Macromolecule binding is measured by an increase in anisotropy as a function of ligand concentration.\cite{214,215} Protein binding to DNA containing FBA 6-MI has been monitored by anisotropy.\cite{55} The unbound DNA is free to rotate in solution and the anisotropy is low; when protein is titrated into the solution binding will decrease the DNA’s rotational diffusion causing anisotropy to increase. The relationship between size and anisotropy is used to determine the stoichiometry of binding, if two or three ligands bind the relative anisotropy may double or triple respectively.\cite{214}

2.6.1 Time-Dependent Anisotropy Decay

The steady-state anisotropy is measured using continuous illumination, representing an average of the anisotropy decay over the excited state lifetime. Time-dependent anisotropy decay uses a pulsed excitation source to obtain additional molecular information. The shape and form of the anisotropy decay depends on the size, shape, and flexibility of the labeled molecule. The individual decay components represent the time-dependent emission for parallel (∥) and perpendicular (⊥) polarization given by the following:

\[
I_\parallel(t) = \frac{1}{3} I(t) [1 + 2r(t)] \tag{2.63}
\]

\[
I_\perp(t) = \frac{1}{3} I(t) [1 - r(t)] \tag{2.64}
\]
where \( r(t) \) represents the time-resolved anisotropy. When the time-resolved anisotropy decay is a multi-exponential decay it is described as follows:

\[
r(t) = r_0 \sum_j g_j \exp \left( \frac{-t}{\theta_j} \right) = \sum_j r_0 j \exp \left( \frac{-t}{\theta_j} \right)
\]

(2.65)

where \( r_0 \) is the limiting anisotropy in the absence of rotational diffusion such that \( r_0 = \sum j r_0 j \), \( \theta_j \) represent the individual rotational times, and \( g_j \) is the associated fractional amplitude for each rotational time (\( \sum g_j = 1.0 \)). The time-dependent decay components of the emission are each measured and used to calculate the time-dependent anisotropy:

\[
r(t) = \frac{I_\parallel(t) - I_\perp(t)}{I_\parallel(t) + 2I_\perp(t)}
\]

(2.66)

The time-dependent anisotropy is analyzed to determine which model best represents the system. The measured values are calculated from the polarized intensity decays given by:

\[
r_m(t_k) = \frac{N_\parallel(t_k) - G N_\perp(t_k)}{N_\parallel(t_k) + 2 G N_\perp(t_k)}
\]

(2.67)

where \( N_\parallel(t_k) \) and \( N_\perp(t_k) \) are the experimental measurements convolved with the IRF and G represent an instrument correction factor for bias in the monochromater for parallel vs. perpendicular polarized light. There are several ways to analyze the anisotropy time-decay; the method used in this thesis is to directly analyze the polarized intensity decays without calculating \( r_m(t_k) \). The instrument response function and polarized decays are used to calculate:

\[
N_\parallel^c(t_k) = \sum_{t=0}^{t=t_k} L_\parallel(t_k) I_\parallel(t - t_k) \Delta t
\]

(2.68)

\[
N_\perp^c(t_k) = \sum_{t=0}^{t=t_k} L_\perp(t_k) I_\perp(t - t_k) \Delta t
\]

(2.69)
where $L_d(t_k)$ and $L_{−d}(t_k)$ are the instrument response function. The calculated polarized intensities are used to minimize the $\chi^2_R$ using the parameter values in the intensity ($\alpha_i$ and $\tau_i$) and anisotropy decay ($r_{0i}$ and $\theta_i$) models given by the following equation:

$$
\chi^2_R = \frac{1}{v} \sum_{t=0}^{n} \frac{1}{\sigma^2_{ik}} [N^i_l(t_k) - N^i_l(t_k)]^2 + \frac{1}{v} \sum_{t=0}^{n} \frac{1}{\sigma^2_{ik}} [N^i_c(t_k) - N^i_c(t_k)]^2
$$

(2.70)

where $v$ is the degrees of freedom (number of points – number of floating parameters) and the weighting factors are directly proportional to the polarized components. Minimizing $\chi^2_R$ requires the correct model of the anisotropy decay law, which depends upon the size and shape of the fluorophore, its local environment, and attachment to a macromolecule. The simplest case, a spherical molecule, displays a single rotational correlation time. The complexity of the anisotropy decay increases for nonspherical fluorophores, if the fluorophore is located in an anisotropic environment, or with internal flexibility of a fluorophore within a larger macromolecule (Figure 2-10). In this thesis the fluorophore is a DNA fluorescent nucleobase analog, which is sensitive to local DNA flexibility and torsional motion of the oligomer.

Time-resolved anisotropy has been used to examine DNA dynamics. Some of the early work examining DNA dynamics used ethidium bromide intercalated into DNA to measure the rotational correlation time of DNA.\textsuperscript{217,218} Further work revealed a multi-exponential decay where the fast depolarization of the probe within the DNA resulted from DNA bending about the short axis and the longer depolarization resulted from torsional motion of the DNA about the long axis.\textsuperscript{219,220} The conformational dynamics of small oligomers has also been investigated. In addition, the use of fluorescent base analogs such as 2-aminopurine 2-AP, an adenine analog, has enabled the examination of both local and global
Figure 2-10 Time-dependent anisotropy

The time-dependent anisotropy decay is bi-exponential when the fluorophore is joined to a macromolecule. The fast DNA segmental motion causes a rapid decay in anisotropy. In contrast, the DNA torsional motion are normally slower than the lifetime of fluorophore preventing the anisotropy from decay to zero. Anisotropic decay for 6-MI in sequence CAFTC07ds adapted from Chapter 5
For example, when 2-AP is located as the central base in a 7-mer and opposite guanine, the anisotropy decay is composed of two components, a fast component of 87 ps, which accounts for 60% of the decay, and long component of 2 ns. The fast short component is attributed to 2-AP being extrahelical. The effects of ligands on DNA flexibility has been investigated with time-resolved anisotropy.

One focus in this thesis is to examine DNA conformational dynamics using the FBA 6-MI. There are two areas of interest: 1) the relationship between sequence and the local dynamics of 6-MI within the DNA and 2) the protein induced perturbation to DNA conformational dynamics.

### 2.7 Time resolved fluorescence resonance energy transfer (TR-FRET)

The above FRET section, describes using FRET as a “spectroscopic ruler” to measure the distance between donor and acceptor pairs. The efficiency of transfer for donor and acceptor pairs represents a fixed distance or the average distance for the solution ensemble. A method to evaluate the conformational distribution or distance distribution of the ensemble is time-resolved fluorescence resonance energy transfer (TR-FRET). Importantly, each system of FRET pairs is dependent on the geometry and dynamics, requiring a specific model to interpret TR-FRET data. The presence of energy transfer has a marked effect on the donor excited-state lifetime. Assuming the donor lifetime is single-exponential and a single distance between donor and acceptor (D-A) the lifetime will decrease while remaining single exponential. The distance between donor and acceptor is defined by a single rate $k_T$ for all donor molecules given by:
\[ I_{DA}(r, t) = I_{D0}^0 \exp\left[-t/\tau_D - k_T(r)t\right] \]  

(2.71)

where \( \tau_D \) is the lifetime of the donor molecule in the absence of acceptor, \( I_{D0}^0 \) is the intensity of donor molecules at \( t = 0 \), and \( k_T(r) \) is the D-A transfer rate given by equation 2.28. Since there is only one distance corresponding to one transfer rate \( k_T(r) = k_T \), the D-A intensity decay is given by:

\[ I_{DA}(r, t) = I_{DA}(t) = I_{D0}^0 \exp\left[-t/\tau_{DA}\right] \]  

(2.72)

The excited state lifetime of the donor in the presence of acceptor, \( \tau_{DA} \), is then equal to:

\[ \frac{1}{\tau_{DA}} = \frac{1}{\tau_D} + k_T(r) \]  

(2.73)

A single-exponential TR-FRET decay is rare for biomolecules, even with the use of external fluorescent probes. The complexity of the decay increases when the efficiency of transfer corresponds to distributed molecular conformations with a range of donor to acceptor distances or to multiple discrete molecular conformations with different donor-acceptor distances. For example, consider the case where excited state decay of a donor molecule is single-exponential. In the presence of acceptor some pairs are closely spaced, displaying a short decay time and others are separated by a greater distance corresponding to a longer decay time. The distribution in distances provides a range of decay times. To compare TR-FRET to steady-state FRET it is important to calculate the average decay rate of the ensemble, which is proportional to the steady-state intensity. The amplitude weighted average lifetime is proportional to the area under the decay and defined as the sum of \( \alpha_i \tau_i \) products. The following equation defines the average transfer efficiency \( \langle E \rangle \):

\[ \langle E \rangle = 1 - \frac{\int I_{DA}(t) dt}{\int I_D(t) dt} = 1 - \frac{\sum_i \alpha_{DAi} \tau_{DAi}}{\sum_i \alpha_{Di} \tau_{Di}} \]  

(2.74)
where \( I_{DA} \) and \( I_D \) are the donor excited state decay in the presence and absence of energy transfer respectively, and \( \langle E \rangle \) represents the ensemble time-averaged transfer efficiency.

The distance between donor and acceptors can be represented as a probability function \( P(r) \), which is distributed along the distance axis \( r \). For a single conformation the distribution is narrow and the time-averaged FRET distance is a good approximation for the system. The presence of a distribution of distances has a noticeable impact on the excited state decays of the donor molecules and the probability distribution function \( P(r) \). A distribution in donor acceptor distances requires a range in transfer rates \( k_T \) to be described in functional form. It is impractical to determine the donor-acceptor probability distribution for a random shape, and in general the most commonly used and appropriate distribution is a Gaussian:

\[
P(r) = \frac{1}{\sigma \sqrt{2\pi}} e^{\frac{-1}{2} \left( \frac{r - \bar{r}}{\sigma} \right)^2}
\]  

(2.75)

where \( \bar{r} \) is the mean of the Gaussian with a standard deviation of \( \sigma \). The distance distribution is usually described by the full width half maximum, which in nanometers is given by 2.354 \( \sigma \). The intensity decay for each donor-acceptor pair at distance \( r \) is given by equation 2.72. However, each of the unique donor-acceptor distance \( r \) cannot be observed, only the weighted average \( P(r) \) can be determined. The weighted average for the donor intensity decay in the presence of the acceptor is a summation of the intensity decays for all accessible distances.

The intensity decay distribution is defined as the integral equation:

\[
I_{DA}(t) = \int_{r=0}^{\infty} P(r) I_{DA}(r, t) dr = I_D^0 \int_{r=0}^{\infty} P(r) \exp \left[ -\frac{t}{\tau_D} - \frac{t}{\tau_D} \left( \frac{R_0}{r} \right)^6 \right] dr  
\]  

(2.76)

Equation 2.76 defines the excited state decay for an ensemble of donor-acceptor pairs as the weighted average of the decays for each unique donor-acceptor distance \( r \). The donor decay time \( \tau_D \) is known and can be fixed so the floating parameters are \( r \) and the FWHM.
excited-state decay of the donor is multi-exponential, the transfer rates are proportional to \( \tau_D^{-1} \) for each fluorophore. The transfer rate from each decay time component is then defined as:

\[
k_{Ti} = \left( \frac{1}{\tau_{Di}} \right) \left( \frac{R_0}{r} \right)^6
\]

(2.77)

and the distance dependent decay times for each component are defined as:

\[
\frac{1}{\tau_{DAi}} = \frac{1}{\tau_{Di}} + \left( \frac{1}{\tau_{Di}} \right) \left( \frac{R_0}{r} \right)^6
\]

(2.78)

If equation 2.72 correctly describes the rate of transfer for each component, the excited state decay of the donor-acceptor pairs spaced at a distance \( r \) is given by:

\[
I_{DA}(r,t) = \sum_i \alpha_{Di} \exp \left[ -\frac{t}{\tau_{Di}} - \frac{t}{\tau_{Di}} \left( \frac{R_0}{r} \right)^6 \right]
\]

(2.79)

The probability distribution and the intensity decay for a sample is then defined as:

\[
I_{DA}(t) = \int_0^\infty P(r) I_{DA}(r,t) \, dr
\]

(2.80)

Importantly, the multi-exponential decay for the donor does not introduce any additional parameters into the analysis. The excited state decay of the donor in the absence of acceptor can be measured in a separate experiment and the parameters for \( \alpha_i \) and \( \tau_i \) can be held constant. Similarly, in this thesis data analysis is performed by globally modeling the donor molecule in the presence and absence of the acceptor, linking the donor decay and amplitude terms between donor and D-A molecules.

TR-FRET has been applied to nucleic acids to characterize its dynamics alone and bound to proteins.\(^{17,224-227}\) For example, a 9 bp oligonucleotide was labeled on the 5’ end with FAM and the complementary strand 5’ terminus was labeled with TAMRA.\(^{228}\) The TR-FRET data was modeled as a distance distribution with a value of \( r \) near 37 Å and a width of
13 Å. The wide width of 13 Å was not expected and attributed to the flexible aminohexyl linker. Protein induced DNA bending has also been investigated using TR-FRET. The TR-FRET of peptides derived from the yeast basic leucine zipper (bZIP) transcription factor GCN4 bound to AP-1 sites in duplex DNA indicate protein induced bend angle of 21° with a distribution of only 2°. TR-FRET has also been used to examine the conformational distribution of a DNA four-way branch point called a Holliday Junction. The Holliday junction was labeled with two separate donor-acceptor pairs (I-II) and (I-IV) (Figure 2-11).
Figure 2-11 TR-FRET probability distribution

A) model for labeling the Holiday junction with fair pairs FAM and TAMRA B) TR-FRET indicates the two conformational distributions for FRET pairs F1R2 and F1R4.
The TR-FRET data indicated that sites I and II were further apart with a narrow distribution of $r$-values while sites I and IV were closer together with a broader distribution. This result was unexpected, since the structure of a Holliday junction suggests a similar distance distribution for both pairs. While HU is known to induce bending in DNA, the degree of bending and the distribution is unknown for many high affinity motifs. In addition, HU is known to bind to DNA constructs with a high degree of conformational flexibility. An interesting question is whether HU binding changes the conformational flexibility of DNA. TR-FRET is a method that can answer these questions.
3 Material and Methods:

3.1 HU Protein

The isolation of HU with high purity (>90%) was accomplished with the *E. coli* overexpressing strain RLM1078 (a gift from Dr. Roger McMacken, John Hopkins). Four 300 ml cell cultures containing 50 μg/ml ampicillin were grown to an optical density (OD) of 2 at 595 nm in Terrific Broth (1 Liter = 12 g Bacto-Tryptone, 24 g Bacto-yeast extract, 4 ml of glycerol in 0.17 M KH$_2$PO$_4$ and 0.72 M K$_2$PO$_4$) at 30 °C with constant shaking (190 rpm). To induce HU overexpression an additional 200 ml of Terrific Broth heated to 64° C was added to each culture flask yielding a final temperature of 42 °C to which ampicillin was added to maintain 50 μg/ml. Using warmed Terrific Broth for thermal induction ensures synchronous activation of all heat shock promoters controlling HU expression. Cell growth continued for an additional hour at 42 °C with vigorous aeration (250 rpm). The cell culture flask was then placed in an ice bath to arrest cell growth followed by centrifugation at 5000 X g for 15 min at 4° C to isolate cell pellets. The resulting cell pellets were rinsed with buffer A (25 mM Hepes pH 7.6, 1 mM DTT, 0.1 mM EDTA) and resuspended in 30 ml buffer A, and frozen stored at -80 °C.

The cells were thawed in a 10° C water bath and 4 M KCl, 25 mg/ml Lysozyme and 0.2 M PMSF were added to a final concentration of 1.0 M, 0.25 mg/ml, and 0.2 mM respectively. The freeze thaw cycle was repeated an additional 3 times. To ensure complete cell lysis, the cell suspension was sonicated on ice for a total of 15 minutes, cycle = 30 second sonication burst with 2 minutes rest intervals, power level = 3. The cell lysate was centrifuged at 100,000 X g for 1 hour at 4° C using a Beckman Ultracentrifuge (60Ti rotor, 8
cm radius 100,000 X g = 33,000 rpm). The supernatant was removed and the volume increased to 50 ml with the addition of Buffer B (25 mM Heps pH 7.6, 75 mM NaCl, 1 mM DTT, 0.1 mM EDTA, and 10% glycerol). The lysate was dialyzed against 3 X 4L of buffer B at 4° C.

The purification of HU from the cell lysate requires three column chromatography steps using an AKTA FPLC to control column loading, washing, and gradient elution; HU was followed by monitoring 230 nm and 254 nm absorbance while other cellular proteins were monitored by absorbance at 280 nm. The dialysate was loaded onto a 5 ml SP-Sepharose FF-HiTrap column (GE Healthcare Biosciences, Pittsburgh PA.) that was equilibrated with 5 column volumes of Buffer B; followed by washing away unbound protein with 10 column volumes of Buffer B. HU was eluted from the column using a linear gradient of 50 column volumes from 75 mM to 750 mM NaCl collecting 5 ml fractions of eluent; HU typically eluted at 0.4 M NaCl (Figure 3-1-A). The peak fractions were analyzed by gel electrophoresis. 20 μL of sample was mixed with 6 μL of reducing sample buffer, heated to 90 °C for 10 minutes, loaded onto a 16.5% SDS-PAGE gel, ran for ~1 hour at 35 mAmps, and visualized with Commassie staining; the HU monomer = 9 kD and fractions containing HU were pooled (Figure 3-1-A). The pooled fractions were then dialyzed against 3 X 4 L of Buffer C (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 75 mM NaCl, 1 mM DTT, and 10% glycerol). The dialyzed HU fractions were loaded onto a HiTrap Heparin HP column (GE Healthcare Biosciences, Pittsburgh PA.), which was equilibrated with 5 column volumes of Buffer C; the unbound protein was washed out with 10 column volumes of Buffer C. HU was eluted using a linear salt gradient of 50 column volumes from 75 mM NaCl to 1 M NaCl;
Figure 3-1 HU purification

(A) The chromatogram for HU dialysate loaded onto Hitrap SP-sepharose FF column indicates the protein elutes ~400 mM NaCl. (B) Combined fractions were further purified using Hitrap Heparin column, chromatogram reveals HU elutes at ~600 mM NaCl. (C) Nuclease was separated from HU using Mono-S HR column, chromatogram indicates HU elutes around ~400 mM NaCl. (D) The intensity of 6-MI as a function of time was used to examine fractions for nuclease activity, using the graph fractions from the second peak were determined to contain DNA nuclease. Protein activity was confirmed with GMSA. The GMSA used 4 nM 5’-overhang with increasing HU concentration, the apparent dissociation constant was determined to be 30 ± 12 nM.
5ml fractions of eluent were collected, HU typically eluted at ~400 mM NaCl (Figure 3-1-B).

The 230 nm peak fractions were analyzed by gel electrophoresis and fractions containing HU were pooled (Figure 3-1-B). The pooled fractions were dialyzed against 3 X 4 L of Buffer D (50 mM Hapes pH 7.6, 1 mM EDTA, 75 mM NaCl, 1 mM DTT, and 10% glycerol). The dialyzed HU fractions were loaded onto a Mono S HR 16/10 column (GE Healthcare Biosciences, Pittsburgh PA.), equilibrated with 5 column volumes of Buffer D and unbound protein was washed out with 10 column volumes of Buffer D. HU was eluted with a linear salt gradient of 50 column volumes from 75 mM NaCl to 1 M NaCl collecting 5 ml fractions of eluent, HU typically eluted at 400 mM NaCl (Figure 3-1-C). The peak fractions were analyzed by gel electrophoresis and fractions from the same peaks containing HU were pooled. The pooled fractions were then dialyzed against 3 X 4 L of Buffer E (50 mM Tris-HCl pH 7.6, 1 mM EDTA, 200 mM NaCl, 10% glycerol). For the nuclease assay, 500 μL aliquots from each fraction were then dialyzed against 3 X 4 L of 50 mM Tris-HCl pH 7.6 and 200 mM NaCl.

The samples were evaluated for nuclease-contamination by monitoring fluorescence intensity, which corresponds to digestion of DNA. Monitoring samples for nuclease activity required double-stranded DNA labeled with the nucleoside analog 6-MI.54 Approximately 1 μg HU (the concentration of HU was determined by Abs. 230 nm, \( \varepsilon_{230} = 37500 \text{ M}^{-1} \text{ cm}^{-1} \)) and 100 nM 6-MI labeled double-stranded DNA were mixed together in the presence of 10 mM Tris-HCl pH 7.6, 5 mM DTT ,and 10 mM MgCl2. Samples were incubated at 30 °C, the fluorescence was measured using a Horiba SPEX Fluoromax-4 spectrofluorometer (Edison, NJ), every 5 minutes for a total of 3 hours; exciting 6-MI at 340 nm and detecting resulting 440 nm emission, background fluorescence was monitored at 650 nm, slits = 4 and 8 nm,
integration time = 10 s, excitation polarizer set to 0 ° and emission polarizer set to 54.7 °. The background corrected fluorescence intensity vs. time was plotted (Figure 3-1-D). The slope of DNA + HU was compared to DNA slope, and fractions with a positive slope were marked as containing nuclease and removed. Based on these results, the second peak from the mono S chromatogram (Figure 3-1-C) was determined to contain nuclease and those samples were not used (Figure 3-1-D). The activity of the purified HU was confirmed by GMSA, in which HU binding affinity for 3’ overhang DNA was determined at a fixed concentration of 3 nM. Visualization on a native gel resolved the free DNA from the HU-DNA complex and an apparent dissociation constant ($K_{Dapp}$) of ~25 nM was calculated from the decrease in band density as a function of increasing HU concentration (Figure 3-1-D). The fractions without nuclease were then frozen in 20 mM Tris-HCl pH 7.6, 1.0 mM EDTA, 400 mM NaCl, and 40% glycerol and stored in -80 °C freezer. This general purification scheme was also applied to the HU mutants.

### 3.2 Oligonucleotides

DNA strands containing the nucleoside analogue 6MI were obtained from Fidelity Inc. (Gaithersburg MD.) in HPLC-purified form. Crude complementary strands were purchased from Integrated DNA Technologies (IDT, Coralville IA), and were purified using a denaturing 7 M urea 20% polyacrylamide gel. DNA bands were identified and isolated from the gel using UV shadowing. DNA was electroeluted from the gel (Schlechier and Schuell, Dassel, Germany) and dialyzed against $d_2$H$_2$O prior to lyophilization and storage. Samples were resuspended in $d_2$H$_2$O and sample concentrations were determined by UV absorption spectroscopy at 260 nm using molar extinction coefficients calculated using standard methods.
for modified or unmodified oligonucleotide. Duplexes were prepared by adding equal moles of the probe strand to a complementary strand in a 50 mM sodium phosphate buffer, pH 7.5. The samples were heated in a water bath at 90 °C for 5 min and then allowed to cool slowly to room temperature in the bath. Annealed samples were stored at -20 °C. The overhang, nicked, and cruciform DNA were prepared by adding equal moles of strands necessary for each construct in 300 mM NaCl, 10 mM Tris-HCl pH 8.0, and 1.0 mM EDTA. The flexible constructs were heated to 64 °C for 4 hours in a large water bath and then allowed to cool slowly to room temperature, approximately 15 hours. Annealed samples were stored at -20 °C.

3.2.1 Labeling of DNA

The oligonucleotides synthesized with an amino modifier C6 linker at the 5’ or 3’ end were obtained from Integrated DNA Technologies (IDT, Coralville IA), and purified using the method outlined above. The amine reactive fluorophores were obtained from Life Technologies (Carlsbad, California) including 5-Carboxyfluorescein succinimidyl ester (FAM) and 5-(and-6)-Carboxytetramethylrhodamine succinimidyl ester (TAMRA). The chemical structure of the two probes is shown in Figure 2.02. The labeled sequences are listed in Table 6-1.

The labeling procedure followed is described in the Life Technologies technical bulletin (MP 00143). Prior to labeling oligonucleotides were concentrated to ~25 μg / μL in ddH₂O using a 3000 MWCO vivaspin 500 centrifugal concentrator (Goettingen, Germany). Approximately a 15-fold mole excess of amine reactive fluorophore (resuspended in 28 μL
DMSO) was mixed with 200 μg of oligonucleotide in the presence of 0.075 M sodium tetraborate pH 8.5. The reaction was incubated for 12 to 16 hours at room temperature (RT) with constant shaking. The excess dye was removed from labeled sample by buffer exchange, the labeled DNA was concentrated using a 6000 MWCO vivaspin 500 centrifugal concentrator (Goettingen, Germany) followed by addition of ddH2O. The buffer exchange was repeated a minimum of 10 times or until no free dye was observed in the flow through, excess dye is small enough to pass through the filter with the flow through. Complete removal of excess dye was evaluated using two methods i. the absorbance of the flow through was measured for free dye ii. the degree of labeling (D.O.L.) was calculated. If either the flow through contained a peak at dye absorbance maximum or the D.O.L. was greater than 100%, the sample was deemed to contain free dye and buffer exchange was repeated. The D.O.L. was calculated as follows:

\[ D.O.L. = \frac{A_f(\lambda_f) \cdot \varepsilon_f}{(A_D(\lambda_D) - (A_f(\lambda_f) \cdot C.F.) \cdot \varepsilon_D)} \]  

(3.1)

where \( A_f(\lambda_f) \) is the absorbance at the fluorophore maximum wavelength(\( \lambda_{max} \)), \( \varepsilon_f \) is the fluorophore extinction coefficient at \( \lambda_{max} \), \( A_D(\lambda_D) \) is the DNA absorbance at \( \lambda_{max} = 260 \text{ nm} \), \( \varepsilon_D \) is the DNA extinction coefficient at DNA \( \lambda_{max} \), and C.F. is the correction factor for fluorophore absorbance at DNA \( \lambda_{max} \) defined as:

\[ C.F. = \frac{A_f(\lambda_D)}{A_f(\lambda_f)} \]  

(3.2)

where \( A_f(\lambda_D) \) is the absorbance of the unreacted dye molecule at 260 nm (DNA wavelength) and \( A_f(\lambda_f) \) is the absorbance of the unreacted dye molecule at its \( \lambda_{max} \). Following removal of free dye and to ensure correct calculation of fluorophore concentration, the sample was buffer exchanged into 20 mM Tris-HCl pH 8.0 and the D.O.L. was recalculated. The labeled oligonucleotide was stored in -20 °C freezer protected from light.
3.2.2 Labeling of protein:

The labeling of HU with cysteine reactive fluorophores required *E. coli* expressing HU containing a single point mutation to cysteine. A number of bacterial strains expressing HU with mutations were developed and obtained from Chaconas et al.\textsuperscript{212} The HU mutants were purified from *E. coli* as described above. The cysteine reactive fluorophore Alexa Fluor 488 maleimide was obtained from Life Technologies (Carlsbad, California). The labeling procedure was followed as described in the Life Technologies technical bulletin (MP 00003).\textsuperscript{231} To ensure reduction of the cysteine thiol 1mM TCEP was incubated with 100 μM mutant HU in reaction buffer (20 mM Tris-HCl pH 7.0, 1.0 mM EDTA, and 200 mM NaCl) for 15 minutes at 25 °C. The sample was then buffer exchanged with oxygen free reaction buffer (the reaction buffer was degassed for 5 minutes and then bubbled with nitrogen gas for 5 minutes, repeated three times) to remove oxygen. The HU mutant and 15-fold mole excess of Alexa Fluor 488 maleimide (resuspended in \textsubscript{dd}H\textsubscript{2}O) were mixed together and incubated under inert conditions at 4 °C for 8 to 12 hours. The excess dye was removed from labeled protein using Dye Removal Columns from Pierce Protein Biology Products (Rockford, IL.). The spin column was loaded with 250 μL of dye removal resin and centrifuged for 30 seconds, 1000 x g. The resin was then buffer exchanged 4 times with 20 mM Tris-HCl pH 7.0, 1.0 mM EDTA, and 150 mM NaCl, followed by centrifugation to remove buffer. The labeled protein was mixed with the resin in the column spin and allowed to incubate for 5 minutes on ice. The spin column was centrifuged 1000 x g for 4 minutes. The flow through was collected containing the labeled protein. The D.O.L. was then determined and if it was greater than 100% the excess free dye was removed from the labeled protein by repeating the above procedure.
3.3 Temperature Melts

The thermodynamic parameters for dsDNA containing 6-MI were determined by absorbance temperature melts. The absorbance at 260 nm as a function of temperature provides thermodynamic information about the global DNA structure while monitoring absorbance at 340 nm allows for examination of the local stability of 6MI. The temperature melts were carried out in 50mM NaPO₄ pH 7.5. For 260 nm melts a sample concentration of 1uM was used, while 340nm melts required a concentration of 15 uM DNA due to only one 6MI per duplex. Temperature melt experiments were carried out in a Beckman DU-650 UV/Vis spectrophotometer. Absorption at 260 nm and 340 nm was monitored from 11 to 95°C at a resolution of 1°C per point and a ramp rate of 1.0°C/min. The buffer absorbance was subtracted at each temperature point. Analysis of temperature melts was done in the program Meltwin (DNA Software Inc., Ann Arbor, MI). The hypochromicity at either 260 or 340 nm was calculated by taking the difference between high and low temperature absorbance divided by high temperature absorbance, which is represented by the following equation:

\[
\%h = 100 \cdot \frac{Abs_{SS\ DNA} - Abs_{DS\ DNA}}{Abs_{SS\ DNA}}
\] (3.3)

3.4 Steady State Fluorescence Measurements:

The steady state quantum yields were determined relative to 6-MI monomer for each single- and double-stranded oligomer containing 6-MI. The quantum yield was determined from an average of at least three experiments and was calculated using the equation given below:\(^1\)
\[
\frac{Q_x}{Q_r} = \frac{A_r(\lambda_r)I(\lambda_x)n_x^2D_x}{A_x(\lambda_x)I(\lambda_r)n_r^2D_r}
\] (3.4)

where \(A\) is the absorbance of the sample at the excitation wavelength, \(I\) is the intensity of the exciting light at the excitation wavelength, \(n\) is the index of refraction of the solution, and \(D\) is the integrated fluorescence emission. In this case, \(x\) refers to the single- and double-stranded oligomers containing 6MI, \(r\) is 6-MI monomer, which was used as a reference.

Absorbance measurements were performed with a Beckman Coulter DU-650 at a concentration of 15 \(\mu\)M 6-MI monomer in the strands to ensure an absorbance of at least 0.1 at 340 nm. For fluorescence measurements, oligomers were diluted either 1:25 or 1:50 and the 6-MI monomer was diluted 1:200. Fluorescent emission measurements were made from 390 nm to 550 nm at 1 nm/pt, with a 1 s integration time and with the excitation polarizer set to 0° and emission polarizer set to 54.7°, using a Horiba SPEX Fluoromax-4 spectrofluorometer (Edison, NJ). Samples were analyzed in 3 mm square quartz cuvettes, while being stirred continuously. Spectral analyses were performed with Grams AI ver. 8 (Thermo Electron Corp.).

**Fluorescence Quenching**

Collisional quenching measurements of fluorescence intensity were performed by titrating 0 to 200 mM KI into a 400 nM solution of DNA. The ionic strength of the solution was held constant at 200 mM with the balance made up with KCl. The concentration of 400 nM DNA was held constant during the titration. Single point emission was measured at 430 nm with an excitation of 340 nm using a Horiba Fluoromax-2 fluorimeter using slits of 2 and a 4 nm band-pass. The intensity was corrected for background and buffer contributions. For 6-MI monomer, quenching was measured by titrating potassium iodide from 0 to 200 mM
into a solution of 400 nM 6-MI monomer. The 6-MI monomer quenching data were evaluated using the Stern-Volmer equation:

\[
\frac{I_0}{I} = \frac{\tau_0}{\tau} = 1 + K_{SV}[Q]
\]

(3.5)

where \(I_0\) and \(I\) are the steady-state intensities, \(K_{SV}\) is the Stern-Volmer constant and \(\tau_0\) and \(\tau\) are the amplitude-weighted fluorescence lifetimes in the absence and presence of quencher \(Q\), respectively. In ss DNA and ds DNA, the dependence on quencher was not linear and the solvent accessible fraction was determined from a non-linear least squares fit of the data using the following equation:

\[
I = \frac{I_{fa}}{1 + K_{SV}[Q]} + I_0 - I_{fa}
\]

(3.6)

where \(I\) represents the intensity, \(I_0\) is the initial intensity, \(I_{fa}\) is the intensity from fluorophores that are accessible to quencher, and \(K_{SV}\) is the Stern-Volmer constant. The fraction accessible to quencher \((f_a)\) was determined from a minimum of at least three separate fluorescence-quenching experiments.

### 3.5 Analysis of Protein-DNA binding parameters

**Gel Mobility Shift Assay**

The gel mobility shift assay was used as an initial evaluation of purified HU for binding activity. HU protein binding to DNA constructs containing either a 3-prime or 5-prime 18 bp overhang were evaluated by GMSA. A 6.5% non-denaturing acrylamide gel was prepared in 22.5 mM Tris-Base, 22.5 mm Boric Acid and 0.625 mM EDTA (TBE). Samples were prepared in a 10 mM Tris-HCl pH 8.0 buffer containing 1.0 mM EDTA, 60 mM KCl, and 4% ficoll, incubating for 20 minutes at 25 °C. The gels were pre-run at a constant 10
V/cm for 1 hour or until stabilization of the current. The samples were loaded at a constant 5 V/cm and after loading all lanes the voltage was increased to 10 V/cm at 4 °C. To visualize the DNA bands, the gel was stained with fluorescent dye Syber Green (Life Technologies, Carlsbad, CA.) for 10 minutes. The fluorescent bands were detected using a Typhoon 9200 imager (GE Healthcare Biosciences, Pittsburgh PA.) and quantified using Image Quant TL software 7.0 (GE Healthcare Biosciences, Pittsburgh PA.). To determine the dissociation constant for HU, binding was quantified in terms of disappearance of free DNA band. Multiple bands arising from protein-DNA complexes, which migrated slower than the free DNA were observed, starting with 1:1 stoichiometry and continuing to the higher-order slowest migrating complexes at high protein concentration. The background in the gel was subtracted using the “rolling ball method”, applied in the Image Quant software. The apparent dissociation constant, $K_{D_{app}}$, was determined assuming one to one binding interactions using the following equations.

$$K_{D_{app}} = \frac{[HU][DNA]}{[DNA-HU_x]} \quad (3.7)$$

$$f = \frac{[DNA-HU_x]}{[DNA]_T} \quad (3.8)$$

Substitution of (3.8) into (3.7) and rearrangement yields the following binding equation:

$$f = \frac{([HU]_T+[DNA]_T+K_D) - \sqrt{([HU]_T+[DNA]_T+K_D)^2-4[DNA]_T[HU]_T}}{2[DNA]_T} \quad (3.9)$$

The $K_D$ refers to the apparent $K_D$ as described above. The optimization of the binding parameters was performed using the non-linear least squares fitting package from Origin 8.0 (Microcal).
**Solution DNA binding experiments**

The anisotropy binding experiments for Msh2-Msh6, IHF and HU were performed by titrating protein while maintaining a constant 1 nM concentration of DNA. The TBP binding experiment was performed by preparing individual samples containing 1nM DNA and increasing concentrations of TBP. A sample volume of 1 ml was used in quartz cuvettes measuring 4 mm x 10 mm. The Msh2-Msh6 binding buffer contained 5 mM MgCl₂, 5 mM Tris-HCl pH 8.0, and 50 mM NaCl. The IHF binding buffer contained 5 mM Tris-HCl pH 7.5, 0.1 mM EDTA, and 70 mM KCl. The HU specific binding (overhang, nicked, and cruciform DNA) buffer contained 40 mM KH₂PO₄-K₂HPO₄ pH 8.0, 1.0 mM EDTA, and 10% ethylene glycol, while non-specific binding (double-stranded and single-stranded DNA) buffer contained 10 mM Tris-HCl pH 8.0, 1.0 mM EDTA, and either 50 or 200 mM NaCl. TBP binding buffer contained 10% ethylene glycol, 5 mM Hepes, pH 8.0, 0.04% Triton X-100, 0.5 mM EDTA, and 60 mM KCl. Msh2-Msh6 was incubated for 3 minutes at 298K for each titration point; for IHF and HU each titration point was incubated for 10 minutes at 283K and for TBP the samples were incubated at 298K for 2 hours. Steady state fluorescence measurements were performed as described above. The excitation wavelength was 340 nm and emission was monitored at 430 nm, with slits of 5 and 15 nm bandpass, respectively. Samples were monitored for a total time of 30 s. All anisotropy measurements were corrected for background and buffer contributions and reported values result from at least 3 independent measurements. Anisotropy was calculated using the following equation:

\[ r = \frac{I_\parallel - G(I_\perp)}{I_\parallel + 2G(I_\perp)} \]  

(3.10)

where \(I_\parallel\) is vertically polarized light, \(I_\perp\) refers to horizontally polarized light and \(G\) refers to the G-factor determined for each experiment. The intensity binding experiments for Msh2-
Msh6, HU and IHF were performed under the same conditions used for anisotropy experiments, except that the polarizers were removed. Proteins were titrated into a 500 pM concentration of DNA (or down to 50 pM for IHF titrations and 200 pM for HU). Anisotropy and intensity binding curves were analyzed assuming a 1:1 binding interaction with the following equation:

\[
\begin{align*}
    r &= r_i + (r_f - r_i) \times \frac{([D]+K_D+[P]) - \sqrt{([D]+K_D+[P])^2 - 4[D][P]}}{2[D]} \\
    &= n_k\left[\frac{L}{1+k[L]}\right]
\end{align*}
\]

where \( r \) refers to either the anisotropy or the intensity at a certain protein concentration, \( r_i \) and \( r_f \) refer to the initial and final anisotropy or intensity values respectively, \([D]\) is the molar concentration of DNA, \([P]\) is the protein concentration; and \( K_d \) is the dissociation constant.

The HU binding to double-stranded and single-stranded DNA anisotropy curves were analyzed based on measured and literature reported stoichiometry of 3 HU’s to 1 DNA assuming identical noninteracting binding sites using the following equation:

\[
    r = r_0 + n_k\left[\frac{L}{1+k[L]}\right]
\]

where \( r \) represents the anisotropy of titrated sample, \( r_0 \) is the anisotropy of the free DNA, \( r_1 \) is the anisotropy with one HU bound, \( k \) represents the microscopic binding constant, \( n \) represents the number of HU bound, this equation assumes the fractional change in anisotropy is the same for each binding protein. The concentration of free HU is represented by \([L]\). The number of ligands was held constant \( n=3 \) based on stoichiometric data. The equation 3.12 was modified to take into account the change in 6-MI intensity with protein binding:

\[
    r = \frac{r_0(1+k[L])+n_k[L](r_1-r_0)}{1+k[L]+n_k[L](R-1)}
\]
where \( R = \Phi_1/\Phi_0 \), with \( \Phi_1 \) representing the fluorescence yield after binding the first ligand. The equilibrium binding constants were evaluated in Origin 8.0 (Microcal) using the Marquardt-Levenberg least-squares algorithm and the appropriate equations within the program. The goodness of fit was evaluated by visual inspection, \( \chi^2 \) values, and the correlation coefficients. The microscopic binding constant \( k \) is related to the stepwise binding constants such that \( K_1 = 3k \), \( K_2=k \), and \( K_3=k/3 \).

**Solution protein-DNA Stoichiometry**

The HU stoichiometric binding experiments were performed by titrating protein into a sample containing a fixed DNA concentration between 10 to 100 fold greater than HU dissociation constant. The DNA anisotropy was monitored as a function of HU concentration. A sample volume of 200 µL was used in quartz cuvettes measuring 3 mm x 3 mm. The binding buffer contained 5 mM MgCl\(_2\), 5 mM Tris-HCl pH 8.0, and 50 mM NaCl. The IHF binding buffer contained 5 mM Tris-HCl pH 7.5, 0.1 mM EDTA, and 70 mM KCl. The HU specific binding (overhang DNA) buffer contained 40 mM KH\(_2\)PO\(_4\)-K\(_2\)HPO\(_4\) pH 8.0, 1.0 mM EDTA, and 10% ethylene glycol, while non-specific binding (single-stranded DNA) buffer contained 10 mM Tris-HCl pH 8.0, 1.0 mM EDTA, and 50 mM NaCl. The HU sample at each titration point was incubated for 10 minutes at 283K. Steady state fluorescence measurements were performed as described above. The excitation wavelength was 340 nm and emission was monitored at 430 nm, with slits of 4 and 8 nm linear bandpass for 40 nM DNA and 2 and 3 nm bandpass for 2000 nM DNA. Samples were monitored for a total time of 30 s. All anisotropy measurements were corrected for background and buffer contributions and reported values result from at least 3 independent measurements. At the relatively high DNA concentration, 10-fold excess the \( K_D \), the DNA binds all of the added
HU leading to a linear increase in anisotropy as a function of molar ratio HU:DNA. The anisotropy as a function of the molar ratio HU:DNA was plotted and the stoichiometry was defined as the point where there is a plateau in anisotropy.

3.6 Fluorescent Lifetime

The Photon Technology International (PTI) TimeMaster instrument was used for time-correlated single photon counting (TCSPC). Samples containing 50 nM DNA were excited with a Becker & Hickl (BDL-375-SMC) 375nm pulsed picosecond- laser diode (rep rate=1MHz, <1mW average power). Emission was detected at 460 nm with emission slits of 15 nm using a 450 nm cut-off filter. Intensity decay data was collected to 65000 in the peak channel with the emission polarizer set to 54.7°. The vertical/vertical (V/V) polarization data used for anisotropy lifetime analyses were collected to 65000 in the peak channel and vertical/horizontal (V/H) polarization data were collected for the same amount of time as the V/V data. Stoichiometric amounts of DNA and protein (50 nM for IHF experiments, and 200 nM for Msh2-Msh6) were incubated for 10 min and data were collected using the same conditions as described above.

Time-resolved fluorescence decays were analyzed using the following equation:

\[ I(t) = \sum_{i=1}^{n} \alpha_i \exp \left(-\frac{t}{\tau_i}\right) \]  

(3.14)

where \( I \) is the intensity at time \( t \) and \( \alpha_i \) represents the fractional subpopulation with a lifetime of \( \tau_i \). The fluorescence decay curves were fit to a multi-exponential model using an iterative reconvolution method. All fitting and analysis were performed with Globals Unlimited. To reconvolve the instrument response function (IRF) with fluorescence lifetime data, the response function of the system was acquired by measuring scattered light from a colloidal
suspension of non-dairy creamer to give a full width at half-maximum of 14 channels corresponding to ~200 psec. The multi-exponential decay model \( I'(t) \) was used to describe the change in fluorescence intensity over time by fitting to the measured fluorescence decay data:

\[
I'(t) = \int L(t')I(t'-t)dt'
\]  

(3.15)

where \( L(t) \) is the IRF and \( I(t) \) is the intensity at time, \( t \). To determine the necessity for an additional decay term we evaluated the goodness of fit through examination of the residuals for systematic oscillations, reduction of the reduced chi-squared upon addition of the decay term and if the addition of a new decay term lead to a unique lifetime. The average lifetime \( \tau_{avg} \) is a useful value which allows for direct comparison of the average excited state lifetime between 6-MI in different constructs defined as follows:

\[
\tau_{avg} = \frac{\sum_\alpha \tau^2_i}{\sum_\alpha \tau_i}
\]  

(3.16)

Another useful value is the amplitude-weighted lifetime \( \tau_{(mol)} \), which is proportional to the integrated area of the excited state decay and is useful for calculating the excited state quantum yield and efficiency of energy transfer is given by:

\[
\tau_{(mol)} = \int_0^\infty I(t)dt = \sum_\alpha \alpha_i \tau_i
\]  

(3.17)

The experimentally measured anisotropy decay, \( r(t) \) is modeled as the sum of one or more exponential decay terms:

\[
r(t) = r_0 \sum_\beta \beta_i e^{-\frac{t}{\theta_i}}
\]  

(3.18)

where \( r_0 \) is the limiting anisotropy, \( \beta_i \) are the weighting factors for each contributing exponential and \( \theta_i \) is the rotational correlation time for component \( i \). The weighting factors \( \beta_i \) are normalized such that \( \sum_\beta \beta_i = 1 \). Two correlation times are used to define the experimentally observed time-dependent anisotropy decay. The molecular motions,
result in anisotropy decay, can be defined as fast local motion of the chromophore with a
rotational correlation time defined as $\theta_L$ and a slower motion due to overall rotation due to
DNA tumbling with a rotational correlation time defined as $\theta_R$. The time dependent
anisotropy decay is defined in terms of the molecular parameters as follows:

$$ r(t) = r_0 (\beta_1 e^{-\left(\frac{t}{\theta_L}\right)} + \beta_2 e^{-\left(\frac{t}{\theta_R}\right)}) $$  \hspace{1cm} (3.19)

This can be rearranged as:

$$ r(t) = r_0 (\beta_1 e^{-t\left(\frac{1}{\theta_L}\frac{1}{\theta_R}\right)} + \beta_2 e^{-\left(\frac{t}{\theta_R}\right)}) $$  \hspace{1cm} (3.20)

The relationship between the experimental expression and the molecular motion is defined as:

$$ \frac{1}{\theta_1} = \frac{1}{\theta_L} + \frac{1}{\theta_R} $$  \hspace{1cm} (3.21)

This expression for the anisotropy decay can then be rearranged recognizing that:

$$ \theta_R = \theta_2 $$  \hspace{1cm} (3.22)

so that the slow component of the observed decay can be equated directly to loss of
anisotropy from the overall tumbling motion. The fast component contains contribution from
both the local chromophore motion as well as the tumbling but the fast molecular motion, $\theta_L$,
can be extracted from the fit of the experimental data as follows:

$$ \theta_L = \frac{\theta_2 \theta_1}{\theta_2 - \theta_1} $$  \hspace{1cm} (3.23)

The experimental anisotropy decay curves, $r(t)$, were fit to a multi-exponential model using
an iterative reconvolution method as described above. The lifetime values determined from
fitting the excited state decay experimental data were fixed during the analysis of time
dependent anisotropy decay. The typical value for $\chi^2$ ranged from 1.05 to 1.32. The standard
deviations for lifetime and anisotropy decay values were determined from a minimum of
three separate experiments.
3.7 Fluorescence Resonance Energy Transfer

To correctly evaluate the inter dye distance from FRET experiments, the Förster distance ($R_0$) must be calculated for each donor and acceptor dye pairs under experimental conditions. The $R_0$ is the distance at which 50% of the donor molecules in the excited state decay through the nonradiative energy transfer pathway while the other 50% decay by radiative and nonradiative pathways that occur in the absence of acceptor. The $R_0$ is defined as follows:

$$R_0 = \frac{9000(\ln 10)}{128\pi^5 N} \left( k^2 \eta^{-4} \Phi_D(\lambda) \right)^{\frac{1}{6}} \text{ (in Å)}$$

(3.24)

The constants can be combined and are equal to 0.211 when defining $R_0$ in Å and $J(\lambda)$ units are M$^{-1}$cm$^{-1}$(nm)$^4$. $k^2$ is the factor indicating the relative orientation in space of the transition dipole for the donor and acceptor which is 2/3, if a dynamic random averaging system of the donor and acceptor is assumed, $\eta$ is the refractive index of the medium assumed to be 1.4 for biomolecules. $\Phi_D$ is the quantum yield of the donor in the absence of acceptor, and $J(\lambda)$ is the spectral overlap integral between emission of the donor fluorophore and absorption of the acceptor fluorophore expressed with units M$^{-1}$cm$^{-1}$(nm)$^4$. The spectral overlap between donor emission and acceptor absorption is defined as:

$$J(\lambda) = \int_{0}^{\infty} F_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda = \int_{0}^{\infty} \frac{F_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda}{\int_{0}^{\infty} F_D(\lambda) d\lambda}$$

(3.25)

$F_D(\lambda)$ is the normalized intensity of donor fluorescence occurring at each wavelength with units of nanometers, from $\lambda$ to $(\lambda + \Delta \lambda)$ and $\epsilon_A$ is the acceptor only extinction coefficient spectrum (M$^{-1}$cm$^{-1}$).

The uncertainty in FRET measurements can be large because the relative orientation of the donor and acceptor fluorophores in solution is unknown. The error is especially high if
the dyes are immobile on the nanosecond timescale. It is possible to define the range in uncertainty for $k^2\text{min}$, the relative orientation can vary from $k^2\text{min}$ to $k^2\text{max}$. The values of $k^2\text{min}$ and $k^2\text{max}$ are directly related to the depolarization of donor and acceptor, and can be calculated based on steady state anisotropy. The $k^2\text{min}$ and $k^2\text{max}$ values are calculated using the following equations:

$$k^2\text{min} = \frac{2}{3} \left( 1 - \frac{\sqrt{\frac{5}{2}r_d + \sqrt{\frac{5}{2}r_a}}}{2} \right)$$  \hspace{1cm} (3.26)

$$k^2\text{max} = \frac{2}{3} \left( 1 + \sqrt{\frac{5}{2}r_d + \sqrt{\frac{5}{2}r_a} + 3 \sqrt{\frac{5}{2}r_d \sqrt{\frac{5}{2}r_a}} \right)$$  \hspace{1cm} (3.27)

where $r_d$ and $r_a$ are the steady state anisotropy values measured for donor only and acceptor only labeled samples. To determine the uncertainty in the FRET determined distance the following equation is used:

$$R_{\text{min}} = \left( \frac{6 \sqrt{k^2\text{min}}}{2/3} \right) \cdot R_{\text{app}}$$  \hspace{1cm} (3.28)

$$R_{\text{max}} = \left( \frac{6 \sqrt{k^2\text{max}}}{2/3} \right) \cdot R_{\text{app}}$$  \hspace{1cm} (2.29)

where $R_{\text{app}}$ is the apparent distance calculated when $k^2=2/3$. The errors in the orientation factors are considered with experimental errors when reporting the uncertainty in the FRET distance measurements.

*End to end distance dependent efficiency of transfer (E)*

134
Once the Förster distance $R_0$ is known the rate of energy transfer, $k_T \text{(s}^{-1}\text{)}$ can be calculated. The rate of energy transfer, from the singlet excited state ($S_1$) of a donor molecule D to the ground state ($S_0$) of an acceptor molecule A is defined as follows:

\[
k_T = \left(\frac{R_0}{R}\right)^6 / \tau_D \quad (3.30)
\]

The transfer rate must be faster than the excited state decay rate for transfer to occur. The efficiency of transfer ($E$) is defined as the fraction of photons absorbed by a donor molecule that are transferred to the acceptor. The fraction is defined as follows:

\[
E = \frac{k_T(r)}{\tau_D^{-1} + k_T(r)} \quad (3.31)
\]

where $k_T$ is divided by the total decay rate of the donor in the presence of the acceptor.

Substituting equation 3.30 into equation 3.31 the energy transfer can be defined in terms of the strong dependence on distance

\[
E = \frac{R_0^6}{R_0^6 + r^6} \quad (3.32)
\]

The above equation indicates energy transfer is dependent on distance to the sixth because of this dependence measurements of the distance ($r$) are only reliable when $r$ is within a factor of 2 of $R_0$. The measurable distances range from $r$ twice the Förster distance ($r = 2R_0$) and energy transfer is 1.54% to $r$ is half the Förster distance ($0.5R_0$) and energy transfer is 98.5%. Distances outside of this range ($r = 0.5R_0$ to $r = 2R_0$) are not reliably measured.

There are several ways to measure transfer efficiency. One method is to measure the relative fluorescence intensity of the donor in the absence ($F_D$) and presence ($F_{DA}$) of acceptor:

\[
E = 1 - \frac{F_{DA}}{F_D} \quad (3.33)
\]
The transfer efficiency can also be calculated based on the excited state of the donor in the absence \( (\tau_D) \) and the presence \( (\tau_{DA}) \) of acceptor

\[
E = 1 - \frac{\tau_{DA}}{\tau_D} = 1 - \frac{\int t_D(t)dt}{\int t_D(t)dt} = 1 - \frac{\sum_i \sigma_{DAi} \tau_{DAi}}{\sum_i \sigma_{DI} \tau_{DI}} \tag{3.34}
\]

Fluorescence energy transfer can also be calculated based on the ratio of donor excited acceptor fluorescence to directly excited acceptor fluorescence, which is defined as \((\text{Ratio})_A\). The calculation of \((\text{Ratio})_A\) uses the following equation:

\[
(\text{Ratio})_A = \frac{F_{DA}(v_1,v')}{F_A(v_2,v'')} \tag{3.35}
\]

\(F_{DA}(v_1,v')\) represents the extracted donor excited acceptor emission spectra, where the excitation wavelength is at donor \(\lambda_{\text{max}}(v')\) and \(F_A(v_2,v'')\) represents the emission of acceptor excited at acceptor \(\lambda_{\text{max}}(v'')\). To generate the extracted donor excited acceptor spectra, the donor emission must be subtracted. This is achieved by spectral subtraction of \(F_D(v_1,v')\), emission of donor in absence of acceptor excited at donor \(\lambda_{\text{max}}\) wavelength, from \(F_{DA}(v_1,v')\). Background contributions were removed by buffer spectral subtraction. The integrated area was then calculated for each spectra and used to determine \((\text{Ratio})_A\). The efficiency of energy transfer \(E\) is related to \((\text{Ratio})_A\) by the following equation:

\[
(\text{Ratio})_A = \left\{ E \cdot d^+ \left[ \frac{\varepsilon^D(v')}{\varepsilon^D(v'')} + \frac{\varepsilon^A(v')}{\varepsilon^A(v'')} \right] \frac{\Phi^A(v_1)}{\Phi^A(v_2)} \right\} \tag{3.36}
\]

where \(\varepsilon^D\) and \(\varepsilon^A\) are the extinction coefficients of the donor and acceptor at their respective excitation wavelengths, \(d^+\) is the fraction of molecules labeled with the donor fluorophore, and \(\Phi^A\) represent the quantum yield of acceptor molecule. The calculated \(E\) is then directly related to the distance between donor and acceptor molecules by the following equation:

\[
E = \frac{1}{1 + \left( \frac{R_0}{R} \right)^6} \tag{3.37}
\]
During steady-state FRET experiments HU was titrated into a fixed 40 nM of labeled DNA (donor only, acceptor only, or doubly labeled DNA) in the presence of 5 mM Tris-HCl pH 8.0, 70 mM KCl, 1.0 mM EDTA, and 0.05% Tergitol NP40 for a total volume of 200 μL. The fluorescence emission measurements were made by exciting the samples at the donor excitation (FAM = 494 nm), or acceptor direct excitation (TAMRA = 555 nm), and scanning donor emission (FAM = 510 to 700 nm) or acceptor emission (TAMRA = 565 to 700 nm) at 1 nm/pt, with a 1 s integration time, with the excitation polarizer set to 0° and emission polarizer set to 54.7°, and with slits set to 4 and 8 nm using a Horiba SPEX Fluoromax-4 spectrofluorometer (Edison, NJ). Samples were analyzed in 3 mm square siliconized glass cuvettes, while being stirred continuously. Spectral analyses were performed with Grams AI ver. 8 (Thermo Electron Corp.).

The above equations describe the donor and acceptor pair separated by a fixed distance. In solution, the reported distance represents an average distance of the ensemble of donor and acceptor molecules, with no information about the distribution. Time-resolved fluorescence resonance energy transfer (TR-FRET) is used to evaluate the conformational distribution. In addition, TR-FRET energy transfer is directly related to steady-state FRET energy transfer by integration of the decay curves for donor in the presence and absence of acceptor. The amplitude weighted average decay is proportional to the area under the decay and defined as the sum of \( \alpha_i \tau_i \) products. The following equation defines the average transfer efficiency (\( \langle E \rangle \)):

\[
\langle E \rangle = 1 - \frac{\int I_{DA}(t)dt}{\int I_D(t)dt} = 1 - \frac{\sum_i \alpha_i \sigma_{DAi} \tau_{DAi}}{\sum_i \alpha_i \sigma_{Di} \tau_{Di}}
\]  

(3.38)
where \( I_{DA} \) and \( I_D \) are the donor excited state decay in the presence and absence of energy transfer respectively and \(<E>\) corresponds to an average of the molecular conformations. TR-FRET is able to recover the D-A probability distribution from the non-exponential decays of the donor, which are not available from steady-state fluorescence measurements. It is impractical to determine the D-A probability distribution for a random shape and in general the most commonly used and appropriate distribution is a Gaussian:

\[
P(R) = \frac{1}{\sigma \sqrt{2\pi}} e^{\left[-\frac{1}{2} \left(\frac{R - \bar{R}}{\sigma}\right)^2\right]} \tag{3.39}
\]

where \( R \) with macron is the mean of the Gaussian with a standard deviation of \( \sigma \). The weighted average for the donor intensity decay in the presence of the acceptor is a summation of the intensity decays for all accessible distances and is given by

\[
I_{DA}(t) = \int_{R=0}^{\infty} P(R) I_{DA}(R,t) dR = I_D^0 \int_{R=0}^{\infty} P(R) \sum_{i=1}^{n} \alpha_i \exp\left[-\frac{(\bar{R} - R_i)^2}{4\sigma_i^2}\right] \tag{3.40}
\]

The expression represents the excited state decay for an ensemble of D-A pairs given by the weighted average of the decays for each D-A distance. Data analysis was performed by globally modeling the donor molecule in the presence and absence of the acceptor, linking the decay and amplitude terms between donor and D-A molecules. TR-FRET can also be modeled as multiple populations with different distance distributions given by the following equation:

\[
P(R) = \sum_{i=1}^{n} X_i \sigma_i^{-1} (2\pi)^{-1/2} \exp\left[-\left(\frac{R - R_i}{2\sigma_i}\right)^2\right] \tag{3.41}
\]

\[
\sigma_j = \frac{FWHM_j}{[2 \cdot (2 \ln 2)^{1/2}]} \tag{3.42}
\]
where $X_i$ is the mole fraction and $\sigma_i$ the standard deviation for each population. Several functional forms of the probability distribution $P(R)$ were evaluated: single discrete distance ($1r$), two discrete distances ($2r$), single Gaussian ($1G$), and two Gaussian components ($2G$). The best fits were evaluated based on residual plots and $\chi^2$, in most cases the best fit was obtained using the $1G$ model.

The PTI TimeMaster instrument was used to acquire TR-FRET data. Samples containing 200 nM DNA, donor only or donor + acceptor in the presence of 5 mM Tris-HCl pH 8.0, 70 mM KCl, 1.0 mM EDTA, and 0.05% NP40 for HU + overhang complex, 5 mM Tris-HCl pH 8.0, 50 mM NaCl, 1.0 mM EDTA, 0.05% NP40 for HU- double-stranded DNA complex, or 5 mM Tris-HCl pH 8.0 70 mM KCl, and 1.0 mM EDTA for IHF, were excited with a Becker & Hickl (BDL-375-SMC) 375nm pulsed picosecond- laser diode (rep rate=1MHz, <1mW average power). Emission was detected at 460 nm for 6-MI and 520 nm for FAM, with emission slits of 15 nm using a 450 nm cut-off filter. Intensity decay data was collected to 20000 in the peak channel with the emission polarizer set to 54.7°.

Stoichiometric amounts of DNA and protein were incubated for 10 min and data were collected using the same conditions as described above. The data was initially modeled using Globals unlimited program. Modeling FRET data for population distribution and the number of populations was achieved using the program FARGOFIT. To correctly model FRET distributions, the donor only and donor + acceptor decays were globally fit linking decay terms and amplitudes between models. The possibility of two FRET populations was evaluated by constructing a model with multiple populations. To evaluate necessity of an additional population and the goodness of a model, the $\chi^2$ was examined for improvement and if the model produced physiologically relevant distance values.
**HU-DNA model**

To generate the computational molecular models of the HU-DNA complex several restraints were used based on experimental data. First, based on end labeled DNA FRET data the HU induced distortion was modeled as a symmetric bend with each HU monomer equally distorting the DNA by $60^\circ$. In addition, using the HU/IHF crystal structures and experimental data, the distance between intercalating prolines was fixed to 9 basepairs. Finally, DNA footprinting indicates one of the prolines intercalates at the single-double strand junction. The above experimental data was used to reduce the number of models to four: i) $\alpha$-HU P64 intercalates at junction and $\beta$-HU P64 intercalates 9bp away in double-strand ii) $\beta$-HU P64 intercalates at junction and $\alpha$-HU P64 intercalates 9bp away in double-strand iii) $\alpha$-HU P64 intercalates at junction and $\beta$-HU P64 intercalates 9bp away in single-strand iiii) $\beta$-HU P64 intercalates at junction and $\alpha$-HU P64 intercalates 9bp away in single-strand. The four models were then created for each FRET DNA construct.

The nature of the pseudo-self-complementary DNA construct in the *Anabaena* HU crystal structure makes it hard to align with models of the overhang constructs. In contrast, IHF was crystalized with a DNA structure containing a single nick to the backbone PDB ID = 1IHF. To produce the models the IHF-DNA crystal structure (PDB ID= 1IHF) was used as a template to dock HU onto DNA constructs. The models required a protein structure-homology model for the Hu$\alpha$$\beta$ heterodimer based on the Hu$\alpha$a crystal structure (PDB ID = 1P51). The protein structure homology model for *E. coli* hup$A$ and hup$B$ genes were created using Swiss-model workspace. The *E. coli* Hu$\alpha$$\beta$ heterodimer was then created by aligning each $\alpha$ and $\beta$ monomer with one respective HU-1P51 crystal-structure $\alpha$-monomer in Pymol.
The Huαβ heterodimer was then aligned with IHF in the IHF-DNA crystal structure 1IHF. The bent DNA constructs were formed using the online webserver 3D DART. To align the FRET DNA structure with the IHF crystal structure DNA required making a “DNA-homology model”, the 9bp sequence between bends in the DNA was mutated to the IHF sequence. This allowed the DNA to be aligned with the IHF sequence. A second bent construct was then generated replacing the 9bp IHF with the sequence from FRET construct. This construct is then aligned with the “homology-model” construct using the bases outside the 9 bp region for alignment.

The distance between the amino acid Ca representing the position of cysteine mutant and guanine N9 representing 6-MI’s position in the construct were measured in Pymol. The theoretical distance between 6-MI and the five mutants (αHU-Q43C, αHU-A78C, αHU-K90C, βHU-Q64C, and βHU-T70C) was measured for each construct with all four models. The quality of each model was evaluated by calculating the RMSD between the FRET and model distance for a particular construct.

3.8 Molecular dynamics simulation

Technological advancements in biological spectroscopy are increasing the available information on macromolecular structure and dynamics. Understanding this information requires rigorous theories for modeling the electronic structure and dynamics of fluorescent probes in large biomolecules. The competing excited state relaxation pathways make modeling of fluorescence spectroscopy particularly difficult. In addition, at this time an \textit{ab initio} or semi-empirical quantum chemistry model is prohibitively expensive. There is
continual improvement in computational hardware and software which is facilitating progress toward accurate *ab initio* quantum mechanical (QM) calculations of electronic excited states and when coupled with classical mechanical molecular dynamics (MD) simulations of biological macromolecules accurate models will be achieved.\textsuperscript{24,242} An important distinction is that QM deals with the electrons in a system and is limited to hundreds of atoms; while molecular mechanics ignores the electronic motions and calculates the energy of system as a function of only the nuclear positions and can evaluate systems containing thousands of atoms. Although MD lacks a detailed description of electronic interactions, it can be used to model the conformation of a fluorescent probe in a macromolecule over hundreds of nanoseconds. The conformation of a probe is important for understanding experimentally observed fluorescence. In terms of 6-MI, the fluorescent populations ($\tau_1$, $\tau_2$, $\tau_3$) observed by TR-fluorescence are assumed to be related to the ground state geometry of 6-MI within the DNA. A sequence dependent molecular conformation of 6-MI may explain the difference in excited state decays.

In MD the molecular motions for large biological systems are simulated by successively integrating Newtonian laws of motion in very small time steps according to empirical force fields. During this process a chronological sequence of structures is generated called a trajectory. The trajectory is then analyzed to obtain molecular information. Force fields are developed to accurately and efficiently approximate physical interactions and experimental observations. Molecular mechanics makes several assumptions for developing force fields, such as Born-Oppenheimer approximation, transferability of force field from one molecule to another, and accurate description of the intra and inter molecular forces.
There are a number of molecular modelling force fields in use today which are based on four components. The four components include functions describing the energetic penalty associated with the deviation of bonds and angles away from their ‘reference’ or ‘equilibrium’ values, a function to describe how the energy changes as bonds are rotated, and a term describing non-bonded interactions. A general functional form for a force field that can be used to model macromolecules is given by:  

\[ V(r^N) = \sum_{\text{bonds}} (l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 + \sum_{\text{torsions}} \left( \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} (4\epsilon_{ij} \left[ \frac{(\sigma_{ij})^{12}}{r_{ij}} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right) \]  

In equation 3.44 \( V(r^N) \) represents the potential energy as a function of the positions (\( r \)) of \( N \) particles. The contributions to equation 2.44 are represented as diagrams in Figure 3-2. The first term in equation 3.44 defines the interaction between pairs of bonded atoms; a harmonic potential is used to model the increase in energy as the bond length \( l_i \) deviates from the reference value \( l_{i,0} \). The second summation is over all valence angles in the system using a harmonic potential, the valence angle is defined as the angle formed between three atoms A-B-C with A and C bonded to B. The third term, the torsional potential defines the change in energy as the bonds in the system rotate. Finally, the fourth term defines the non-bonded contributions to the potential energy. The term is calculated between all pairs of atoms (\( i \) and \( j \)) that are in different molecules or separated by at least three bonds in the same molecule (i.e. have a 1, \( n \) relationship where \( n \geq 4 \)). In the simplest force fields the non-bonded terms are usually modelled using a Coulomb potential for electrostatic interactions and a Lennard-Jones potential for van der Waals interactions.  

\[ 242,243 \]
\[ V(r^N) = \sum_{\text{bonds}} (l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 + \sum_{\text{torsions}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) + \sum_{\text{Coulombic}} \frac{q_i q_j}{4\pi \epsilon_0 r_{ij}} + \sum_{\text{van der Waals}} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \]

The equation represents potential energy force field. Each component of the force field is represented by the adjacent diagram. The equation is used to describe bonded and non-bonded interactions between atoms in a molecular mechanic simulation.
The potential energy force field has been improved for better representation of molecular systems. In equation 3.44 the bond and angles terms use Hooke’s law formula to describe how the energy varies with the square of displacement (at the bottom of the well this is a great approximation); a more appropriate model of deviations away from equilibrium is the inclusion of cubic and higher order terms. In addition, An improper torsional potential can be introduced to establish planarity of ring systems and cross terms for modeling the coupling of internal coordinates, which improves the agreement vibrational spectra.242

A large portion of the improvement to the force field is focused on the non-bonded interactions. The non-bonded terms in a force field are divided into two categories electrostatic interactions and van der Waals interactions. The non-bonded forces do not depend on specific bonding relationship between atoms, the interactions are considered ‘through-space’ interactions and are usually modeled as a function of some inverse power of distance. A cut-off distance is used to limit the number of forces which must be evaluated at each time step. The electrostatic interactions describe elements that attract electrons (electronegative) and elements with less attractions or repulsion for electrons, the distribution of these elements in a molecule give rise to local molecular charge. A computational efficient method to describe charge distribution in a molecule is approximating the molecule as single element using multipolar expansions, which includes the charge \( (q) \), dipole \( (\mu) \), quadrupole \( (\Theta) \), and octopole \( (\Phi) \); where the lowest non-zero moment is the most important component in the model. Using this method the Coulombic energies can be approximated with a set of distances, orientations, and multipolar expansions for a larger sets of pair-wise atomic interactions. This approximation however, fails for atoms at close range and needs to be replaced with an explicit point charge.242
Non-bonded interactions in the system cannot be completely explained by electrostatics. The deviation of rare gases from ideal gas behavior is an obvious example because these atoms contain no dipole-dipole or dipole-induced-dipole interactions. To explain this behavior a different term is necessary, known as van der Waals forces. The force describes an attractive inductive effect due to instantaneous dipoles in a molecule which can induce a dipole in neighboring molecules and a repulsive effect due to the Pauli principle, forbidding electrons to occupy the same space. The best known van der Waals potential function is the Lennard-Jones function (given in equation 3.44.) contains two adjustable parameters: the collision diameter \( \sigma \), the separation between atoms for which the energy is zero, and the energy well depth \( \epsilon \). The parameters are graphically illustrated in Figure 3-2; in this figure \( r^* \) is the separation distance where the energy is at a minimum.

Non-bonded interactions requires assignment of atomic partial charge, which is a computational challenge because partial charge is not a precisely defined quantity. One method used to address charge assignment is accurate reproduction of defined molecular potentials. An early example is the modeling of the electrostatic potential of formamide, methanol, and formic acid.\(^{244}\) The approach was to place a positive point charge at various locations beyond the van der Waals contact surface and calculation of an \textit{ab initio} molecular electrostatic potential. The partial atomic point charges were then determined by least squares fitting the classical electrostatic Coulomb potential to the molecular orbital potential with fixed electroneutrality and experimental dipole vectors.
Applying the electrostatic potential method to larger molecules was problematic. The atomic point charges derived this way were overly sensitive to conformational changes, and that partial charges on buried atoms can be assigned artificially high charges. One approach to alleviate this problem is to derive charges using the restrained electrostatic potential (RESP) method. The method applies a harmonic penalty function to non-hydrogen atoms partial charge magnitude during least squares fitting. The restraints reduce charges on some atoms and vary less with molecular conformation.

The MD model software Assisted Model Building with Energy Refinement (Amber) has been built around the concept of accurate pair-wise charges using the RESP method to biological building blocks such as amino acid nucleic acid. The residues are constrained to have a unit charge, which makes the model transferable to other systems; new systems can be built by linking building blocks without refitting for each model. The Cornell et al. force field was a major breakthrough implementing RESP derived point charges for Coulombic interactions on all non-bonded atoms and atoms separated by more than three bonds, the harmonic bond and angle terms were parameterized to fit observed vibrational modes, and most torsional terms were defined with two atoms on the rotational axis using a three-fold symmetric potential. VDW parameters were optimized in Monte Carlo simulations to reproduce experimental densities and enthalpies of vaporization. Validation of the force field was achieved by successfully reproducing a variety of protein crystal structures. The force field was further revised in 1999, mainly to include additional torsional terms for nucleic acids (designated parm99 in Amber) has proved fairly successful in predicting nucleic acid structures, energies, and dynamics. However, in longer simulations (more than 10 ns) an overpopulations of the alpha/gamma = ( g+, t) backbone
were seen; the parm99 force field was reparameterized using high level quantum mechanical data to generate the parmbsc force field, simulations were extensively compared to experimental data. The set of validation simulations includes two of the long trajectories for the DNA duplex (200 ns each) and the large variety of NA structures (15 different NA families and 97 individual structures).

Before the start of an MD simulation the system must be parameterized to account for the potential force acting on each atom. A molecular dynamics simulation is a deterministic method, meaning the state of the system at any time point can be predicate from its current state. The sets of atomic positions are derived in order by applying Newton’s equation of motion. During the MD simulation the force on each atom is assumed to be constant over a specified time step. The atoms are moved to new positions based on the applied forces, an updated set of forces is then computed; the sequence is then repeated. Evaluation of system requires selecting the appropriate time step to ensure realistic collisions and efficient trajectory production. Generally, position, velocity, and force at time \( t \) specify position and velocity at time \( (t + dt) \), but a number of integration algorithms have been designed to improve the computational efficiency of this task.

The Amber molecular dynamics software package is a suite of programs that work together to carry out and analyze molecular dynamic simulations. The Amber suite contains a number of force fields parameterized for protein, nucleic acid, and carbohydrates simulations with explicit solvent; other force fields can be used within the Amber. The principal flow of information within Amber is shown in Figure 3-3, there are three main steps shown top to bottom in the figure: system preparation, simulation, and trajectory analysis.
Information flow in Amber, the molecular dynamics simulation suite (adapted from Case et al.)

Figure 3-3 Amber flowchart
The main preparation programs include: *antechamber* used to assemble force fields for residues of organic molecules that are not part of the standard libraries and *LEaP* which constructs biopolymers from the component residues, solvates the system, and prepares lists of force field terms and their associated parameters; the parameter topology and input coordinate files are used for MD minimization and production of MD trajectories (Figure 3-3).\textsuperscript{246} In the Amber suite the main minimization and molecular dynamics programs are *sander* and *pmemd* (Figure 3-3). The parallel programs use the message passage interface (MPI) to communicate among processors. The *pmemd* program is optimized for performance but does not support all options found in *sander*.\textsuperscript{246} The task of analyzing MD trajectories is tackled using *ptraj* and *mm-pbsa* programs. The *ptraj* program is able to assemble the total time course of the MD simulation from multiple trajectories, strip groups of atoms (such as solvent), and execute a number of common analysis tasks. The *mm-pbsa* program is used to estimate the energies and entropies from the snapshots contained within trajectory files.\textsuperscript{246}

A goal for the work presented in this thesis is to extract a conformation from an *in silico* ensemble that is an approximation of what the TCSPC set-up extracts from the *in vitro* ensemble. To achieve this goal requires analyzing the sequence dependent conformation of 6-MI within duplex DNA. The conformations can then be grouped based on their corresponding excited state decay. The hypothesis is sequences with similar excited state decays will also have analogous 6-MI conformations. To determine the conformation of 6-MI within an oligomer requires the following steps: i) create DNA sequence ii) replace guanine base with 6-MI iii) add ions and solvate system iv) energy minimize and produce trajectory for an oligomer v) analyze the trajectory for 6-MI conformation.
Amber molecular dynamics simulations were run on the CAFTC07, ATFAA15, aAAFTA10, GTFTG32, ATFTA19, CTFAC07, CAGTC07 (no 6-MI substitution), ATGAA07, and A(mis)_ATFAA07 DNA double strands. Molecular dynamic simulations require an initial structure; The B-form DNA conformation was chosen based on its prevalence in vivo. The initial B-form DNA structures were generated using the Make-Na server with Nucleic Acid Builder. The DNA structures were then modified by replacing the specified guanine with 6-MI.

No software is available for mutating a guanine base to 6-MI nucleobase, so an algorithm was developed in Mathematica 7.0. The algorithm is based on a previously described method. The Transformers method assumes that nucleobases are planar and that the position of the glycosidic nitrogen in a DNA structure remains fixed under mutation of that nucleobase. The transformers method requires parameterizing the position of the specified nucleobase by three atoms (A, B, C). The glycosidic nitrogen N9 is designated Atom A. The nitrogen furthest from the glycosidic nitrogen is designate Atom B. For pyrimidines, the carbon β to the glycosidic nitrogen was designated Atom C. For purines, the non-glycosidic nitrogen on the five-membered ring was designated Atom C. A transformation matrix was then generated which placed the nucleobase in the reference orientation: Atom A at the Cartesian origin (0, 0, 0), Atom B along the x-axis (x, 0, 0), and Atom C in the xy-plane (x’, y’, 0). The inverse of the transformation matrix was then applied to the Cartesian coordinates of an energy minimized 6-MI nucleobase in the reference orientation. The energy-minimized structure of 6-MI nucleobase was generated in Gaussian 03. The specified guanine base was then mutated with the transformed 6-MI nucleobase using the program PDB editor to manipulate the PDB file.
The transformation matrix used in the Transformers algorithm is defined by three rotational matrices and one translation matrix. The translation matrix represents a matrix which translated Atom A to the Cartesian origin. Defining the three rotation matrices requires applying the translation matrix to Atom A, B, and C. Next, a vector (norm) was then defined, which is normal to the plane defined by Atom B and C. Norm is calculated as the cross product of Atom B and Atom C. The first rotation matrix (rotate-x) was defined as a matrix to rotate vector norm about x-axis to eliminate the y-component. The second rotation matrix (rotate-y) was defined as a matrix, which rotated vector (norm) about the y-axis to eliminate the x-component. The rotate-x and rotate-y matrices were then applied to Atom B and C to move them into the x-y plane. The final rotation matrix (rotate-z) was defined by rotating Atom B about the z-axis to eliminate its y-component. These matrices are then multiplied together to generate the transformation matrix.

To execute the isothermal-isobaric ensemble (NPT) simulation using the Amber program suite requires a force field be defined for all atoms. The nucleic acids in the simulation were parameterized using the parmbsc0 force field. A modified force field was necessary for fluorescent base analog 6-MI. This first required a geometry optimization of 6-MI nucleobase with HF/6-31G* level of theory in Gaussian03. A single-point energy calculation at the MP2/6-31G* level of theory was then used to calculate the electrostatic potential of the geometry optimized 6-MI in Gaussian03. The modified force field was then generated using the parameters obtained from the single-point calculation in the antechamber program. The LEaP program was then used to prepare the modified structures for MD. Sodium and chloride counterions were added for electric neutrality; monovalent ion
parameters were specified using ions08 parameters. The system was placed in an octahedral unit cell and solvated with TIP3P explicit water model, the water molecules are rigid structures with a point charge on each atom and a single Lennard-Jones VDW force centered on the oxygen atom. The system was then energy minimized in two steps with periodic boundary conditions and a 10 angstrom cutoff. In the initial minimization, DNA coordinates were restrained with a harmonic potential of 500 kcal mol$^{-1}$ Å$^{-2}$ and the geometry of the solvent and ions were relaxed using 500 steps of steepest decent followed by 500 steps along the conjugate gradient. The second minimization process relaxed the entire system with 1000 steps of steepest decent and 1500 steps along the conjugate gradient.

Energy was added to the system by heating to 300 K at constant volume over a period of 20 ps, maintaining weak harmonic restraints of 10 kcal mol$^{-1}$ Å$^{-2}$ on the solute atoms. The production MD integration time step was 2 fs for a total of 10 ns. To maintain the temperature of the system the Langevin thermostat was set at 300 K. The SHAKE constraint was applied to all bonds involving hydrogen atoms. During the subsequent simulation a constant pressure periodic boundary with $P = 1$ atm was maintained using isotropic position scaling with relaxation time of 2 ps. The output files and trajectory files were updated every 2 ps or 1000 steps. The simulations were analyzed using several methods to determine the simulation had equilibrated and examine the structural similarities and differences between constructs.

The system was deemed equilibrated by analyzing the rmsd and total energy of the system. The final 9ns of MD were analyzed as snapshots of every 10 ps stripped of water. Each 10 ps snapshot was then analyzed using the program 3DNA. The base-pair steps were parameterized by slide, shift, rise, tilt, roll, and twist based at the local reference frame.
Applying 6-MI enhanced fluorescence to examine Protein–DNA interactions in the picomolar range

4.1 Introduction

Many cellular processes including the regulation of replication, transcriptional regulation, recombination and DNA compaction require the action of DNA-binding proteins. A commonly used and powerful method for studying these interactions is fluorescence spectroscopy. Typically, the fluorescent probes used for studying these interactions are covalently attached to the DNA and can be divided into two categories: external or internal. External fluorescent probes are extremely useful for investigating binding interactions and have the benefit of exhibiting high quantum yields, long lifetimes, and the ease of incorporation onto an oligomer. Disadvantages of using an external probe may arise from the molecular properties of the probe, such as the chemical nature of the probe, size, charge, and method of attachment to DNA. For example, attachment to DNA through a 6-carbon linker limits the ability of the probe to report on DNA local structure and dynamics and may cause the probe to be a poor reporter of protein-induced structural perturbations upon binding.

Internal probes, in many cases fluorescent nucleoside analogs, have greatly increased the amount of information obtained regarding local DNA structure and dynamics. These probes can be viewed as providing complementary information to external probes, as they are more sensitive to local interactions. Nucleoside analogs minimally distort DNA structure and are able to base stack with adjacent bases. This structural similarity of internal probes to nucleic acid bases and the ability to hydrogen bond with their cognate base have made these probes ideal for examining protein-DNA interactions. In addition, the fluorescent
properties of internal probes are highly sensitive to changes in DNA structure and can be used
to monitor local protein-induced perturbations.\textsuperscript{13-15} For example, when coupled with time-
resolved fluorescence measurements the internal probe can provide dynamic resolution of
DNA structure on the single base scale.\textsuperscript{6}

A deterrent to using internal probes is a significant reduction in the quantum yield with
incorporation into single-stranded DNA, which is either maintained or further decreased upon
formation of duplex DNA.\textsuperscript{2,5} This reduction in quantum yield lowers sensitivity and
introduces a concentration limit for using internally labeled oligomers. Many DNA-binding
proteins exhibit dissociation constants in the low nanomolar range, which is below the signal
to noise threshold for many fluorescent nucleoside analogs. Thus, the study of these
interactions using internal fluorescent probes has remained a challenging problem.

We herein report on the identification of a specific DNA sequence that enhances 6-MI
fluorescence upon duplex formation and demonstrate its utility in examining protein-DNA
interactions. The enhanced 6-MI fluorescence was initially observed in the recognition
sequence of the \textit{E. coli} protein Integration Host Factor (IHF) when studying IHF-DNA
interactions (Figure 1; Table 4-1). The local sequences (ATFAA and AAFTA) give rise to
the enhanced fluorescence, in which adenine and thymine are the nearest neighbors of the 6-
MI probe. A sequence-dependent enhancement of 6-MI fluorescence upon duplex formation
was previously observed by Knutson and co-workers, when the 6-MI was flanked by thymine
residues (i.e. GTFTG).\textsuperscript{69} The fluorescent pentamers identified in this study are significantly
brighter than the GTFTG sequence identified by Knutson and co-workers ($\Phi_{\text{rel}} = 0.27$ vs. 0.75
for ATFAA) (Tables 1 and 2). Lifetime and quantum yield measurements indicate a decrease
in dynamic, collisional quenching of the excited state by adjacent bases as a plausible
mechanism for the increased fluorescence. The significantly increased fluorescence of the
Figure 4-1 IHF crystal structure

The IHF-DNA X-ray co-crystal structure depicted with the positions of 6-MI in the ATFAA and AAFTA sequences highlighted (AAFTA10 = magenta, ATFAA15 = green, ATFAA07 = navy blue). The IHF consensus binding sequence is shown in yellow. Figure was generated with Pymol using coordinates from pdbid, 1IHF.
ATFAA sequence allows measurement of binding at DNA concentrations as low as 50 pM and anisotropy measurements at a 1 nM concentration of DNA. The increased sensitivity of the ATFAA sequence context is used to investigate sequence-specific and non-sequence-specific protein-DNA interactions. We further demonstrate that the enhanced 6-MI fluorescence coupled with time-resolved lifetime and anisotropy measurements provides unique insight into DNA structure in the context of protein-DNA interactions.

4.2 Results and Discussion:

4.2.1 Duplex formation leads to enhanced fluorescence of 6-MI in distinct sequences

We have examined the fluorescence behavior of the nucleoside analogue, 6-methylisoxanthopterin (6-MI) in several different DNA molecules. The incorporation of this analogue minimally perturbs DNA duplex stability as measured by thermal melting (Figure 4-2). In our studies, fluorescent enhancement of nucleoside analog 6-MI in duplex DNA was observed in oligomers containing the consensus sequence for the DNA-binding protein Integration Host Factor (IHF) (Figure 4-1). 6-MI was incorporated at several unique positions flanking the IHF consensus sequence to study the effect of IHF binding on local DNA dynamics (Table 4-1). In previous studies of 6-MI and other fluorescent nucleoside analogs,60,65 fluorescence intensity remains quenched upon duplex formation as observed with the CAFTC07 oligomer (Φ_{rel,ds} = 0.24 ± 0.05, Φ_{rel,ss} = 0.19 ± 0.04) (Figure 2; Table 4-2). In contrast to the previously observed quenching, 6-MI displayed an ~4-fold or greater increase in fluorescence intensity upon duplex formation within the ATFAA15 oligomer (Φ_{rel,ds} =0.82 ± 0.03, Φ_{rel,ss} =0.20 ± 0.05) and AAFTA10 oligomer (Φ_{rel,ds} =0.73 ± 0.03, Φ_{rel,ss} =0.15 ± 0.05) (Figure 4-3, Table 4-2). Examination of local sequence demonstrates that the identity of the n ± 2 base differed between the highly fluorescent duplex sequences and the
Table 4-1 Sequences of 6-MI-containing and complementary oligomers

<table>
<thead>
<tr>
<th>Sequence name</th>
<th>Sequence$^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAFTC07</td>
<td>5'-TAT GCA FTC ACT ATG AAT CAA CTA CTT AGA TGGT-3'</td>
</tr>
<tr>
<td>ATFAA15</td>
<td>5'-TAT GCA GTC ACT $^F$AT AAT CAA CTA CTT AGA TGGT-3'</td>
</tr>
<tr>
<td>AAFTA10</td>
<td>5'-ACC ATC TAA $^F$TA GTT GAT TCA TAG TGA CTG CAT A-3'</td>
</tr>
<tr>
<td>ATFAA07</td>
<td>5'-TAT GAT FAA ACT ATG AAT CAA CTA CTT AGA TGGT-3'</td>
</tr>
<tr>
<td>GTFTG32</td>
<td>5'-ACT AGA GAT CCC TCA GAC CCT TTT AGT CAG $^F$TFT GGA-3'</td>
</tr>
<tr>
<td>GAFAA15</td>
<td>5'-GCT TGT TCA CGG $^F$GA $^F$AAG GGG AAG ACC ACG G-3'</td>
</tr>
<tr>
<td>JX$^2$</td>
<td>5'-CCA GAA $^F$TA AGT TGA GTC CTT GCT AGG ACG GAG G-3'</td>
</tr>
<tr>
<td>Msh_ATFAA15</td>
<td>5'-TAT GCA GTC ACT $^F$AT AAT CAA CTA CTT AGA TGGT-3'</td>
</tr>
<tr>
<td></td>
<td>ATA CGT CAG TGA T-T TTA GTT GAT GAA TCT AC</td>
</tr>
<tr>
<td>Msh_ATFAA09</td>
<td>5'-ATG TGA $^F$AT AAT ATG GTA TAT ATC TGC TGA AGG AAA T-3'</td>
</tr>
<tr>
<td></td>
<td>TAC ACT TAC TTA TAC CAT T TTA TAG ACG ACA ACC TTT A</td>
</tr>
<tr>
<td>TBP_ATFAA23</td>
<td>5'-GCC CAT TCG CTA TAA AAG $^F$GA $^F$TA AGA GC-3'</td>
</tr>
<tr>
<td></td>
<td>CGG GTA AGC GAT ATT TTC CCT ACT TCT CG</td>
</tr>
<tr>
<td>HU_ATFAA15</td>
<td>5'-TAT GCA GTC ACT $^F$AT AAT CAA CTA CTT AGA TGGT-3'</td>
</tr>
<tr>
<td></td>
<td>ATA CGT CAG TGA TAC TTA</td>
</tr>
<tr>
<td>4WJ$^3$</td>
<td>5'-CCA GAA $^F$TA AGT TGA GTC CTT GCT AGG ACG GAG G-3'</td>
</tr>
<tr>
<td></td>
<td>5' CCT TCA ACC ACC GCT CAA CTC AAC TTC ATT CTG G-3'</td>
</tr>
<tr>
<td></td>
<td>5' CCT CCG TCC TAG CAA GGG GCT GCT ACC GGA AGG G</td>
</tr>
<tr>
<td></td>
<td>5' CCC TTC CGG TAG CAG CCT GAG CGG TGG TTG AAG G</td>
</tr>
</tbody>
</table>

$^1$The 6-MI position is indicated in red (F) and the pentamer sequences containing 6-MI are highlighted in yellow. Boldface highlights the protein-binding recognition sequence or +T insertion. $^2$JX is the sequence containing 6-MI used to form the $^3$4WJ, a four-way formed with the given strands.
The temperature-dependent melting curve monitored at 260 nm of the WT 34bp dsDNA (black) based on the H1 consensus binding sequence of Integration Host Factor (IHF). Also shown are curves generated with dsDNA containing 6MI, CAFTC07ds (red), ATFAA15ds (green), and AAFTA10ds (blue). A similar Tm (330±1.0K) is obtained for all the duplexes, demonstrating that introduction of 6-MI does not perturb DNA stability. Analysis of temperature melts was performed as outlined in Chapter 3.4
quenched duplex sequence. For the fluorescent duplexes, ATFAA15 and AAFTA10 the n ± 2 positions were adenine whereas for the quenched duplex, CAFTC07 the n ± 2 positions were cytosine (Table 4-1). The highly fluorescent duplexes, ATFAA15 and AAFTA10, contained a single purine adjacent to the 6-MI with adenine located at the n+1 (ATFAA15) and the n-1 position (AAFTA10). As the base composition of the ATFAA and the AAFTA sequences are the same, these findings suggest that subtle changes in base sequence can lead to altered fluorescence properties as has been previously observed for the fluorescent nucleoside analog, 6MAP.60 Finally, to determine whether duplex context influences the fluorescence properties, we replaced the CAFTC sequence with the ATFAA sequence (ATFAA07) in our original duplex sequence. As shown in Figure 4-3, formation of the ATFAA07 duplex leads to enhanced 6-MI fluorescence similar to the ATFAA15 and AAFTA10 duplexes and in distinct contrast to the CAFTC07 duplex (Table 4-2).

Previous studies, focused on sequence-dependent quenching of 6-MI in DNA, showed that the presence of purines adjacent to 6-MI (or 3-methylisoxanthopterin (3-MI)) resulted in greater fluorescence quenching than pyrimidines.21,54,64 In a recent study, Knutson and co-workers reported on another sequence which exhibited enhanced 6-MI fluorescence upon incorporation into duplex DNA. Specifically, they observed a 70% increase in quantum yield from 0.12 to 0.19 when 6-MI was flanked by thymine residues at the n ± 1 position (i.e. GTFTG).69 The enhanced fluorescence observed for 6-MI in the GTFTG sequence, was attributed to restricted dynamic motion within the duplex.69 Unlike this study, the effect of bases extending beyond the n ± 1 position was not considered; although we note the n ± 2 residues are G for this sequence.
Fluorescence spectra of ss DNA and ds DNA containing 6-MI are shown based on their quantum yields measured relative to 6-MI monomer (black) using an excitation wavelength of 340 nm. The $\Phi_{\text{rel}}$ of 6-MI in ssDNA (teal, pink, brown, dark green) is highly quenched ($\Phi_{\text{rel}} < 0.2$). The dsDNA molecules containing either the AFTAA or the AAFTA sequences exhibit significantly higher quantum yields ($\Phi_{\text{rel}} \geq 0.7$) (blue, magenta, green). While formation of the CAFTC07 duplex has only a slight effect on the quantum yield. (B) The quantum yields of different DNA sequences containing 6-MI are shown relative to 6-MI monomer (ex=340nm), indicating the ATFAA enhanced fluorescence is maintained in different constructs (green, cyan, red, magenta, blue, orange). The 6-MI in GAFAA15ds (magenta) is highly quenched $\Phi_{\text{rel}}$ is <0.1). Although formation of the TFT32ds duplex leads to an increase in $\Phi_{\text{rel}}$ to 0.27, the $\Phi_{\text{rel}}$ of ATFAA constructs is at least 0.7 or greater.
Table 4-2 Parameters derived from fitting of time-resolved fluorescence decay curves

Parameters\(^1\) derived from fitting of time-resolved fluorescence decay as outlined Chapter 3.7

<table>
<thead>
<tr>
<th>DNA molecule</th>
<th>(\alpha_1)(^2) ±0.01</th>
<th>(\tau_1)(^2) (ns) ±0.12</th>
<th>(\alpha_2) ±0.01</th>
<th>(\tau_2) (ns) ±0.02</th>
<th>(\alpha_3) ±0.01</th>
<th>(\tau_3) (ns) ±0.04</th>
<th>(\tau_{(\text{mod})}) (ns) ±0.03</th>
<th>(\Phi_{\text{rel}}) SS/(\Phi_{\text{rel}}) TR ±0.05</th>
<th>Solvent Exposed Fraction ±0.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-MI</td>
<td>1</td>
<td>6.57</td>
<td>6.57</td>
<td>1/1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAFTC07ss</td>
<td>0.62</td>
<td>0.45</td>
<td>0.21</td>
<td>2.59</td>
<td>0.17</td>
<td>7.27</td>
<td>5.14</td>
<td>0.19/0.31</td>
<td>0.32</td>
</tr>
<tr>
<td>CAFTC07ds</td>
<td>0.81</td>
<td>1.05</td>
<td>0.16</td>
<td>1.97</td>
<td>0.04</td>
<td>8.04</td>
<td>2.65</td>
<td>0.24/0.23</td>
<td>0.5</td>
</tr>
<tr>
<td>ATFAA15ss</td>
<td>0.49</td>
<td>0.61</td>
<td>0.31</td>
<td>2.9</td>
<td>0.2</td>
<td>6.78</td>
<td>4.72</td>
<td>0.20/0.39</td>
<td>0.46</td>
</tr>
<tr>
<td>ATFAA15ds</td>
<td>-</td>
<td>-</td>
<td>0.05</td>
<td>3.62</td>
<td>0.95</td>
<td>7.19</td>
<td>7.1</td>
<td>0.82/1.06</td>
<td>0.17</td>
</tr>
<tr>
<td>ATFAA15ds + IHF</td>
<td>-</td>
<td>-</td>
<td>0.1</td>
<td>3.74</td>
<td>0.9</td>
<td>6.99</td>
<td>6.82</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>ATFAA07ss</td>
<td>0.61</td>
<td>0.52</td>
<td>0.23</td>
<td>2.91</td>
<td>0.17</td>
<td>6.96</td>
<td>4.76</td>
<td>0.15/0.33</td>
<td>0.31</td>
</tr>
<tr>
<td>ATFAA07ds</td>
<td>-</td>
<td>-</td>
<td>0.05</td>
<td>3.12</td>
<td>0.95</td>
<td>6.93</td>
<td>6.84</td>
<td>0.73/1.02</td>
<td>0.2</td>
</tr>
<tr>
<td>ATFAA07ds + IHF</td>
<td>-</td>
<td>-</td>
<td>0.07</td>
<td>2.85</td>
<td>0.93</td>
<td>6.87</td>
<td>6.75</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>AAFTA10ss</td>
<td>0.65</td>
<td>0.66</td>
<td>0.26</td>
<td>2.27</td>
<td>0.09</td>
<td>7.38</td>
<td>3.71</td>
<td>0.11/0.25</td>
<td>0.68</td>
</tr>
<tr>
<td>AAFTA10ds</td>
<td>-</td>
<td>-</td>
<td>0.07</td>
<td>2.79</td>
<td>0.93</td>
<td>7.82</td>
<td>7.7</td>
<td>0.7/1.14</td>
<td>0.14</td>
</tr>
<tr>
<td>AAFTA10ds +IHF</td>
<td>0.07</td>
<td>1.77</td>
<td>0.17</td>
<td>4.73</td>
<td>0.76</td>
<td>7.7</td>
<td>7</td>
<td>0.82/1.06</td>
<td>0.5</td>
</tr>
<tr>
<td>Msh_ATFA A15ds</td>
<td>0.14</td>
<td>1.24</td>
<td>0.43</td>
<td>4.02</td>
<td>0.43</td>
<td>6.65</td>
<td>5.56</td>
<td>0.7/0.75</td>
<td>0.14</td>
</tr>
<tr>
<td>Msh_ATFA A15ds +Msh2-Msh6</td>
<td>0.18</td>
<td>1.24</td>
<td>0.48</td>
<td>3.86</td>
<td>0.34</td>
<td>6.65</td>
<td>5.03</td>
<td>0.78/1.06</td>
<td>0.5</td>
</tr>
<tr>
<td>Msh_ATFA A09ds</td>
<td>-</td>
<td>-</td>
<td>0.08</td>
<td>2.89</td>
<td>0.92</td>
<td>7.37</td>
<td>7.22</td>
<td>0.78/1.06</td>
<td>0.5</td>
</tr>
<tr>
<td>Msh_ATFA A09 +Msh2-Msh6</td>
<td>-</td>
<td>-</td>
<td>0.09</td>
<td>2.87</td>
<td>0.91</td>
<td>7.37</td>
<td>7.2</td>
<td>0.82/1.06</td>
<td>0.5</td>
</tr>
</tbody>
</table>

\[ I(t) = \sum_{i=1}^{x} \alpha_i e^{-\left(\frac{t}{\tau_i}\right)} \]

\(^1\)Fluorescence decay curves were fit to a sum of exponentials: \(^2\)For all parameters standard deviations are given in parentheses.
In this study, we address whether the mechanism of enhanced fluorescence observed in the ATFAA15 and AAFTA10 sequences arises from a unique, constrained conformation of 6-MI, as observed in the GTFTG sequence of Knutson and co-workers\textsuperscript{69} or from a largely extrahelical population of 6-MI as a consequence of the surrounding sequence Figure 4-4.

4.2.2 Sequence-dependent solvent exposure of 6-MI

To parse out these effects, the relative levels of solvent exposure of 6-MI in different sequence contexts were measured by fluorescence quenching. Quenching of 6-MI fluorescence intensity was measured through titration of KI, which collisional quenches the excited state. Stern-Volmer analysis of time-resolved and steady-state fluorescence quenching of 6-MI monomer, reveals linear relationships between the fluorescence intensity and quencher demonstrating that the process is predominately dynamic (Figure 4-5). When 6-MI was incorporated into DNA, the Stern-Volmer plots deviate from linearity indicative of a population not accessible to solvent (Figure 4-6). In ssDNA, solvent exposure of 6-MI was reduced by 40% or more for all sequences (Figure 4-6, Table 4-2). For the CAFTC07ds duplex the fraction of 6-MI accessible to solvent increases to 50 ± 6% from 30% in the single strand. In contrast, the sequences that exhibit enhanced fluorescence experience a decrease in solvent exposure upon duplex formation: 17 ± 2% for ATFAA15ds, 20 ± 7% for AAFTA10ds, and 14 ± 3% for ATFAA07ds (Figure 4-6, Table 4-2).

Similar behavior has been observed for 2-aminopurine (2-AP), where steric constraints introduced by specific double-stranded DNA sequences led to a reduction in probe solvent exposure. For example, when 2-AP was placed in the TATA duplex sequence it exhibited a slower rotational correlation time and reduced solvent exposure, consistent with a restrained
The proposed models for enhancement of fluorescence of 6-MI. Model A: 6-MI within the pentamer sequence ATFAA is largely extrahelical. Thus, experiencing an environment similar to the monomer. Model B: 6-MI is basepaired with cytosine in a sterically constrained conformation with minimal interaction with adjacent bases.
Figure 4-5 6-MI monomer KI quenching

The comparison between the excited state ($\tau_0/\tau$) and steady-state ($F_0/F$) KI quenching of 6-MI monomer reveals there is minimal difference in quenching rates. The similarities in quenching rates demonstrate KI quenching of 6-MI fluorescence is largely collisional.
The Stern-Volmer plots for KI fluorescence quenching of duplex DNA (CAFTC07ds, ATFAA15ds, and ATFAA07ds) reveals non-linear quenching of 6-MI when incorporated into duplex DNA. Deviation from linearity is consistent with a fraction of the population not accessible to quencher. (B) The solvent accessibility in ssDNA and dsDNA measured relative to 6-MI monomer by fluorescence collisional quenching with KI. For duplexes containing either the ATFAA or the AAFTA sequence, the formation of dsDNA leads to a reduction in 6-MI solvent accessibility relative to ssDNA and 6-MI monomer, which suggests 6-MI is not extrahelical in these sequences (ATFAA15ds, AAFTA10ds, ATFAA07ds). Solvent exposure is increased in the quenched duplex, CAFTC07ds. The solvent exposure was calculated using the modified Stern-Volmer equation (3.6) given in Chapter 3.
motion of the probe. Likewise, 3-MI a close structural analog of 6-MI, is minimally quenched by acrylamide, when incorporated into duplex DNA. Our findings are consistent with hydroxyl radical footprinting methods, which also point to differences in base solvent exposure as a consequence of sequence context. This method requires that the OH∙ cleave the DNA backbone. The hydroxyl radical cleavage intensity for the IHF consensus sequence used in this study were predicted using the OH Radical Cleavage Intensity Database, which determines hydroxyl radical cleavage propensities at a particular position using a sliding tetramer window algorithm based on experimentally observed cleavage patterns. The predicted cleavage intensity for the IHF consensus sequence at the 7th position (CAFTC07) is 1.6 while for the 15th position (ATFAA15) it is only 0.68 (Figure 4-7). Similarly, predicted cleavage intensity for the IHF complementary sequence at the 10th position (cAFT10) is only 0.78. Thus, fluorescence quenching results are in good agreement with predicted OH∙ cleavage propensities and further suggest that 6-MI in the ATFAA and AAFTA sequence contexts is relatively protected from solvent and is unlikely to have a large population that is extrahelical.

4.2.3 The excited state lifetime of 6-MI is sequence dependent

To further probe the influence of local sequence on 6-MI photophysics, we examined the excited state decay of 6-MI in the different sequence contexts. In aqueous solution 6-MI monomer exhibits a mono-exponential excited state lifetime of ~6.5 ns. Upon incorporation
Figure 4-7 Predicted OH cleavage pattern

Cleavage intensity for WT oligomer predicted using (O)H (R)adical (C)leavage (I)ntensity (D)atabase (ORChID database). The predicted cleavage intensities are consistent with the measure fraction of 6-MI accessible to KI quencher. For example, there is a larger degree of cleavage observed for the CAFTC07 position (1.6) compared to the ATFAA15 position (0.69)
of 6-MI into ssDNA, much shorter lifetimes are observed, characterized by a complex, multiexponential decay of the excited state (Figure 4-8, Table 4-2). The decay rates are attributed either directly to the rate of quenching for a particular conformation or to the rate of change to an efficiently quenched conformation.

The excited state decay of the 6-MI-containing ssDNA molecules are similar and can be understood in terms of the observed decay components and fractional subpopulations (Figure 4-8, Table 4-2). The largest fractional subpopulations ($\alpha_1 = 0.5$-$0.8$) typically correspond to the shortest lifetime ranging from $0.45$-$0.66$ ns. The next largest fractional subpopulations ($\alpha_2 = 0.21$-$0.31$) are associated with an intermediate lifetime of $2.7$ ns on average. The smallest fractional subpopulation ($\alpha_3 = 0.09$-$0.20$) is associated with the longest lifetime of $\sim 7.0$ ns (Figure 4-8, Table 4-2).

An examination of the excited state lifetimes of the dsDNA sequences that exhibit enhanced fluorescence upon duplex formation reveal pronounced changes in excited state dynamics relative to the single strand. A significant shift in the fractional subpopulations for dsDNA is observed compared to ssDNA; where the largest subpopulation for the dsDNA sequences ($\alpha_3 = \sim 0.9$) is associated with the longest-lived component ($\tau_3 = \sim 7.0$ ns), while the largest population for ssDNA ($\alpha_1 = \sim 0.55$) is associated with the shortest lifetime component ($\tau_1 = \sim 0.5$ ns) (Figure 4-8, Table 4-2). Importantly, this shift in populations is not observed for the quenched duplex, CAFTC07. In addition, the excited state lifetimes for ATF AA15ds, AAFT A10ds, and ATF AA07ds are predominately characterized by a single exponential decay with a lifetime similar to that of 6-MI monomer ($\tau = \sim 6.5$ns). The mean lifetime of the enhanced dsDNA sequences is $\sim 3$ ns longer than for ssDNA, consistent with decreased quenching interactions between 6-MI and adjacent bases. Thus, the excited state is either a
homogenous population or a continuum of conformations; in both cases interbase quenching interactions are minimized.

The multi-exponential decay observed for 6-MI in ssDNA is similar to that detected for 2AP and suggests that the mechanism of dynamic quenching is probably comparable for both probes. Studies reporting on the composition of excited state subpopulations for 2AP are consistent with discrete populations existing in thermal equilibrium with an energy barrier between states. The multiexponential decay for 6-MI in ssDNA is similarly interpreted, suggestive of the presence of discrete subpopulations for each decay component. Importantly, overlap between these populations is unlikely due to the large differences between the decay times (Table 4-2).

The relative quantum yields based on the time-resolved measurements reflects the radiative component of the fluorescence. Similar to the steady-state data, we observe a significant increase in quantum yield upon duplex formation for the ATFAA15, ATFAA07 and AAFTA10 sequences. For these sequences, the steady-state quantum yields are approximately 30% lower than those determined by time-resolved measurements and are strongly suggestive of a static quenching mechanism. Interestingly, this effect is more pronounced in the single strands, where steady-state quantum yields are more than 50% lower than those determined by time-resolved methods and thus measurements and are strongly suggestive of a static quenching mechanism. Interestingly, this effect is more pronounced in the single strands, where steady-state quantum yields are more than suggest that static quenching mechanisms are even greater than in ds DNA (Table 4-2). It is also possible that the discrepancy between absolute values
Figure 4-8 6-MI excited state decay

(A) Fluorescence lifetime decay curves (ex=375, em=460) of 6-MI containing ssDNA and dsDNA molecules compared to monomer (black). The significantly faster decays for ssDNA (brown, dark green, pink, teal) indicate dynamic quenching of 6-MI occurs upon incorporation into ssDNA oligomers. The dsDNA molecules ATFAA15ds (green), ATFAA07ds (blue), and AAFTA10ds (magenta) exhibit fluorescence decays similar to 6-MI monomer (black), indicative of a lack of quenching. Notably, the CAFTC07 duplex exhibits a fast decay (red). The instrument response function is shown in gray. (B) Fractional populations determined from time-resolved fluorescence decay curves (A) of 6-MI-containing ssDNA and dsDNA molecules relative to monomer. Duplex formation leads to a shift in the peak population to the longest lifetime component (α3) for sequences ATFAA15ds, AAFTA10ds, and ATFAA07ds, consistent with a reduction in the non-radiative rate. For ssDNA oligomers, the largest fractional population is associated with the shortest-lived component (α1). The CAFTC07 duplex exhibits a population distribution similar to the single strands where the largest population is associated with the shortest-lived component (α1).
of the steady-state and time-resolved quantum yields results from a fast decay rate that is not resolved with the current instrumentation.

4.2.4  Local motion of 6-MI is decreased in ATFAA sequence

The mobility of 6-MI within the different sequence contexts was directly assessed by time-resolved fluorescence anisotropy measurements. A two-component exponential decay was found to correctly model the time-dependent anisotropy decay for DNA ss and ds molecules containing 6-MI. All of the ssDNA molecules exhibit a fast rotational correlation time that ranges from 0.13-0.42 ns, indicative of a considerable amount of local motion of 6-MI in the ssDNA (Table 4-3, Figure 4-9). Interestingly, the AAFTA10 sequence exhibits the fastest anisotropic decay of the four single strands. The local motion reflects the stability of the probe, which arises in part from base stacking interactions and has been shown to be sequence dependent.\textsuperscript{34,262} In the CAFTC07 duplex the rotational correlation time associated with 6-MI internal motion ($\theta_l$) is increased to 0.61 ns, consistent with reduced rotational mobility of the 6-MI due to base stacking and base pairing interactions (Figure 4-9, Table 4-3). The ds DNA molecules that give rise to enhanced fluorescence all exhibit a markedly different behavior in which the internal motion of the probe is dramatically reduced and the fastest anisotropic decay observed is $\sim$1.2 ns or greater (Table 4-3). We consider this rotational time too long to be attributed solely to the local motion of 6-MI and suggest this decay also reflects rotation of the duplex or segmental motion.\textsuperscript{262,34} The longer correlation times observed ($\theta_R$) are attributed to the overall tumbling or rotational motion of the DNA molecule.\textsuperscript{34}
Table 4-3 Parameters derived from fitting of time-resolved anisotropy decay curves

Parameters derived from fitting of time-resolved anisotropy decay curves

<table>
<thead>
<tr>
<th>DNA molecule</th>
<th>$\beta_1^2$ (±0.02)</th>
<th>$\theta_L$ (ns) (±0.06)</th>
<th>$\beta_2$ (±0.02)</th>
<th>$\theta_R$ (ns) (±0.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAFTC07ss</td>
<td>0.25</td>
<td>0.30</td>
<td>0.75</td>
<td>3.7</td>
</tr>
<tr>
<td>CAFTC07ds</td>
<td>0.20</td>
<td>0.61</td>
<td>0.80</td>
<td>8.8</td>
</tr>
<tr>
<td>ATFAA15ss</td>
<td>0.22</td>
<td>0.32</td>
<td>0.78</td>
<td>4.4</td>
</tr>
<tr>
<td>ATFAA15ds</td>
<td>0.09</td>
<td>1.66</td>
<td>0.91</td>
<td>16.8</td>
</tr>
<tr>
<td>ATFAA15ds + IHF</td>
<td>0.04</td>
<td>1.68</td>
<td>0.96</td>
<td>26.3</td>
</tr>
<tr>
<td>ATFAA07ss</td>
<td>0.33</td>
<td>0.13</td>
<td>0.67</td>
<td>3.3</td>
</tr>
<tr>
<td>ATFAA07ds</td>
<td>0.16</td>
<td>1.16</td>
<td>0.84</td>
<td>15.6</td>
</tr>
<tr>
<td>ATFAA07ds + IHF</td>
<td>0.1</td>
<td>1.18</td>
<td>0.91</td>
<td>25.9</td>
</tr>
<tr>
<td>AAFTA10ss</td>
<td>0.22</td>
<td>0.42</td>
<td>0.78</td>
<td>5.2</td>
</tr>
<tr>
<td>AAFTA10ds</td>
<td>0.13</td>
<td>1.41</td>
<td>0.87</td>
<td>17.0</td>
</tr>
<tr>
<td>AAFTA10ds + IHF</td>
<td>0.10</td>
<td>0.9</td>
<td>0.9</td>
<td>24.7</td>
</tr>
<tr>
<td>Msh_ATFAA15ds</td>
<td>0.14</td>
<td>0.73</td>
<td>0.86</td>
<td>14.6</td>
</tr>
<tr>
<td>Msh_ATFAA15ds + Msh2-Msh6</td>
<td>0.16</td>
<td>2.11</td>
<td>0.84</td>
<td>20.7</td>
</tr>
<tr>
<td>Msh_ATFAA09ds</td>
<td>0.16</td>
<td>2.33</td>
<td>0.84</td>
<td>18.0</td>
</tr>
<tr>
<td>Msh_ATFAA09ds + Msh2-Msh6</td>
<td>0.20</td>
<td>2.81</td>
<td>0.80</td>
<td>25.5</td>
</tr>
</tbody>
</table>

$^1$Fitting of anisotropy decay curves was done with the following expression: $r(t) = r_0(\beta_1 e^{-\frac{t}{\theta_L}} + \beta_2 e^{-\frac{t}{\theta_R}})$. $^2$For all parameters errors are given in parentheses.
Figure 4-9 6-MI TR-anisotropy

(A) Time-dependent anisotropy decay curves of 6-MI (ex=375nm, em=460nm) demonstrate the limited local motion of the 6-MI in the dsDNA sequences containing the pentamer sequences (green, blue, magenta). The time-dependent anisotropy of ssDNA (light-red, dark green, teal and pink) is a bi-exponential decay composed of the local motion of 6-MI (~200ps) and the overall tumbling of DNA. Duplex formation (red, green, blue, magenta) leads to slower tumbling of DNA. (B) A comparison of the quenched and fluorescently enhanced time-resolved fluorescence anisotropy decay curves of the 6-MI-containing ssDNA and dsDNA molecules. The local motion of 6-MI in the duplex ATFAA15ds (green) is reduced compared to the CAFTC07ds duplex (red). In general, duplex formation (green, red) leads to slower overall motion of the DNA. The time-resolved anisotropy of ssDNA (dark green, brown) exhibits a bi-exponential decay consisting of significant local motion of 6-MI (~200ps) and overall tumbling of the DNA (~4 ns) (Table 4-3). Analysis of rotational data was performed as outlined in Chapter 3 section 7.
The average tumbling anisotropic decay of the ss DNA is 4.1 ns and this relatively short correlation time suggests that the ss DNA does not behave as a linear rod (Table 4-3). The $\theta_R$ rotational correlation time of the ds DNA is approximately four times longer than that observed for ss DNA and reflects the increased rigidity of the DNA molecule upon duplex formation (Figure 4-9, Table 4-3). The decreased internal motion of 6-MI is consistent with a model in which duplex formation leads to restrained motion of the probe that causes reduced solvent accessibility and decreased collisional quenching of the excited state with adjacent bases.

4.2.5 Use of 6-MI-enhanced fluorescence to examine protein-DNA interactions

Given the high fluorescence quantum yields exhibited by these sequences (ATFAA15ds and AAFTA10ds) we investigated the utility of these sequences for studying protein-DNA interactions in fluorescence binding experiments. We also specifically incorporated the ATFAA sequence into four different DNA sequences (ATFAA07ds, JX, Msh_ATFAA09ds, TBP_ATFAA23ds) (Table 4-1) and in each case upon duplex formation enhanced fluorescence was observed relative to the ss DNA ($\Phi_{rel} = 0.15 \pm 0.03$) (Table 4-2, Figure 4-3). Additionally, the JX sequence was used to form a four-stranded DNA four way junction (4WJ) (Table 4-1), which allowed us to assess the utility of the ATFAA sequence in unconventional DNA structures. Even within the 4WJ construct 6-MI exhibited enhanced fluorescence similar to that observed in duplex sequences ($\Phi_{rel} = 0.95$) (Figure 4-3). In general, the ATFAA sequence, when incorporated into different DNA constructs, retained the enhanced fluorescence ($\Phi_{rel} \geq 0.7$) (Figure 4-3), making it ideal for investigating structural and dynamic perturbations of DNA in protein-DNA complexes and other structural contexts.
Table 4-4 Equilibrium Protein-DNA Dissociation constants (Kd) measured with 6-MI containing duplexes

<table>
<thead>
<tr>
<th>DNA construct</th>
<th>$K_d$ values (nM) $^1$</th>
<th>$K_d$ values (nM) $^2$</th>
<th>$K_d$ values (nM) $^3$</th>
<th>$K_d$ values (nM) $^4$</th>
<th>Literature $K_d$ Values (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATFAA15ds</td>
<td>1.2 ± 0.8</td>
<td>3.7 ± 0.7</td>
<td>2.4 ± 0.8</td>
<td>1.2 ± 0.4</td>
<td>9 ±1$^5$</td>
</tr>
<tr>
<td>ATFAA07ds</td>
<td>1.5 ± 1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAFTA10ds</td>
<td>1.1 ± 0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSH ATFAA15ds</td>
<td></td>
<td>10.2 ± 3.5</td>
<td></td>
<td>28.0 ±1.6$^6$</td>
<td></td>
</tr>
<tr>
<td>MSH ATFAA09ds</td>
<td></td>
<td>11.0 ± 1.9</td>
<td>13 ±3</td>
<td>28.0 ±1.6$^6$</td>
<td></td>
</tr>
<tr>
<td>TBP ATFAA23ds</td>
<td></td>
<td>10.6 ± 3.1</td>
<td></td>
<td>5 ±2$^7$</td>
<td></td>
</tr>
<tr>
<td>HU ATFAA15ds</td>
<td></td>
<td>4.6 ± 2.2</td>
<td>1.5 ±0.3</td>
<td>16$^8$</td>
<td></td>
</tr>
</tbody>
</table>

$^1$Measured with constant [DNA] = 2000 pM using fluorescence anisotropy.  $^2$Measured with constant [DNA]=1000 pM using fluorescence anisotropy.  $^3$Measured with constant [DNA]=500 pM using fluorescence intensity.  $^4$Measured with constant [DNA]=50 pM using fluorescence intensity.  $^5$From Vitko et al.$^{45}$.  $^6$From Zhai and Hingorani.$^{43}$  $^7$From Perez-Howard et al.$^{41}$.  $^8$Apparent Kd measured by GMSA.$^{37}$
The equilibrium fluorescence anisotropy binding curves show the versatility and sensitivity of the ATFAA sequence, where only 1 nM DNA is needed to accurately measure equilibrium binding affinity of four different protein-DNA complexes. Proteins investigated include sequence specific DNA binding proteins (IHF, TBP) and non-sequence specific DNA binding proteins (Msh2-Msh6, HU). The 6-MI was excited at 340 nm and emission was detected at 440 nm. Representative error bars are shown on the TBP-DNA binding curve; the other curves have similar error. (B) The anisotropy binding curves for IHF binding to DNA with 6-MI labeled in different positions demonstrates binding can be observed independent of probe position in DNA. The relative change in anisotropy is similar for all three constructs. The maximal anisotropy is observed when 6-MI is located in the center of the DNA (ATFAA15ds). Analysis of protein binding was performed as outlined in Chapter 3 section 6.
We specifically exploited the enhanced fluorescence to examine four different high affinity protein-DNA binding interactions at DNA concentrations of 1 nM and less (Figures 4-10). DNA binding interactions were studied with two non-sequence-specific DNA binding proteins (HU and MutS Homolog (Msh) 2-6) and two sequence-specific DNA-binding proteins (Integration Host Factor, IHF and TATA-box Binding Protein, TBP). These proteins represent good model systems for our measurements as their binding interactions have been well characterized by several methods. For all the protein-DNA interactions examined, saturable binding curves are observed when measured by fluorescence anisotropy (Figure 4-12). As the 6-MI is located within the DNA duplex, the starting anisotropy is in the range of 0.16-0.18 reflecting the anisotropy of the free duplex. Addition of protein leads to anisotropy increases of 0.06-0.10. These changes are well outside the error range of our measurements and the increase in anisotropy observed is consistent with protein binding. The magnitudes of the anisotropy changes are not uniform as the masses of the proteins are different and in some cases a profound change in shape accompanies protein binding, such as with the IHF protein.

The incorporated 6-MI minimally perturbs DNA structure and protein-DNA interactions even when a portion of the pentamer sequence is contained within the consensus region as in the IHF-DNA complex formed with ATFAA15ds (Table 4-1). The binding affinity measured was comparable to previously reported values (Table 4-4). Similarly, minimal perturbations were observed when the ATFAA15ds sequence was used to investigate the HU-DNA interaction, as shown by the $K_d$ value (Table 4-4). The HU_ATFAA15ds construct consisted of a single-stranded and duplex portion where the pentamer sequence is located at the junction of the duplex and single strand - the point at which HU putatively binds with
The signal to noise calculated using protein binding experimental conditions reveal the minimum concentration of DNA necessary for accurate measurements. The minimum concentration was defined as the point where the signal to noise deviates from linearity. The average signal to noise with polarizers at 1 nM = 500 and without polarizers at 0.05 nM = 900 for pentamer sequences (ATFAA07, ATFAA15). While the signal to noise is decreased for CAFTC07, with polarizers at 1nM = 256 and without polarizers at 0.05 nM = 598. The signal to noise was calculated as $(I_{\text{signal}} - I_{\text{background}})/(I_{\text{background}})^{0.5}$.

The anisotropy binding curves for IHF binding to 1 nM quenched DNA CAFTC07ds(■) or 1nM unquenched DNA ATFAA15ds(◊) demonstrate the benefit of the enhancement in fluorescence due to pentamer ATFAA. Inaccurate anisotropy values are the result of a low S/N ratio for AFT07ds. (D) Fluorescence intensity binding curves using 50pM AFT07ds (○) and 1nM CAFTC07ds(■) reveal the necessity for higher DNA concentrations due to a poor signal to noise ratio. The Intensity change is ~1.3X for 1nM CAFTC07ds concentration and a dissociation constant (Kd= 1.9 ±0.6 nM) was determined.

**Figure 4-11 6-MI S/N vs. Sequence Context**

(A) The signal to noise calculated using protein binding experimental conditions reveal the minimum concentration of DNA necessary for accurate measurements. The minimum concentration was defined as the point where the signal to noise deviates from linearity. (B) The average signal to noise with polarizers at 1 nM = 500 and without polarizers at 0.05 nM = 900 for pentamer sequences (ATFAA07, ATFAA15). While the signal to noise is decreased for CAFTC07, with polarizers at 1nM = 256 and without polarizers at 0.05 nM = 598. The signal to noise was calculated as $(I_{\text{signal}} - I_{\text{background}})/(I_{\text{background}})^{0.5}$.

(C) The anisotropy binding curves for IHF binding to 1 nM quenched DNA CAFTC07ds(■) or 1nM unquenched DNA ATFAA15ds(◊) demonstrate the benefit of the enhancement in fluorescence due to pentamer ATFAA. Inaccurate anisotropy values are the result of a low S/N ratio for AFT07ds. (D) Fluorescence intensity binding curves using 50pM AFT07ds (○) and 1nM CAFTC07ds(■) reveal the necessity for higher DNA concentrations due to a poor signal to noise ratio. The Intensity change is ~1.3X for 1nM CAFTC07ds concentration and a dissociation constant (Kd= 1.9 ±0.6 nM) was determined.
nanomolar affinity.\textsuperscript{103} Even in this altered structural context, enhanced fluorescence is observed and protein binding is measured with high sensitivity (Table 4-4).\textsuperscript{268}

For all the protein-DNA interactions investigated, the presence of 6-MI minimally perturbed the investigated binding interactions as shown by the measured $K_d$ values (Table 4-4), which are either in good agreement with or below previously reported values.\textsuperscript{103,138,222,267}

Small discrepancies with reported values are not surprising given that some of the reported values were measured using DNA concentrations higher than the $K_d$ value, which can lead to an elevation of the measured $K_d$ value through non-specific binding. For example, Msh2-Msh6 binds to both pentamer-containing DNA substrates with a slightly higher affinity (Figure 4-10, Table 4-4) relative to the previously reported value (10 nM vs. 28 nM).\textsuperscript{270} This difference in $K_d$ values is attributed to the 25-fold lower concentration used in this study (1 nM vs. 25 nM). The lowest concentrations used in this study are based on analysis of signal to noise ratios (S/N) for 6-MI fluorescence in dsDNA. The point at which the signal does not track linearly with concentration is estimated to be the limiting usable concentration for experiments (Figure 4-11). We compare these measurements with the CAFTC07ds because of the similarity in overall sequence; although we note that for CAFTC07ds $\Phi_{rel} = 0.24$, which is higher than 6-MI in other sequence contexts\textsuperscript{6,64,69} (Figure 4-3). From our S/N analysis, we find that in the presence of polarizers the concentration after which the signal is not linear with concentration is approximately 1nM for ATFAA15ds and ATFAA07ds, and is 5nM for CAFTC07ds (Figure 4-11). A comparison of 1 nM ATFAA15ds with 1 nM CAFTC07ds in IHF binding experiments demonstrates the benefit of the higher S/N, as binding is not detected with the CAFTC07ds (Figure 4-11). Thus, we estimate that a minimum concentration of 1 nM is needed for fluorescence anisotropy measurements when using duplexes containing the ATFAA sequence.
Fluorescence intensity binding curves using 500pM DNA reveal the enhanced fluorescence from pentamer ATFAA improves signal to noise at lower DNA concentrations. The lower DNA concentration minimally affects the observed binding curves and consistent dissociation constants ($K_d = 2.0 \pm 0.5$ nM) are determined. Equilibrium binding curves measured by fluorescence intensity with decreasing concentrations of 6-MI-containing DNA (exc = 340; em = 440 nm). DNA concentrations as low as 50 pM can accurately measure protein binding affinity using fluorescence intensity. Even at the lowest DNA concentration, binding is well determined as shown by the consistent dissociation constants obtained, ($K_d = 2.0 \pm 0.5$ nM) (Table 4-4). Analysis of binding curves outlined Chapter 3.6

Figure 4-12 Equilibrium binding curves using picoMolar [6-MI]
Greater sensitivity in binding experiments was obtained by monitoring fluorescence intensity instead of anisotropy, which allows for the use of sub-nanomolar DNA concentrations (Figure 4-12, Table 4-4). Intensity data obtained using 500 pM of DNA substrates ATFAA15ds, HU_TFA15ds, and MSH_TFA09ds yielded $K_d$ values that are consistent with the anisotropy data (Table 4-4, Figure 4-12). To fully explore the sensitivity of the pentamer sequence, we further decreased the TFA15ds concentration to 50 pM. At this DNA concentration, we detected IHF binding and measured a $K_d$ value of 1.2 ± 0.3 nM, consistent with measurements performed at 500 pM and 1.0 nM DNA (Figure 4-12, Table 4-4). We employed the same S/N analysis as described above to estimate the lowest usable concentration for the ATFAA-containing sequences (Figure 4-11). Of note, IHF binding to the quenched CAFTC07ds induced a 130% increase in fluorescence intensity when a 1 nM concentration was used, which is larger than the 30-40% change detected for the ATFAA- or AAFTA-containing sequences (Figure 4-11). However, IHF binding experiments performed with 50 pM AFT07ds do not detect binding because of the relatively low signal (Figure 4-11). This finding illustrates that even though the relative intensity change may be smaller for the pentamer-containing sequences the overall sensitivity is potentially better because of the higher signal of the initial point.

Both 2-aminopurine (2AP) and 6-MI (in other sequence contexts) lack sufficient fluorescent intensity to examine DNA distortions or to measure binding affinities at sub-nanomolar concentrations. The concentration of 2-AP used in other studies\textsuperscript{10,272-276} is 25 nM, which suggests that this is potentially the limiting concentration for 2AP. This concentration is 500-fold higher than the sensitivity limit of the 6-MI pentamer sequence determined in this study. Similarly, 6-MI fluorescence in other sequence contexts is very quenched and in previous studies a concentration of 10 nM was used,\textsuperscript{62} which is 200-fold greater than the
lowest concentration of ATFAA used in this work. The 10 nM concentration may not be the
limiting concentration for all 6-MI-containing duplexes given the strong dependence of 6-MI
quantum yield on sequence. Nevertheless, the advantages of the 6-MI enhanced fluorescence
are striking as it enables the examination of protein-DNA interactions using low nanomolar to
sub-nanomolar concentrations of DNA (Figures 4-10). Also, the sensitivity afforded by the
enhanced 6-MI fluorescence leads to more accurate determinations of high affinity binding
interactions at low DNA concentrations.

4.2.6 Influence of ATFAA Sequence Location on Binding Measurements

The enhanced fluorescence is detected regardless of the location of the pentamer sequence
in the DNA duplex and thus can be used effectively to investigate protein-DNA interactions.
In the case of TATA-Box binding protein (TBP), the ATFAA sequence was located 3’ to the
TATA recognition sequence (Table 4-1). The titration of TBP yielded an increase in
anisotropy with a $K_d$ consistent with previous measurements demonstrating that the ATFAA
sequence reports on TBP binding without interfering with it (Figure 4-10, Table 4-4). The
effect of pentamer location was further investigated with the Msh2-Msh6 and IHF binding
interactions. As shown in figure 4-13, similar $K_d$ values were obtained with different
sequences, illustrating that the location of the 6-MI does not influence the measurement.
Specifically, the ATFAA sequence was located either directly next to the protein-binding site
(Msh_TFA15ds) or 10 base pairs away from the +T insertion loop on the 5’ side
(Msh_TFA09ds) (Figure 4-14). In the IHF duplex sequence the ATFAA sequence was placed
in two distinct positions: adjacent to the binding site on the 5’ side (ATFAA15ds), and 8
base pairs away on the 5’ side (ATFAA07ds) (Figure 4-1). As noted above, the AAFTA
sequence, which is also highly fluorescent, is located on the complementary strand in the
Figure 4-13 Protein-induced intensity change vs. 6-MI sequence position

(A) The increase in 6-MI fluorescence intensity upon protein binding is a sensitive indicator of structural distortions near the protein-binding site. In the Msh2-Msh6-DNA binding interaction (top), when the probe is located adjacent to the +T insertion loop the intensity change was larger (Msh_TFA15ds, 45%) than when located 10 bp away (Msh_TFA09, 20%). (B) In the IHF-DNA binding interaction (bottom), the intensity increase is greatest when 6-MI is located within the IHF consensus sequence (cAFT10ds, 43%), and is decreased when located adjacent to the consensus sequence (TFA15ds, 35%) or 8bp away (TFA07ds, 26%). The location of 6-MI in the IHF or Msh2-Msh6 duplexes had little to no effect on the $K_d$ values (Table 4-4). Analysis of binding outlined Chapter 3.6.
region of the consensus sequence (Figure 4-1, Table 4-1). For all three duplexes, similar $K_d$ values were obtained despite the difference in position on the duplex (Table 4; Figure 9), demonstrating that the pentamer sequence does not need to be located at the protein-binding site to accurately measure $K_d$ values confirming the utility and versatility of the pentamer sequence for studying high-affinity protein-DNA interactions (Figure 4-12).

4.2.7 Detection of DNA Structural Perturbations with 6-MI in Protein-DNA interactions

A distinct advantage of the pentamer sequence relative to external probes is the ability to examine more than just protein binding affinities; 6-MI fluorescence intensity also monitors any perturbations or distortions to the DNA structure. Any protein-induced structural changes in base stacking or base pairing results in an increase in 6-MI fluorescence intensity.\textsuperscript{14} Specifically, we observe that 6-MI reports on structural perturbations according to location within the duplex and the largest intensity changes are associated with the regions of greatest perturbation. In this study all four of the proteins used distort the DNA to varying extents upon binding as determined by x-ray crystallography and all four protein-DNA complexes exhibit an increase in fluorescence intensity upon protein binding (Figure 4-12).\textsuperscript{164,165,277,278}

We explored 6-MI fluorescence in detail with the Msh2-Msh6 and IHF proteins. In the case of Msh2-Msh6 binding to the two DNA constructs, the smallest intensity increase was observed for Msh2-Msh6 binding 10 bp away from the 6-MI (Msh_ATFAA09ds) (25%), while an increase of 45% is observed if 6-MI is located at the +T insertion loop (Msh_ATFAA15ds) (Figure 4-13). The crystal structure of the Msh2-Msh6-DNA complex reveals that perturbation of DNA structure occurs largely at the mismatch binding site, with a reported bend angle of 42° (Figure 4-14).\textsuperscript{279} The smaller intensity change observed for the Msh_ATFAA09ds complex suggests that the distortion induced by the protein is
The X-ray co-crystal structure of Msh2-Msh6 bound to DNA containing a single +T-loop (yellow) with the positions of 6-MI highlighted. The location of 6-MI in the Msh_ATFAA15 duplex is shown in blue and in the Msh_ATFAA09 duplex is shown in red. The Phe residue (green) inserts at the +T loop, distorts the DNA backbone and is proximal to the 6-MI in the Msh_TFA15 duplex. Figure was generated with pymol using the coordinates from pdbid, 2O8F.

Figure 4-14 Msh2-Msh6 crystal structure
substantively weakened 10 bp away from the binding site (Figure 4-14); however, a shorter DNA construct was used in the crystal structure so a direct correlation with bend angle cannot be made for this site.

Similarly, IHF, which has been shown to bend DNA by >160º, induces an increase in 6-MI intensity that is dependent on the location of 6-MI relative to the binding site (Figure 4-13). The greatest intensity increase (AAFTA10ds; 43%) occurs when 6-MI is located within the consensus sequence on the complementary strand. The intensity increase is reduced when 6-MI is located adjacent to the 5' side of the consensus sequence (ATFAA15ds; 35%) and an even smaller intensity change is observed, when 6-MI is located 8 bases away from the consensus sequence (ATFAA07ds; 26%) (Figure 4-13). As shown by the IHF-DNA crystal structure, protein binding induces large changes to the DNA structure and the 6-MI intensity changes reveal these are greatest in the consensus sequence (Figure 4-1). These observations demonstrate that 6-MI is a sensitive reporter of local conformation and protein-induced distortion of the DNA helix.

These effects are further explored using time-resolved fluorescence lifetime and anisotropy measurements of 6-MI in the protein-DNA complexes, which provide greater detail of local DNA structural changes that occur upon protein binding. For all of the complexes studied, we observe that the mean lifetime decreases upon protein binding and the magnitude of the change is dependent on location of 6-MI relative to protein binding site. Not surprisingly, the largest changes are associated with those sequences where the probe is located directly in the binding region. For the IHF and Msh2-Msh6 protein binding leads to an increase in dynamic quenching, as reflected by a redistribution of the subpopulations to shorter decay components
Dynamic quenching upon IHF binding is the greatest for AAFTA10ds, where a shortening of the mean lifetime by 700 ps is observed (Table 4-2; Figure 4-15). The larger

![Graph A](image1.png)
![Graph B](image2.png)
![Graph C](image3.png)
![Graph D](image4.png)

**Figure 4-15 Protein-induced change to lifetime vs. 6-MI sequence position**

(A) Time-resolved fluorescence spectroscopy (ex=375nm, em=460nm) indicates with IHF binding there is no dynamic quenching of 6-MI excited state for ATFAA07ds, which is 8bp away from the consensus sequence (dark blue). (B) There is minimal quenching as 6-MI is moved adjacent to the consensus sequence, ATFAA15ds (green, dark green) (C) Dynamic quenching is greatest when 6-MI is located in on the complementary strand in the consensus sequence (AAFTA10) (pink, purple). (D) Time-resolved fluorescence decay curves reveal dynamic quenching of Msh_ATFAA15ds (orange) with Msh2-Msh6 binding (brown), where the location of 6-MI is adjacent to the T-bulge on the 5’ side. In contrast, no detectable dynamic quenching occurs when the location of 6-MI is 10 basepairs away on the 5’ side of T-bulge, Msh_ATFAA09ds (dark purple with protein, pink without protein). Analysis of anisotropic decay given in Chapter 3.7
subpopulation associated with the shorter decay component is attributed to increased collisional quenching with adjacent bases as a consequence of increased flexibility of 6-MI upon IHF binding. As the distance between 6-MI and the protein binding site increases (Figure 4-1) the amount of dynamic quenching decreases, where IHF binding to ATFAA15ds leads to a 200 ps decrease in mean lifetime, while only a 90 ps decrease in mean lifetime is detected for the ATFAA07ds duplex (Table 4-2; Figure 4-15). The position specific dynamic quenching in the IHF-DNA complex demonstrates the high sensitivity of 6-MI and the ATFAA sequence to local structural perturbations.

In the case of the Msh2-Msh6-DNA complex, when 6-MI is located adjacent to the +T loop - the putative site of protein binding, the greatest change in excited state dynamics is detected. Specifically, the mean lifetime of Msh_ATFAA15ds decreases by ~500 ps upon protein binding, which contrasts with the Msh_ATFAA09ds duplex where no appreciable dynamic quenching is observed and the probe is 10 bp away from the protein binding site (Table 4-2; Figure 4-14).

For both protein-DNA complexes, the increase in dynamic quenching and corresponding decrease in mean lifetime observed upon protein binding would suggest that the fluorescence intensity decreases. As noted above, steady state fluorescence measurements show an increase in intensity upon binding (Figure 4-13). These observations are strongly suggestive that the increase in fluorescence intensity detected in the steady state measurements results from a decrease in static quenching of the 6-MI, as previously suggested by Knutson and co-workers. One source of the decrease could be a reduction in the amount of hypochromicity as base stacking interactions are disrupted as a consequence of the protein-induced distortions.
Time-resolved anisotropy measurements are used to further understand the effect of protein binding on the local motion of 6-MI, as reflected in the rotational correlation time. As expected protein binding (either IHF or Msh2-Msh6) increases the correlation time associated with DNA tumbling ($\theta_R$) by 7-10 ns reflecting the larger size and mass of the protein-DNA complex (Figures 4-16). This increase in correlation time is readily detected in the time-dependent anisotropy decay curves as a shallower slope and is a good indicator of protein binding.

The shorter correlation time is attributed to local motion of 6-MI, and similar to the lifetime measurements, the largest changes in local motion are detected when the probe is located in the consensus sequence (Figure 4-1). Binding of IHF to the AAFTA10ds leads to increased flexibility of 6-MI as detected by a shorter rotational correlation time of the probe ($\theta_L$) of ~900ps, which is not seen for the other IHF duplexes (Figure 4-16, Table 4-3). Significantly, in the X-ray crystal structure, salt bridge and van der Waals contacts are observed between Arg 42 and Arg 46 and the DNA backbone of the residues adjacent to 6-MI in the AAFTA10 sequence. Additionally, replacement of the adenine next to 6-MI with either 2-AP or diaminopurine reduces binding affinity 10-20 fold,$^{140,164}$ indicating that these protein-backbone interactions in the IHF-DNA complex are key to recognition and binding, and probably give rise to the increased mobility observed.

Interestingly, binding of Msh2-Msh6 to the +T-insertion loop results in a longer correlation time for 6-MI local motion ($\theta_L$) (2.11 ns bound vs. 0.73 ns free) suggesting that protein binding stabilizes the probe when located adjacent to the +T (Figure 4-16, Table 4-3). This decrease in motion is attributed to the stacking of residues Phe 432 and Met 452 with the +T insertion as shown in the x-ray co-crystal structure (Figure 10).$^{280}$ The 6-MI result further
Figure 4-16 Protein-induced change to $r(t)$ vs. 6-MI sequence position

(A) Time-resolved anisotropy decay curves of DNA labeled with 6-MI at three different positions in the presence and absence of IHF indicate local base motion is sensitive to protein binding. (A-B) For all three duplexes, the overall tumbling time of the DNA ($\theta_R$) is increased with IHF binding as expected (navy blue, dark green, violet). When 6-MI is located in the consensus sequence (B), IHF binding leads to an increase in 6-MI local motion ($\theta_L$) AATFAA10(violet). (C-D) Time-resolved anisotropy decay curves of DNA labeled with 6-MI at two different positions in the presence and absence of Msh2-Msh6 demonstrate 6-MI local motion is sensitive to protein binding. For both duplexes, the overall tumbling time of the DNA ($\theta_R$) is increased with protein binding as expected. (D) Binding to the Msh_ATFAA15 duplex (orange) leads to a longer correlation time $\theta_L$, suggesting protein binding stabilizes the region around the +T insertion loop (olive green). (C) The absence of a similar increase in $\theta_L$ in the Msh_ATFAA09 duplex, suggests this stabilization is not propagated, but remains close to the +T site (violet, purple). Time resolved analysis given in Chapter 3.7
suggests that stabilization of the +T loop, also leads to a stabilization of the surrounding bases. Similar to the IHF duplexes, the effect is spatially dependent as no change in local motion of 6-MI is detected for the Msh_ATFAA09 duplex, where 6-MI is 10 bases away from the +T loop (Figure 4-16, Table 4-3). Thus, local motion of the probe as detected by time-resolved anisotropy is a sensitive indicator of protein-DNA interactions and can provide unique insight into the nature of the interaction.

4.3 Summary

We report on the discovery of two pentamer sequences, ATFAA and AAFTA, which lead to enhanced fluorescence of nucleoside analog 6-MI upon formation of duplex DNA. We determined the enhanced sequence is solvent inaccessible in a constrained conformation supporting Model A. The ATFAA sequence was incorporated into a number of different DNA duplexes and retained the enhanced fluorescence. This observation of enhanced fluorescence regardless of DNA structure suggests that the ATFAA sequence construct is a good tool for any experiment where high sensitivity and minimal substrate perturbation is desired and we demonstrate that concentrations as low as 50 pM can be used effectively. These sequences were used to measure binding interactions with four distinct DNA-binding proteins. We show that the ATFAA sequence works equally well with sequence-specific and non-sequence-specific DNA-binding proteins. Furthermore, time-resolved and steady-state fluorescence measurements demonstrate that 6-MI fluorescence is very sensitive to local distortion and can report on different degrees of distortion at unique sites within one protein-DNA complex. Given the high sensitivity, versatility and overall utility in measuring binding, we anticipate that these sequence constructs, ATFAA and AAFTA, will be broadly used for investigating protein-DNA interactions.
5 Photophysical Characterization of the sequence specific 6-MI enhanced fluorescence in duplex DNA

5.1 Introduction

Fluorescence spectroscopy is an established technique for studying conformational dynamics of biological systems.\(^1\) Until recently however, the utility of fluorescence spectroscopy for examining DNA and RNA macromolecules has been hindered by technical challenges. Several developments have facilitated the application of fluorescence spectroscopy to DNA/RNA macromolecules. First, advances in the synthesis of oligonucleotides, discovery of new fluorescent probes, and their ease of attachment to DNA and RNA macromolecules has increased the accessibility of fluorescent probes for studying DNA/RNA dynamics.\(^2\) Second, the realization that DNA and RNA molecules are not merely static genetic databases, but that their structure and dynamics play a critical role in many cellular processes including replication of DNA, transcription, translation, DNA compaction, and the regulation of enzymatic activity has caused the field to expand rapidly. For example many DNA binding proteins bend DNA either to prevent or to recruit binding of other proteins.\(^{261}\) New technical developments in fluorescence spectroscopy now allow for the examination of many unanswered questions related to these interactions, such as how specific nucleic acid structures and conformational dynamics influence protein recognition and binding. Fluorescence spectroscopy can provide evidence about the structure and conformational dynamics of DNA due to the sensitivity of the probes and the wealth of information available from their measured properties. However significant ground work needs to be done on specific probes to correlate experimental data with the local probe environment, with the goal of obtaining information on DNA structure and dynamics.
In general, fluorescent probes that are covalently attached to DNA or RNA macromolecules can be divided into two categories: external probes, attached by a carbon linker and internal probes, designed as fluorescent base analogs.\textsuperscript{2,3} External probes are excellent reporters on the global dynamics of an oligomer due to their high quantum yields, while internal probes, also known as nucleoside analogs, report on local DNA structure and dynamics.\textsuperscript{2,3,6} Importantly, the structural similarity between internal probes and native nucleosides allow for measurements of subtle features, such as local base perturbations, since minimal distortion to DNA structure occurs when nucleoside analogs are incorporated.\textsuperscript{5,281} The stability of internal probes and their sensitivity to local environment results in an excellent reporter of protein-induced perturbations to DNA structure and conformational dynamics.\textsuperscript{31,54,60,61,262,272,282,6} Correctly interpreting the biophysical meaning of observed changes in a probe’s fluorescent properties requires a detailed understanding of both the probe in solution and incorporated into a DNA or RNA macromolecule.

The fluorescence observed from nucleoside analogs is directly influenced by the structure and dynamics of its local environment.\textsuperscript{16} Changes in local environment or molecular motion can be reflected in the fluorescent emission spectrum, excited state lifetimes, and anisotropic properties of a fluorophore.\textsuperscript{1} However, correlating these measurements with the underlying physical processes remains a challenge, since many of these measurements report on several competing mechanisms. For example, excited state lifetime and fluorescence intensity data is difficult to interpret because of the competing non-radiative pathways that also de-populate the excited state (\(k_{\text{ic}}\) internal conversion, \(k_{\text{et}}\) energy transfer, \(k_{\text{Q}}\) quenching, \(k_{\text{isc}}\) intersystem crossing, \(k_{\text{p}}\) photochemistry, \(k_{\text{po}}\) phosphorescence).\textsuperscript{17} In the absence of non-radiative decay, the excited state lifetime (\(\tau_{\text{e}}\)) and quantum yield is inversely proportional to the strength of absorption (dipole strength), which means a stronger absorber has a shorter excited state
lifetime and a higher quantum yield. Since non-radiative decay pathways are almost always available to de-populate the excited state in solution, observed radiative lifetimes are rarely as long as the calculated lifetimes and quantum yields are lower than would be expected by the dipole strength alone. Understanding which non-radiative components are competing with the excited state decay and to what extent can be challenging. A considerable amount of effort has been put into the characterization of several nucleoside analogs excited states leading to proposed mechanisms for nucleoside analog fluorescence.

This study focuses on the use of 6-methylisoxanthopterin (6-MI), a fluorescent guanine analog that is able to hydrogen bond to cytosine in duplex DNA, as a reporter of DNA dynamics (Figure 1-2). The excitation energy of 6-MI is lower (ex\text{max} = 340\text{nm} / em\text{max} = 430\text{nm}) than that of natural nucleic acid bases (260nm), such that it is spectrally distinct from the rest of the DNA oligomer. This property makes it an excellent tool for studying DNA dynamics and structure on both global and local distance scales. Another spectral property of 6-MI that facilitates its use as a fluorescent reporter is its high quantum yield compared to other fluorescent nucleoside analogs like 2-aminopurine (2-AP), which mimics adenine. Theoretical calculations of 6-MI monomer electronic states suggest a reversal of excited states (similar to 2-AP) as a cause for high quantum yield, where 6-MI lowest excited state is of $\pi\pi^*$ character in contrast to the $n\pi^*$ character of guanine’s lowest excited state. Like 2-AP in solution, 6-MI in solution exhibits a high quantum yield $\Phi = 0.7$. Incorporation of 6-MI into ssDNA results in quenching of fluorescence, which in most cases is maintained or increased upon duplex formation. The degree of fluorescence quenching is dependent on the identity of the bases adjacent to the 6-MI, and fluorescence quenching is greatest when 6-MI is inserted between purines. Compared to 2-AP, the mechanism of 6-MI fluorescence quenching has been investigated to a lesser extent. One proposed mechanism for quenching within DNA is
interactions between 6-MI and neighboring bases cause a re-ordering of excited state energies, such that the lowest excited state exhibits nπ* character.\textsuperscript{66}

The excited state decay of 6-MI monomer is a single exponential with a lifetime of \( \sim 6.5 \text{ ns} \), but when incorporated into either ss- or dsDNA the decay is multi-exponential.\textsuperscript{65} The multi-exponential fluorescence decay is modeled as three lifetimes components of \( \sim 500 \text{ ps} \), \( \sim 2 \text{ ns} \), and \( \sim 6.5 \text{ ns} \).\textsuperscript{65} Each lifetime component corresponds to a subpopulation of structures with different quenching characteristics that do not interconvert rapidly on the time scale of the excited state decay. Previous studies indicated the subpopulation with a \( \sim 500 \text{ ps} \) lifetime arises from 6-MI stacking with adjacent bases, while the smallest subpopulation (6.5 ns lifetime) corresponds to an extrahelical conformation of 6-MI, and the \( \sim 2 \text{ ns} \) population is an intermediate conformation between the two states.\textsuperscript{65}

Importantly, work probing DNA and RNA dynamics has demonstrated that duplex DNA is minimally distorted upon introduction of 6-MI and that the fluorescent properties of 6-MI are sensitive to local environment.\textsuperscript{64} In addition, 6-MI has been used in a number of studies of protein-nucleic acid interactions including investigations of the cleavage activity of HIV-1 integrase,\textsuperscript{61} probing the binding and destabilization of G tetrad structures by unwinding protein (UP1),\textsuperscript{62} understanding the dynamics of single stranded DNA\textsuperscript{62}, watching the formation of the RecA-DNA filament\textsuperscript{11}, and probing RNA dynamics in the Tetrahymena group I ribozyme.\textsuperscript{63}

In a previous study, we reported on the sequence dependent enhancement of 6-MI fluorescence upon formation of dsDNA( see Chapter 4).\textsuperscript{55} The enhancement of fluorescence was shown to be a consequence of a specific sequence ATFAA (either 5’ or 3’ ordered), where F represents 6-MI. Based on time-resolved fluorescence and KI quenching experiments we hypothesized a single solvent inaccessible conformation of 6-MI was responsible for the enhanced fluorescence. Accordingly, the time dependent anisotropy of 6-MI revealed reduced
flexibility in the double stranded pentamer sequence compared to sequences where 6-MI fluorescence was quenched. Knutson and co-workers identified an additional sequence where duplex DNA formation enhanced 6-MI fluorescence. Knutson et al. proposed that the enhancement of fluorescence was the result of thymines flanking the 6-MI (i.e. TFT), which conformationally constrained the 6-MI. Similar to the ATFAA pentamer sequence, the flanking thymines prevented dynamic quenching with the adjacent bases.

In contrast to earlier work on the photophysical properties of 2-AP and 6-MI where the analog was fluorescently quenched in single-strand and duplex DNA, the pentamers (ATFAA/AAFTA/ATFTA) represent a unique system in which formation of duplex DNA leads to different and interesting photophysical properties. In this study we take advantage of the unique photophysics of this system and examine how small changes in sequence context cause large differences in the fluorescent properties of 6-MI providing insight into the mechanisms for enhancement and quenching of fluorescence. We are interested in exploring the relationship between 6-MI conformational flexibility and dynamic quenching of fluorescence, since in our previous study we hypothesized that 6-MI enhanced fluorescence was the result of surrounding bases constraining its motion. To increase the conformational flexibility of 6-MI we systematically introduced a base-bulge or DNA mismatch opposite the 6-MI and examined the effect on its photophysical properties both in the enhanced pentamer ATFAA/AAFTA and in the quenched sequence CAFTC. Photophysical differences between the two classes of sequences report on the physical basis of enhanced or quenched fluorescence within duplex DNA. We also used the small perturbations to examine another sequence ATFTA which exhibit enhancement of fluorescence upon duplex formation. Additionally, we investigated the role of sequence context on 6-MI fluorescence by measuring the fluorescent properties of several other sequences. Finally, we use molecular dynamics simulations (section
5.2.3) to compare a molecular model of 6-MI within the enhanced versus the quenched sequences. Combining what we learn from simulations with experimental data, we provide a model for how local DNA structure leads to enhanced fluorescence in dsDNA.

5.2 Results

In Chapter 4, we reported a distinct, sequence dependent increase in 6-MI fluorescence of ~4 fold or greater upon duplex formation relative to 6-MI in ssDNA (Table 5-2).\textsuperscript{55} This dramatic enhancement of 6-MI fluorescence in dsDNA was specific to the pentamer sequences, ATFAA or AAFTA. In the case of a similar sequence, CAFTC, which differs only at the n ± 2 positions, significant fluorescence quenching was observed in both ssDNA and dsDNA. This suggests that subtle changes in base sequence can lead to altered fluorescence properties. Typically, incorporation of 6-MI into dsDNA results in a multi-exponential decay where the shorter lifetimes are attributed either directly to the rate of quenching for a particular conformation or to the interconversion rate to an efficiently quenched conformation. Each decay components may correspond to a population with multiple conformations but overlap between these populations is unlikely due to separation between the decay times. The AAFTA/ATFAA pentamer sequences exhibit predominantly a single exponential decay with lifetimes similar to monomer 6-MI (τ=6.5ns), suggesting that interbase quenching interactions are minimized (Table 5-2). In addition, we report herein of a related pentamer sequence ATFTA with a relative quantum yield of ~0.85 and a species
Table 5-1 6-MI containing oligomers

<table>
<thead>
<tr>
<th>Sequence</th>
<th>5'</th>
<th>3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAFTC07</td>
<td>TAT GCA FTC ACT ATG AAT CAA CTA CTT AGA TGGT</td>
<td>GTA TGA ATC AGT ATG GTA TFF ATC TGC TGA AGG AAA</td>
</tr>
<tr>
<td>ATFAA15</td>
<td>TAT GCA GTC ACT ATF AAT CAA CTA CTT AGA TGGT</td>
<td>GTA TGA ATC AGT ATG GTA TFF ATC TGC TGA AGG AAA</td>
</tr>
<tr>
<td>AAFTA10</td>
<td>ACC ATC TAA FTA GTT GAT TCA TAG TGA CTG CATATA</td>
<td>GGA TGG GAA GGG CAT CGT TTG</td>
</tr>
<tr>
<td>ATFTT19</td>
<td>ATG TGA ATC AGT ATG GTA TFF ATC TGC TGA AGG AAA</td>
<td>GTA TGC TGA AGG AAA</td>
</tr>
<tr>
<td>TTFCA15</td>
<td>GGT GCT CAA AAT CTC</td>
<td>GTA TGC TGA AGG AAA</td>
</tr>
<tr>
<td>TGGTG09</td>
<td>GGT GCT CAA AAT CTC</td>
<td>GTA TGC TGA AGG AAA</td>
</tr>
<tr>
<td>AAFCC14</td>
<td>CAG CTG TAC TGG</td>
<td>GTA TGC TGA AGG AAA</td>
</tr>
<tr>
<td>AAFGG10</td>
<td>GGA TGG GAA GGG CAT CGT TTG</td>
<td>GTA TGC TGA AGG AAA</td>
</tr>
<tr>
<td>TFFCA15</td>
<td>GTG GCT CAA AAT CTC</td>
<td>GTA TGC TGA AGG AAA</td>
</tr>
<tr>
<td>TGGTG09</td>
<td>GTG GCT CAA AAT CTC</td>
<td>GTA TGC TGA AGG AAA</td>
</tr>
<tr>
<td>AAFCC14</td>
<td>CAG CTG TAC TGG</td>
<td>GTA TGC TGA AGG AAA</td>
</tr>
<tr>
<td>AAFGG10</td>
<td>GGA TGG GAA GGG CAT CGT TTG</td>
<td>GTA TGC TGA AGG AAA</td>
</tr>
</tbody>
</table>

Table 5-1: First four sequences represent DNA used to characterize fluorescent properties of 6MI in AFTAA sequence followed by sequences used to characterize the AFX trimer where X= A,G,C,T and 6-MI surrounded by purine/pyrimidine. Perturbation to Duplex DNA was achieved by the introduction or removal of base in the complementary strand. (Blue) To disrupt hydrogen bonding to the 6MI a mismatch Adenosine was introduced A(mis)_ATFAA15 and A(mis)_CAFTC07.
Table 5-2 Parameters derived from fitting of time-resolved fluorescence decay curves for bulge/mismatch

<table>
<thead>
<tr>
<th>Sequence</th>
<th>$a_{(1)}$</th>
<th>$\tau_{(1)}$</th>
<th>$a_{(2)}$</th>
<th>$\tau_{(2)}$</th>
<th>$a_{(3)}$</th>
<th>$\tau_{(3)}$</th>
<th>$\tau_{(mol)}$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\Theta_1$</th>
<th>$\Theta_2$</th>
<th>Q.Y. ($\Phi$)</th>
<th>Solvent SS/TR</th>
</tr>
</thead>
<tbody>
<tr>
<td>6MI mono</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6.57</td>
<td>6.57</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.00</td>
<td>1.05</td>
</tr>
<tr>
<td>ATFAA15ss</td>
<td>0.49</td>
<td>0.61</td>
<td>0.31</td>
<td>2.9</td>
<td>0.2</td>
<td>6.78</td>
<td>2.55</td>
<td>0.22</td>
<td>0.32</td>
<td>0.78</td>
<td>5.7</td>
<td>0.20/0.39</td>
<td>0.46</td>
</tr>
<tr>
<td>ATFAA15ds</td>
<td>0</td>
<td>0</td>
<td>0.06</td>
<td>4.01</td>
<td>0.94</td>
<td>7.27</td>
<td>7.07</td>
<td>0.09</td>
<td>1.6</td>
<td>0.91</td>
<td>17.7</td>
<td>0.82/1.08</td>
<td>0.17</td>
</tr>
<tr>
<td>5(+T)_ATFAA15</td>
<td>0.13</td>
<td>1.82</td>
<td>0.52</td>
<td>4.71</td>
<td>0.35</td>
<td>7.72</td>
<td>5.39</td>
<td>0.11</td>
<td>0.65</td>
<td>0.89</td>
<td>16.5</td>
<td>0.69/0.83</td>
<td>0.39</td>
</tr>
<tr>
<td>3(+A)_ATFAA15</td>
<td>0.23</td>
<td>2.32</td>
<td>0.7</td>
<td>4.21</td>
<td>0.07</td>
<td>7.76</td>
<td>4.02</td>
<td>0.18</td>
<td>0.38</td>
<td>0.82</td>
<td>15.9</td>
<td>0.40/0.59</td>
<td>0.45</td>
</tr>
<tr>
<td>A(mis)_ATFAA15</td>
<td>0.33</td>
<td>1.21</td>
<td>0.56</td>
<td>3.76</td>
<td>0.1</td>
<td>7.43</td>
<td>3.25</td>
<td>0.12</td>
<td>0.63</td>
<td>0.88</td>
<td>16.1</td>
<td>0.37/0.49</td>
<td>0.44</td>
</tr>
<tr>
<td>5(-T)_ATFAA15</td>
<td>0</td>
<td>0</td>
<td>0.35</td>
<td>4.28</td>
<td>0.65</td>
<td>7.01</td>
<td>6.05</td>
<td>0.1</td>
<td>1.49</td>
<td>0.9</td>
<td>19.4</td>
<td>0.73/0.92</td>
<td>0.3</td>
</tr>
<tr>
<td>3(-A)_ATFAA15</td>
<td>0.17</td>
<td>2.22</td>
<td>0.22</td>
<td>5.18</td>
<td>0.61</td>
<td>7.86</td>
<td>6.31</td>
<td>0.09</td>
<td>1.57</td>
<td>0.91</td>
<td>18.0</td>
<td>0.72/0.96</td>
<td>0.31</td>
</tr>
<tr>
<td>AAFTA10ss</td>
<td>0.65</td>
<td>0.66</td>
<td>0.26</td>
<td>2.27</td>
<td>0.09</td>
<td>7.38</td>
<td>1.68</td>
<td>0.22</td>
<td>0.42</td>
<td>0.78</td>
<td>5.2</td>
<td>0.15/0.26</td>
<td>0.68</td>
</tr>
<tr>
<td>AAFTA10ds</td>
<td>0</td>
<td>0</td>
<td>0.08</td>
<td>2.87</td>
<td>0.92</td>
<td>8</td>
<td>7.59</td>
<td>0.13</td>
<td>1.92</td>
<td>0.87</td>
<td>17.2</td>
<td>0.73/1.13</td>
<td>0.14</td>
</tr>
<tr>
<td>5(+A)_AAFTA10</td>
<td>0.12</td>
<td>1.4</td>
<td>0.14</td>
<td>4.49</td>
<td>0.74</td>
<td>7</td>
<td>5.98</td>
<td>0.12</td>
<td>1.31</td>
<td>0.88</td>
<td>17.8</td>
<td>0.61/0.93</td>
<td>0.54</td>
</tr>
<tr>
<td>3(+T)_AAFTA10</td>
<td>0.61</td>
<td>1.06</td>
<td>0.3</td>
<td>3.69</td>
<td>0.09</td>
<td>8.1</td>
<td>2.48</td>
<td>0.15</td>
<td>1.05</td>
<td>0.85</td>
<td>15.9</td>
<td>0.36/0.38</td>
<td>0.79</td>
</tr>
<tr>
<td>CAFTA07ss</td>
<td>0.47</td>
<td>0.46</td>
<td>0.26</td>
<td>2.13</td>
<td>0.27</td>
<td>6.66</td>
<td>2.57</td>
<td>0.29</td>
<td>0.50</td>
<td>0.71</td>
<td>5.4</td>
<td>0.19/0.39</td>
<td>0.32</td>
</tr>
<tr>
<td>CAFTC07ds</td>
<td>0.81</td>
<td>1.05</td>
<td>0.16</td>
<td>1.97</td>
<td>0.04</td>
<td>8.04</td>
<td>1.49</td>
<td>0.2</td>
<td>0.61</td>
<td>0.8</td>
<td>8.8</td>
<td>0.24/0.27</td>
<td>0.5</td>
</tr>
<tr>
<td>5(+A)_CAFTC07</td>
<td>0.61</td>
<td>0.46</td>
<td>0.3</td>
<td>1.4</td>
<td>0.09</td>
<td>7.03</td>
<td>1.33</td>
<td>0.76</td>
<td>1.63</td>
<td>0.24</td>
<td>20.6</td>
<td>0.21/0.20</td>
<td>0.9</td>
</tr>
<tr>
<td>3(+T)_CAFTC07</td>
<td>0.89</td>
<td>0.37</td>
<td>0.05</td>
<td>1.9</td>
<td>0.06</td>
<td>7.47</td>
<td>0.87</td>
<td>0.74</td>
<td>1.1</td>
<td>0.26</td>
<td>24.5</td>
<td>0.19/0.14</td>
<td>0.96</td>
</tr>
<tr>
<td>A(mis)_CAFTC07</td>
<td>0.79</td>
<td>0.3</td>
<td>0.08</td>
<td>1.51</td>
<td>0.13</td>
<td>7.4</td>
<td>1.32</td>
<td>0.67</td>
<td>0.78</td>
<td>0.33</td>
<td>24.4</td>
<td>0.18/0.20</td>
<td>0.83</td>
</tr>
<tr>
<td>5(-C)_CAFTC07</td>
<td>0.66</td>
<td>0.99</td>
<td>0.28</td>
<td>1.66</td>
<td>0.06</td>
<td>7.1</td>
<td>1.54</td>
<td>0.23</td>
<td>0.11</td>
<td>0.77</td>
<td>7.1</td>
<td>0.20/0.24</td>
<td>0.59</td>
</tr>
<tr>
<td>ATFTFA19ss</td>
<td>0.23</td>
<td>0.71</td>
<td>0.39</td>
<td>3.25</td>
<td>0.38</td>
<td>7.29</td>
<td>4.21</td>
<td>0.27</td>
<td>0.47</td>
<td>0.73</td>
<td>5.5</td>
<td>0.31/0.64</td>
<td></td>
</tr>
<tr>
<td>ATFTFA19ds</td>
<td>0</td>
<td>0</td>
<td>0.05</td>
<td>0.32</td>
<td>0.95</td>
<td>8.03</td>
<td>7.66</td>
<td>0.15</td>
<td>1.58</td>
<td>0.85</td>
<td>20.1</td>
<td>0.90 / 1.18</td>
<td></td>
</tr>
<tr>
<td>5(+T)_ATFTFA19</td>
<td>0</td>
<td>0</td>
<td>0.07</td>
<td>3.31</td>
<td>0.93</td>
<td>7.41</td>
<td>7.1</td>
<td>0.11</td>
<td>0.84</td>
<td>0.89</td>
<td>20.3</td>
<td>0.72 / 1.08</td>
<td></td>
</tr>
<tr>
<td>3(+T)_ATFTFA19</td>
<td>0.03</td>
<td>0.76</td>
<td>0.53</td>
<td>3.79</td>
<td>0.43</td>
<td>6.9</td>
<td>5.06</td>
<td>0.12</td>
<td>0.74</td>
<td>0.88</td>
<td>18.9</td>
<td>0.42 / 0.77</td>
<td></td>
</tr>
</tbody>
</table>

1Fluorescence decay curves were fit to a sum of exponentials
2For all parameters standard deviations are given in parentheses
3species concentration-weighted lifetime.
Analysis of excited state decay outlined in section 3.7
Table 5-3 Parameters derived from fitting of time-resolved fluorescence decay curves for 6-MI sequences

<table>
<thead>
<tr>
<th>Sequence</th>
<th>α₁ (±0.01)</th>
<th>τ₁ (±0.03)</th>
<th>α₂ (±0.01)</th>
<th>τ₂ (±0.05)</th>
<th>α₃ (±0.02)</th>
<th>τ₃ (±0.04)</th>
<th>τ₃ (mol) (±0.03)</th>
<th>β₁ (±0.02)</th>
<th>Θ₁ (±0.06)</th>
<th>β₂ (±0.02)</th>
<th>θ₂ (±0.9)</th>
<th>Q.Y. (Φ) SS / TR ±0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>6MI mono</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6.57</td>
<td>6.57</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>TTFCA18ss</td>
<td>0.67</td>
<td>0.43</td>
<td>0.2</td>
<td>1.14</td>
<td>0.13</td>
<td>6.34</td>
<td>1.34</td>
<td>0.2</td>
<td>0.81</td>
<td>0.8</td>
<td>5.1</td>
<td>0.11/0.2</td>
</tr>
<tr>
<td>TTFCA18ds</td>
<td>0.83</td>
<td>0.29</td>
<td>0.13</td>
<td>0.99</td>
<td>0.04</td>
<td>6.3</td>
<td>0.62</td>
<td>0.77</td>
<td>1.06</td>
<td>0.23</td>
<td>7.1</td>
<td>0.1/0.09</td>
</tr>
<tr>
<td>GAFAA15ss</td>
<td>0.68</td>
<td>0.34</td>
<td>0.19</td>
<td>1.35</td>
<td>0.13</td>
<td>6.03</td>
<td>1.27</td>
<td>0.31</td>
<td>0.41</td>
<td>0.69</td>
<td>2.2</td>
<td>0.15/0.19</td>
</tr>
<tr>
<td>GAFAA15ds</td>
<td>0.63</td>
<td>0.26</td>
<td>0.32</td>
<td>0.92</td>
<td>0.05</td>
<td>5.29</td>
<td>0.72</td>
<td>0.4</td>
<td>1.45</td>
<td>0.6</td>
<td>10.4</td>
<td>0.14/0.11</td>
</tr>
<tr>
<td>AAFCC14ss</td>
<td>0.22</td>
<td>0.4</td>
<td>0.52</td>
<td>1.74</td>
<td>0.15</td>
<td>6.06</td>
<td>1.9</td>
<td>0.29</td>
<td>0.68</td>
<td>0.71</td>
<td>5.5</td>
<td>0.17/0.29</td>
</tr>
<tr>
<td>AAFCC14ds</td>
<td>0.5</td>
<td>0.37</td>
<td>0.39</td>
<td>1.48</td>
<td>0.11</td>
<td>5.21</td>
<td>1.34</td>
<td>0.15</td>
<td>0.45</td>
<td>0.85</td>
<td>8.0</td>
<td>0.18/0.20</td>
</tr>
<tr>
<td>AAFGG10ss</td>
<td>0.53</td>
<td>0.36</td>
<td>0.22</td>
<td>1.71</td>
<td>0.25</td>
<td>5.82</td>
<td>2.02</td>
<td>0.35</td>
<td>0.31</td>
<td>0.65</td>
<td>2.6</td>
<td>0.14/0.31</td>
</tr>
<tr>
<td>AAFGG10ds</td>
<td>0.49</td>
<td>0.52</td>
<td>0.31</td>
<td>2.97</td>
<td>0.2</td>
<td>6.48</td>
<td>2.47</td>
<td>0.38</td>
<td>0.71</td>
<td>0.62</td>
<td>13.8</td>
<td>0.19/0.38</td>
</tr>
<tr>
<td>TGFTG09ss</td>
<td>0.53</td>
<td>0.45</td>
<td>0.33</td>
<td>1.58</td>
<td>0.15</td>
<td>5.76</td>
<td>1.62</td>
<td>0.15</td>
<td>0.32</td>
<td>0.85</td>
<td>2.5</td>
<td>0.16/0.25</td>
</tr>
<tr>
<td>TGFTG09ds</td>
<td>0.62</td>
<td>0.32</td>
<td>0.19</td>
<td>1.57</td>
<td>0.19</td>
<td>5.75</td>
<td>1.59</td>
<td>0.44</td>
<td>0.58</td>
<td>0.56</td>
<td>6.9</td>
<td>0.06/0.24</td>
</tr>
<tr>
<td>GTFTG32ss</td>
<td>0.21</td>
<td>0.21</td>
<td>0.33</td>
<td>1.52</td>
<td>0.47</td>
<td>4.2</td>
<td>2.49</td>
<td>0.24</td>
<td>0.39</td>
<td>0.76</td>
<td>5.8</td>
<td>0.11/0.37</td>
</tr>
<tr>
<td>GTFTG32ds</td>
<td>0.07</td>
<td>0.39</td>
<td>0.61</td>
<td>2.37</td>
<td>0.31</td>
<td>4.12</td>
<td>2.77</td>
<td>0.19</td>
<td>0.8</td>
<td>0.81</td>
<td>13.9</td>
<td>0.19/0.43</td>
</tr>
<tr>
<td>TAFAG05ss++</td>
<td>0.72</td>
<td>0.21</td>
<td>0.19</td>
<td>1</td>
<td>0.09</td>
<td>4.7</td>
<td>0.77</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.04/0.12</td>
</tr>
<tr>
<td>TAFAG05ds++</td>
<td>0.77</td>
<td>0.34</td>
<td>0.17</td>
<td>0.93</td>
<td>0.04</td>
<td>5.59</td>
<td>0.74</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.04/0.11</td>
</tr>
<tr>
<td>TAFAG08ss+</td>
<td>0.81</td>
<td>0.3</td>
<td>0.14</td>
<td>2.21</td>
<td>0.06</td>
<td>7.28</td>
<td>0.99</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>/0.15</td>
</tr>
<tr>
<td>TAFAG08ds+</td>
<td>0.83</td>
<td>0.21</td>
<td>0.13</td>
<td>1.2</td>
<td>0.03</td>
<td>5.89</td>
<td>0.51</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>/0.08</td>
</tr>
<tr>
<td>CTFCT04ss+</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>3.45</td>
<td>0.75</td>
<td>7.01</td>
<td>6.12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>/0.93</td>
</tr>
<tr>
<td>CTFCT04ds+</td>
<td>0.88</td>
<td>0.26</td>
<td>0.09</td>
<td>1.04</td>
<td>0.03</td>
<td>6.3</td>
<td>0.51</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>/0.08</td>
</tr>
</tbody>
</table>

1Fluorescence decay curves were fit to a sum of exponentials. 2For all parameters standard deviations are given in parentheses. 3species concentration-weighted lifetime. Analysis of excited state decay outlined in section 3.7

1Literature values from sequences64
5'- ACT GCT AFA GAT TTT CCA CAC -3' 5'- ACT FCT AGA GAT TTT CCA CAC -3'
++Literature values from sequences69
5'- ACT AFA GAT CCC TCA GAC CCT TTT AGT CAG TGT GGA -3'
5'- ACT AGA GAT CCC TCA GAC CCT TTT AGT CAG TFT GGA -3'
weighted mean lifetime ($\tau_{mol}$) of ~8 ns (Table 5-2). In this study we have examined 6-MI enhanced fluorescence in greater detail by systematically perturbing duplex DNA to gauge the contribution of surrounding bases on the stability and photophysical properties of 6-MI. These experimental studies were complemented with molecular dynamics simulations to model the sequence-dependent conformation of 6-MI and correlate simulated conformations with observed photophysical properties.

The results are presented in the following order. The first section outlines some of our findings on the relationship between 6-MI fluorescence and sequence contest. Following this sequence is the effects of the three local DNA distortions on 6-MI photophysics, which include: (i) a base-bulge on the complementary strand, (ii) a base-bulge adjacent to the 6-MI, or (iii) an adenine / 6-MI mismatch on 6-MI photophysics (Table 5-1). These local perturbations are expected to probe the effect of base-stacking interactions as well as base pair hydrogen bonding on 6-MI conformation and flexibility. The final section outlines the molecular dynamic analysis of the sequence dependence of 6-MI conformation.

5.2.1 The influence of sequence context on 6-MI photophysics

Previous studies of 6-MI quantum yields have centered on a relatively small number of sequences. (Table 5-3). To date most of the reported work has examined the effect of adjacent or n+1 flanking bases on the QY of 6-MI and 3-MI, a homolog of 6-MI where the following order of increasing QY T ~ C >> G ~ A was observed. We observed a similar ordering in the ssDNA sequences (Table 5-3, Figure 5-2). Since we reported on pentamer sequences with identical bases adjacent to 6-MI, yet profoundly different fluorescent properties where CAFTC dsDNA is quenched and AAFTA dsDNA is enhanced, we elected to further examine the effect of sequence context on 6-MI fluorescence. We specifically
The quantum yield relative to 6-MI monomer indicates quenching of fluorescence for 6-MI in single and double-stranded DNA. The emission max for 6-MI monomer occurs at 431nm, single strands containing 6-MI emission max ~430nm with relative Q.Y. between 0.1 and 0.2. Samples were excited at 340nm, the emission maximum for dsDNA is slightly red shifted to ~435nm The quantum yield is similar between AAFX sequences where X=G,T,C. Thus the n+2 bases appear to play an important role in the enhancement of 6-MI fluorescence.

Figure 5-1 Relative QY for quenched sequences
examined the QY of sequences where two adenines were 5' to 6-MI and the 3' side was occupied by C, T or G (Table 5-1) and observed the following order for the AAFX sequences in which the highest quantum yield is observed for X=T (QY_{rel} = 0.7), while the other bases exhibit relative quantum yields of 0.2 or lower (Table 5-2 and 5-3). This suggests that the 3' position of T may be important for a sequence to exhibit enhanced fluorescence. This is consistent with the finding of Knutson and co-workers who determined that the GTFTG sequence also exhibits enhanced fluorescence. Interestingly, replacing the guanines with adenines in the GTFTG duplex DNA (ATFTA19ds) produces a QY_{rel} of 0.90 ± 0.05 slightly greater than the ATFAA sequence (Figure 5-2, Table 5-1). In fact the ATFTA19 ssDNA exhibits a QY_{rel} of 0.31 ± 0.05, significantly larger compared to other ssDNA sequences (Table 5-2 and 5-3). Thus, while the identity of the n+2 bases adjacent to the 6-MI are important for enhancement of fluorescence, the identity of n+1 base is an important factor in fluorescence quenching.

We further found that replacing the adenine 5' to 6-MI with guanine (GFT09 = TGFTG) but keeping the 3' thymine reduces the QY by 91% (Table 5-3, Figure 5-1) suggesting that the observed fluorescence behavior is unique to adenine and not purines. Nevertheless, it is clear that the effect of sequence context is complex since the pentamer sequence ATFAA also exhibits enhanced fluorescence with thymine 5' to 6-MI and adenine 3' to the probe. Investigations of a greater number of sequences need to be done before any distinct conclusions can be drawn.

The relative quantum yields were also determined from time-resolved measurements. Importantly, these quantum yields only reflect the radiative component of the fluorescence. For the majority of the sequences investigated, the
Fluorescence lifetime decay curves (ex=375, em=460) of 6-MI containing ssDNA and dsDNA molecules compared to monomer (black). Dynamic quenching of 6-MI occurs upon incorporation into ssDNA oligomers, indicated by the significantly faster decays. The degree of quenching is dependent on surrounding bases where thymine adjacent to 6-MI leads to a decrease in quenching, consistent with previously reported data. Comparison of the AAFX where X is T,G,C data set suggest 3’ pyrimidines results in decreased quenching. The instrument response function is shown in gray.

(B) The excited state decay for single and double stranded sequences GTFTG32 and ATFTA19 reveals a reduction in quenching when n+2 is A(light blue, pink). While quenching is maintained in GTFTG32ds, there is a strong reduction in ATFTA19ds. The decay profile for ATFTA19ds is actually longer than 6-MI monomer. One possibility is 6-MI in ATFTA19ds is shielded from quenching with solvent.
Fractional populations determined from time-resolved fluorescence decay curves (shown in Figure 5-2) of 6-MI-containing ssDNA and dsDNA molecules relative to monomer. For ssDNA oligomers, the largest population is associated with the shortest-lived component ($\alpha_1$). The dsDNA exhibits a population distribution similar to the single strands where the largest population is associated with the dynamical quenched species. When adenine is adjacent to 6-MI the size of ($\alpha 1$) decreases slightly. Analysis of excited state decay outlined in section 3.7.

**Figure 5-3 Sequence dependent excited state population**

[Bar chart showing fractional populations for different DNA sequences]
steady-state quantum yields are approximately 30% lower than those determined by time-resolved measurements and are strongly suggestive that a static quenching mechanism contributes to some of the observed fluorescence behavior (Table 5-2 and 5-3).

The fluorescent behavior of these sequences was further examined using time-resolved fluorescence methods. All of the ssDNA molecules have similar excited state decays in which the largest population is dynamically quenched with an average lifetime of 400 ps (Figure 5-2, Table 5-2 and 5-3). Typically, duplex formation of the AAAX sequences alters the fluorescence properties, leading to significant dynamic quenching and lower QY values. This behavior is in distinct contrast to the pentamer sequences AAFTA, ATFAA, and ATFTTA, which have minimal dynamic quenching (Table 5-2, 5-3). In all cases the quenched duplexes either have similar decay profiles to the ssDNAs or the fractional populations of the shortest decay components are increased relative to the ssDNA (Figure 5-3). The TGFTG09ds exhibits a slight increase in $\tau_{mol}$ to 4.25 ± 0.03 ns due to a decrease in the fractional population of the mid-range lifetime, the quantum yield of the duplex is still lower than the single-strand emphasizing that the population of the shortest lifetime component appears to be most significant for determining quantum yield (Table 5-3).

We proposed that the identity of the n+2 base influences the local motion of 6-MI.55 We further investigated this hypothesis by comparing 6-MI rotational correlation time of sequences with a different base at the 5’ n+2 position (Figure 5-4). We found the relationship between 6-MI conformational rigidity and 5’ n+2 suggests that adenine reduces 6-MI motion >> guanine > thymine = cytosine. This is however, an approximation since there are too many variations in the pentamer sequence between constructs to make a definitive statement. A greater data set is required. We do see that when both the 5’ and 3’ n+2 position are adenine
Time-resolved fluorescence anisotropy decay curves of 6-MI incorporated in ssDNA and dsDNA molecules reveal the local motion of 6-MI is dependent on adjacent bases. In general, duplex formation leads to slower overall motion of the DNA. For GTFTG32 and ATFTA19, the time-resolved anisotropy of ssDNA (light red and light blue) exhibits a bi-exponential decay consisting of significant local motion of 6-MI (<0.5 ns) and overall tumbling of the DNA (<5 ns) (Table 5-2, Table 5-3). Upon duplex formation the local motion of 6-MI in the ATFTA19ds (red) is hindered significantly compared to 6-MI in GTFTG32ds (blue) consistent with n+2 adenine altering base dynamics. The rotational correlation times were obtained analyzing the anisotropic decay using equation 3.18 as outlined in Chapter 3. (B) The time-resolved anisotropy of ssDNA exhibits a bi-exponential decay consisting of significant local motion of 6-MI (<1 ns) and overall tumbling of the DNA (<5 ns) (Table 5-2, Table 5-3). Notable, duplex formation of TFC18 (wine) leads to faster anisotropic decay. Analysis of excited state anisotropy outlined in section 3.7.
the local motion of 6-MI is hindered (Table 5-2 and 5-3). For example as shown in figure 4.0, changing the n+2 guanine of GTFTG32 to adenine (ATFTA19) increases the local rotational correlation time ($\theta_L$) from 0.80 ± 0.06 ns to 1.06 ± 0.06 ns, consistent with a conformationally restrained base (Table 5-2, 5-3). This is similar to the reported rotational data where altering cytosine in CAFTC07 to adenine in AAFTA10 increases $\theta_L$ by ~1.4 ns (Table 5-2).\(^5\) In addition, single-stranded GTFTG32ss and ATFTA19ss $\theta_L$ is ~0.45 ns indicating duplex formation is important for reducing 6-MI motion (Table 5-2 and 5-3). Interestingly, two adenines adjacent to 6-MI either 5’ AAFCC/AAFGG or 3’ GAFAA does not significantly increase $\theta_L$ for 6-MI (Table 5-3). Thus, hindering 6-MI motion requires an adenine at the 5’ and 3’ n+2 position. It is important to note, formation of duplexTFC18 reduces 6-MI rotational correlation time from 0.81 ± 0.06 ns to 1.06 ± 0.06 ns (Table 5-3). However, $\beta_1$ increases by 43% from single-strand to double-strand for TFC18 suggesting a larger population of 6-MI exhibit this motion and as shown in Figure 5 the anisotropic decay for TTFCA18ds is rapid compared to other dsDNA. For the majority of the ssDNA sequences 6-MI $\theta_L$ is < 0.5 ± 0.06 ns corresponding to a base capable of freely rotating (Table 5-2 and 5-3).

It is clear from the range of data that there is no simple pattern of quenching which can be explained by nearest neighbor bases. One might expect that purines, with lower lying excited states, would more effectively interact with, and quench, the 6-MI excited state but more complicated effects are apparently at work. The fact that the n +/- 2 bases are so important suggests that it is the local structure or dynamics which are critical to the quantum yield. To explore these issues we look to the effect of increasing the conformational instability of the DNA and the influence on QY. These are introduced via bulges and mismatches placed in various positions relative to the 6-MI.
5.2.2  Local DNA structural deformations effect 6-MI photophysics

We employ the following nomenclature to describe the modified duplexes used for these studies, which specifically references all perturbations to the 6-MI-containing strand only. Thus, single base bulges are an addition or subtraction of a base as follows: i) unpaired base adjacent to 6-MI on the 5' side 5(+X) where X represents the unpaired base and 3(+Y) would be an unpaired base on the 3' side of 6-MI ii) an unpaired base on the complementary strand 5' or 3' to the 6-MI bp is denoted as 5(-X) or 3(-Y), respectively in which the missing base on the 6-MI containing strand is -X or -Y. Thus, using the TFA15 sequence (Table 5-1) a perturbed sequence containing an unpaired thymine base adjacent to 6-MI on the 5' side would be designated 5(+T)_TFA15 (Table 5-1). Sequences containing a 6-MI:adenine mismatch are denoted as A(mis)_sequence (Table 5-1).

The complementary strand base-bulge alters local DNA structure

A number of techniques were applied to examine the impact of a complementary strand base-bulge on the photophysics of 6-MI. Correlating the structural impacts of the specific base-bulge with 6-MI structure and dynamics will help determine which factors are most influential in quenching fluorescence. We first determined the thermal stability of the DNA to confirm the base-bulge on the complementary strand led to an increase in local conformational variability by examining absorption temperature melts at 340 nm, a wavelength at which only 6-MI absorbs. The melting temperature calculated from this data provides a measurement of duplex stability, which reports indirectly on local perturbations to the duplex. As shown in figures 5-5 and 5-6, based upon monitoring 340 nm the base-bulge
Above are plots of the change in absorbance at 260 nm (ΔAbs.) relative to the initial absorbance (Abs.i) vs the temperature in Kelvin. The shift in Tm for constructs containing a perturbation indicate reduced DNA stability. A Comparison of dsDNA melting profile to sequences containing bulge or mismatch displays a shift in Tm to lower temperature suggests decreased stability to global DNA structure. In addition, the sequences exhibit a similar amount of hypochromicity. The samples contained 500nM dsDNA bulge/mismatch containing oligomers in 50mM NaPO₄. Analysis of the temperature melts was performed as outlined in Chapter/Section 3.4.
Above are plots of the change in absorbance ($\Delta\text{Abs.}$) at 340 nm relative to the initial absorbance ($\text{Abs}_i$) versus temperature in Kelvin (K). Absorbance melts for 15uM dsDNA and bulge/mismatch containing oligomers in 50mM NaPO$_4$ monitored at 340nm. A decrease in hypochromicity is observed for strands with introduction of bulge/mismatch. Comparison of dsDNA melting profile to sequences containing bulge or mismatch displays a shift in Tm to lower temperature suggesting decreased stability to global DNA structure. Analysis of the temperature melts was performed as outlined in Chapter/Section 3.4.

Figure 5-6 Temperature melt 340 nm
Table 5-4 Temperature Melts for absorbance at 340nm and 260nm

<table>
<thead>
<tr>
<th></th>
<th>Tm (K)</th>
<th>ΔTm (K)</th>
<th>Tm (K)</th>
<th>ΔTm (K)</th>
<th>Hypochrom.</th>
<th>Hypochrom.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>340nm</td>
<td>260nm</td>
<td>340nm</td>
<td>260nm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFT07ds</td>
<td>334.4±0.9</td>
<td>327.4±0.9</td>
<td>0.34±0.06</td>
<td>0.27±0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5(+A)_CAFTC07ds</td>
<td>331.5±0.9</td>
<td>-2.9±0.9</td>
<td>323.9±0.9</td>
<td>-3.5±0.9</td>
<td>0.22±0.02</td>
<td>0.23±0.03</td>
</tr>
<tr>
<td>3(+T)_CAFTC07ds</td>
<td>330.5±0.0</td>
<td>0.0±0.0</td>
<td>323.5±0.0</td>
<td>3.5±0.0</td>
<td>0.22±0.01</td>
<td>0.2±0.05</td>
</tr>
<tr>
<td>A(mis)_AFT07ds</td>
<td>330.9±0.6</td>
<td>1.0±0.6</td>
<td>322.9±0.6</td>
<td>1.0±0.6</td>
<td>0.14±0.03</td>
<td>0.26±0.04</td>
</tr>
<tr>
<td>5(-C)_CAFTC07ds</td>
<td>331±1.0</td>
<td>0.6±1.0</td>
<td>322.9±1.0</td>
<td>1.0±1.0</td>
<td>0.25±0.01</td>
<td>0.23±0.01</td>
</tr>
<tr>
<td>TFA15ds</td>
<td>333.3±0.3</td>
<td>327.0±0.3</td>
<td>0.12±0.02</td>
<td>0.26±0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5(+T)_ATFAA15ds</td>
<td>331.6±0.3</td>
<td>1.7±0.7</td>
<td>323.1±0.7</td>
<td>4.9±0.9</td>
<td>0.09±0.02</td>
<td>0.23±0.01</td>
</tr>
<tr>
<td>3(+A)_ATFAA15ds</td>
<td>327.3±0.0</td>
<td>6.0±0.0</td>
<td>323.3±0.0</td>
<td>3.7±0.0</td>
<td>0.2±0.03</td>
<td>0.2±0.05</td>
</tr>
<tr>
<td>A(mis)_TFA15ds</td>
<td>330±0.0</td>
<td>3.3±0.0</td>
<td>321.4±0.0</td>
<td>5.6±0.6</td>
<td>0.08±0.04</td>
<td>0.24±0.01</td>
</tr>
<tr>
<td>5(-T)_ATFAA15ds</td>
<td>331.4±0.0</td>
<td>1.9±0.0</td>
<td>322.9±0.0</td>
<td>4.1±0.0</td>
<td>0.07±0.01</td>
<td>0.24±0.03</td>
</tr>
<tr>
<td>3(-A)_ATFAA15ds</td>
<td>330.7±0.0</td>
<td>2.6±0.6</td>
<td>322.7±0.6</td>
<td>4.3±0.3</td>
<td>0.09±0.03</td>
<td>0.21±0.04</td>
</tr>
<tr>
<td>cAFT10ds</td>
<td>332.7±0.0</td>
<td>327.0±0.0</td>
<td>0.31±0.02</td>
<td>0.26±0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5(+A)_AAFTA10ds</td>
<td>331.1±0.0</td>
<td>1.6±0.0</td>
<td>324.4±0.0</td>
<td>2.6±0.0</td>
<td>0.25±0.02</td>
<td>0.25±0.02</td>
</tr>
<tr>
<td>3(+T)_AAFTA10ds</td>
<td>331.1±0.0</td>
<td>1.6±0.0</td>
<td>321.4±0.0</td>
<td>5.6±0.0</td>
<td>0.12±0.03</td>
<td>0.2±0.01</td>
</tr>
</tbody>
</table>

Table (5-4) Observed Tm for 15uM in 50mM NaPO₄ duplex DNA containing 6-MI monitored by absorbance at 340nm and Tm for 1uM DNA monitored at 260nm. The ΔTm represents the difference between altered DNA and duplex DNA. The fraction of hypochromicity for duplex formation is reported for 340 nm and 260 nm. Analysis of the temperature melts was performed as outlined in Chapter/Section 3.4
in the complementary strand destabilizes the duplex by 1.7 K on average (Table 5-4) indicating structural deformation in both the fluorescently enhanced (ATFAA15) and quenched (CAFTC07) sequences. This correlates with increased solvent exposure another indicator of duplex perturbation. The base-bulge on the complementary strand increases solvent exposure by 76% and 18% for ATFAA15 and CAFTC07 respectively (Table 5-2, Figure 5-7). The larger percentage increase in solvent accessibility for ATFAA15 is partly ascribed to the difference in solvent exposure of the unperturbed duplex; 6-MI adopts a solvent-inaccessible conformation in ATFAA15ds, whereas 6-MI in CAFTC07ds has greater solvent accessibility. The solvent exposure indicates a similar degree of structural perturbation from the base-bulge on the complementary strand for both fluorescently enhanced and quenched sequences. Validating our hypothesis, the base-bulge perturbation to DNA is independent of sequence and comparison of changes to 6-MI photophysics in different sequences is an acceptable method to determine the influence of 6-MI conformational flexibility on the quenching or enhancement of fluorescence. An important observation is the Tm calculated from the change in 340 nm absorbance is on average ~7.5 ± 1.5 K higher than the Tm calculated based upon the change in 260nm absorbance (Table 5-4). A possible explanation for the discrepancy is we are monitoring a single base at 340 nm versus the complete construct at 260 nm.

We are interested in how the location of the bulge influences 6-MI quantum yield which can couple with spectra shifts, changes in hypochromicity, excited state lifetime and rotational correlation time. Shown in Figure 5-8 and listed in Table 5-2 are the QY\text{rel} for all the sequences studied including the complementary strand base-bulge. Both the steady state and
The solvent accessibility in ssDNA and dsDNA is measured relative to 6-MI monomer by fluorescence collisional quenching with KI. For duplexes containing either the ATFAA or the AAFTA sequence, the formation of dsDNA leads to a reduction in 6-MI solvent accessibility relative to ssDNA and 6-MI monomer, which suggests 6-MI is not extrahelical in these sequences (ATFAA15ds, AAFTA10ds). Solvent exposure is increased in the quenched duplex, CAFTC07ds. A single-base bulge increases 6-MI solvent exposure notably in CAFTC07ds the sequence with no enhanced fluorescence. The greatest increase in solvent exposure occurs when the single-base bulge is located 3-prime and adjacent to 6-MI. The fraction change in fluorescence versus [KI] concentration was analyzed using a modified Stern-Volmer equation in section 3.4 equation 3.6.

Figure 5-7 Fraction solvent accessible
The Quantum yield relative to monomer for oligomers indicates there is fluorescence quenching when 6-MI is mismatched with adenine for enhanced sequence ATFAA15ds (A), A(mis)_ATFAA15ds (cyan). In addition, the QY decreases when there is a base bulge adjacent on 3-prime side of 6-MI 3(+A)_ATFAA15ds (same behavior for (B) AAFT10 and (C) ATFTA19) (magenta). A blue-shift in the spectrum is observed for oligomers containing a bulge or mismatch. The quantum yield is reduced ~50% with the introduction of a single-base bulge adjacent to the 6-MI on the 3-prime side 3(+A)_ATFAA15ds (Φ=0.4±0.02) and adenine mismatch A(mis)_ATFAA15ds (Φ=0.37±0.03) across from 6-MI relative to ATFAA15ds (Φ=0.82±0.03). (D). There is minimal change in QY with the quenched sequence CAFTC07ds Samples were excited at 340nm in 50mM NaPO₄ pH 7.5.
Time-resolved quantum yields are given as well as detailed fluorescence lifetimes and time-resolved anisotropy measurements. It is clear that the complementary strand base-bulge has minimal impact on the quantum yield, both in the highly fluorescent sequence (ATFAA15ds) and the quenched sequence (CAFTC07). It is interesting to note distorting the quenched sequence 5(+C)_AFT07ds did not lead to a higher quantum yield either, we had expected reducing base stability would lead to an increase in fluorescence. This result provides insight into the mode of fluorescence quenching (Table 5-2, Figure 5-8).

As expected, based on QY$_{rel}$ data, there is only a modest change to the excited state lifetime with the introduction of the complementary base-bulge (Table 5-2, Figure 5-9). The changes are reflected in shifts in the fractional subpopulations (Figure 5-10, Table 5-2). For the complementary base-bulge 5(-T)_ATFAA15, there is a 35% shift in the fractional subpopulation from $\tau_3 =$ 7.16 ns to $\tau_2 =$ 4.2 ns, which suggests a small change in the conformations that 6-MI adopts compared to the conformations in ATFAA15ds (Figure 5-10, Table 5-2). Similarly, composition of the excited state subpopulations for the base-bulge 3(-A)_ATFAA15 is also altered compared to ATFAAads. Three decay components are required to model the lifetime with the appearance of a faster component $\tau_1 =$ 2.21 ns indicating 6-MI can adopt a greater number of conformations in the 3(-A)_ATFAA15 construct (Figure 5-10, Table 5-2). In contrast, the excited state decay for complementary base-bulge in the quenched sequence 5(-C)_AAFTA07 resembles the decay of CAFTC07ds, indicating the greater sensitivity of 6-MI excited state to conformation in the enhanced sequence to perturbations on the complementary strand. The modest influence of the complementary base-bulge on excited state decay is paired with a nominal change in 6-MI conformational flexibility. The time-dependent anisotropy decay of each duplex reports on internal motion.
Fluorescence lifetime decay curves (ex=375, em=460) of 6-MI incorporated into ssDNA and dsDNA molecules compared to monomer (black). The significantly faster decay of single-stranded DNA (red) is consistent with an increase in the non-radiative rate that occurs upon incorporation into DNA. The similarity of ATFA15/AAFTA/ATFTA19 dsDNA (blue) and 6-MI monomer (black) fluorescence decays represents a reduction in this rate upon duplex formation. The 3’ base bulge produces a faster decays compared to the enhanced duplexes ATFA15/AAFTA/ATFTA19 (magenta) indicating this perturbation alters 6-MI microenvironment. Similarly, dynamic quenching occurs when 6-MI mismatched with adenine for the enhanced duplexes ATFA15 (cyan). In addition, the 3’ base-bulge and adenine mismatch alter the excited state decay for the quenched duplex CAFTC07ds. The instrument response function is shown in gray. The excited state decays was modeled using a sum of exponentials given in Section 3.7 equation 3.14.

Figure 5-9 6-MI lifetime base-bulge/mismatch
Fractional populations determined from time-resolved fluorescence decay curves (shown in Figure 5-9) of 6-MI-containing DNA molecules relative to monomer indicate duplex formation leads to a shift in the peak population to the longest lifetime component ($\alpha_3$) for sequences containing ATFAA15/AAFTA10/ATFTA19. For ssDNA oligomers, the largest fractional population is associated with the shortest-lived component ($\alpha_1$). The CAFTC07ds exhibits a population distribution similar to the single strands where the largest population is associated with the shortest-lived component ($\alpha_1$). Also, perturbation in CAFTC07 does not lead to population shift toward $\alpha_3$. Perturbation of DNA structure through single-base bulge or mismatch leads to a shift in the population distribution in ATFAA15/AAFTA10/ATFTA. In all three cases the 3-prime single-base bulge results in largest population shift to $\alpha_1$. In addition adenine/6-MI mismatch (A(mis) ATFAA15ds and A(mis) CAFTC07ds) population is also shifted toward the quenched species. The excited state decays was modeled using a sum of exponentials given in Section 3.7 equation 3.14.

Figure 5-10 Fractional population perturbation
of 6-MI, which is dependent on the sequence context as well as the type and location of perturbation. A two-component exponential decay was necessary to correctly model the time-dependent anisotropy decay for ss and dsDNA molecules containing 6-MI. All sequences exhibit a faster rotational correlation time ($\theta_L$), representing the local motion of 6-MI. The internal motion $\theta_L$ reflects the stability of the probe, arising in part from base stacking interactions.\textsuperscript{34,262} Interestingly, a base-bulge on the complementary strand has little to no effect on the internal motion of 6-MI within the enhanced duplex 5(-T)$_{ATFAA15}$ and 3(-A)$_{ATFAA15}$ or quenched duplex 5(-C)$_{AAFTA07}$ (Figure 5-11 Table 5-2). Thus, the perturbation to nucleic acid dynamics from a base-bulge appears to be highly localized to the adjacent bases. The effect on 6-MI flexibility is similar for the quenched/ enhanced, reaffirming our early hypothesis (Figure 5-11, Table 5-2). Based on temperature-melt and time-dependent anisotropy data, the base-bulge on the complementary strand leads to instability of the DNA structure but has little effect on local dynamics for 6-MI, concurrent with the observed QY and excited state decay.

The hypochromicity of 6-MI in DNA is a useful parameter obtained from 340 nm absorption temperature melts. Hypochromism originates from the interaction between one particular excited electronic state of a given chromophore and different electronic states of the neighboring chromophores.\textsuperscript{18} The hypochromicity at 340 nm for ATFAA15ds (0.12± 0.01) is lower relative to AAFTA10ds (0.26± 0.02) and AAFTA07ds (0.33± 0.06) (Table 5-4, Figure 5-6). The difference in hypochromicity at 340 nm between the highly fluorescent duplexes ATFAA15ds and AAFTA10ds, suggests enhancement in fluorescence corresponds to a decrease in the non-radiative rate of decay, not an increase in the transition rate from
Figure 5-11 6-MI r(t) vs base-bulge A

(A) Time-resolved fluorescence anisotropy decay measurements for ATFAA15 constructs containing base bulge adjacent adjacent to 6-MI (5(+T) = green, 3(+A) = magenta) or adenine/6-MI mismatch (A(mis) = cyan) indicate distortions to the local DNA structure increase local motion of 6-MI. However, a base-bulge on the complementary strand results in minimal change to 6-MI dynamics (5(-T) = wine). The change is not as dramatic as it is for the single-base bulge in the quenched CAFTC07 ds (B) The pattern is similar but magnitude of the change is larger, indicating greater stability of 6-MI in the ATFAA sequence. The time-resolved anisotropic decays were modeled using a sum of exponentials given in Section 3.7 equation 3.18.
Time-resolved fluorescence anisotropy decay measurements for AFTAA10 constructs containing base bulge adjacent adjacent to 6-MI (5(+T) = green, 3(+A) = magenta) indicate a slightly greater change to 6-MI dynamics with the 3’ bulge, consistent with AFTAA15 data. (B) The pattern is similar for ATFTA19 base-bulge. In ATFTA19 the change 6-MI dynamics is not as apparent. The time-resolved anisotropic decays were modeled using a sum of exponentials given in Section 3.7 equation 3.18.
the ground state to the excited state. In addition, while the hypochromicity of 6-MI is ~30% smaller for the base-bulge on the complementary strand compared to ATFAA15ds, the difference in QY$_{rel}$ between complementary base-bulge and ATFAA15ds is only 11% (Table 5-2 and 5-4), which supports the idea that ground-state electronic interactions are not the main factor for the enhanced fluorescence. Another indicator of a reduction in stacking interactions is a shift in the emission spectra, which is associated with reduced stacking interactions of nucleoside analogs with adjacent bases.$^{262}$ The 6-MI emission is blue-shifted for the complementary base-bulge oligomers compared to the fluorescently enhanced and quenched duplexes (Figure 5-8).

We were surprised to observe the insensitivity of 6-MI QY to the base-bulge on the complementary strand. We had expected any change to the DNA conformation would be reflected in 6-MI QY. We know from temperature melt data and 6-MI solvent exposure the base-bulge on the complementary strand leads to conformational instability. The base-bulge has a minimal effect on 6-MI rotational correlation time therefore the excited state decay is similar to duplex DNA. We noted the effect on 6-MI was independent of sequence context as similar changes were observed in the quenched and fluorescently enhanced constructs.

*Base-bulge adjacent to the 6-MI is associated with increased motion and fluorescence quenching of 6-MI*

In contrast to the base-bulge on the complementary strand, there is a substantial change in the quantum yield and rotational correlation time when the base-bulge is adjacent to 6-MI. As
shown in Figure 5-5 and Table 5-4, the base-bulge adjacent to 6-MI destabilizes the duplex by $2.2 \pm 3.4$K on average), consistent with a structural deformation to the DNA construct. The 340 nm Tm data is in agreement with 260 nm results. For example, the 340 nm $\Delta$Tm for 5(±T) ATFAA15 is $\sim1.7 \pm 1.8$ K compared to the 260 nm $\Delta$Tm of $4.9 \pm 1.2$ (Table 5-4). The behavior is similar for the other enhanced sequence AAFTA10 and the quenched sequence CAFTC07 (Table 5-4). The reduction in thermal stability at 340 nm is slightly greater when the base-bulge is 3-prime to 6-MI, most notable in 3(±T) ATFAA15 the $\Delta$Tm is $6K \pm 1.8$ K (Table 5-4, Figure 5-6). The instability of a base-bulge adjacent to 6-MI is associated with increased solvent exposure. The base-bulge 5-prime adjacent to 6-MI increases the solvent exposure by 120% for 5(±T) ATFAA15 compared to ATFAA15ds (Table 5-2, Figure 5-7). Likewise, when the base-bulge is 3-prime adjacent to 6-MI the solvent exposure increases by 160% for 3(±A) ATFAA15 (Table 5-2, Figure 5-7). The trend in increased solvent exposure is similar for adjacent base-bulges in the enhanced sequence AAFTA10 (Figure 5-7, Table 5-2). For the quenched sequence CAFTC07 the increase in solvent exposure from a base-bulge adjacent to the 6-MI is smaller compared to the highly fluorescent sequences (Table 5-2, Figure 5-7). In summary, the greater reduction in 340 nm DNA thermo stability and increase in 6-MI solvent exposure suggests destabilization of 6-MI within the DNA is larger for the adjacent base-bulge compared to one on the complementary strand.

Time resolved anisotropy supports this hypothesis, revealing the internal motion of 6-MI increases when adjacent to a base-bulge compared to the bulge on the complementary strand (Table 5-2, Figure 5-11). The degree of 6-MI local motion is dependent on the location of base-bulge, either 5-prime or 3-prime relative to the 6-MI. For example, when the base-bulge is 5-prime adjacent to 6-MI, such as 5(±T) ATFAA15 and 5(±A) AAFTA10 the rotational
correlation time of 6-MI decreases by $0.95 \pm 0.12$ and $0.61 \pm 0.12$ ns, respectively (Figure 5-11, Table 5-2). Interestingly, for sequences with enhanced fluorescence the perturbation to 6-MI base dynamics is greater when location of the base-bulge is 3-prime to 6-MI. The rotational correlation time decreases by $1.22 \pm 0.12$ and $0.87 \pm 0.12$ ns for $3(+A)_{ATFAA15}$ and $3(+T)_{AAFTA10}$, respectively (Figure 5-11, Table 5-2). We observed similar behavior with the recently discovered fluorescently enhanced sequence ATFTA19ds, such that greater perturbation to 6-MI dynamics occurs when the base-bulge is 3-prime to 6-MI (Figure 5-11, Table 5-2). Thus, the identity of the 3-prime adjacent base does not determine if the base-bulge will alter 6-MI. However, we did notice the local rotational correlation time for $3(+T)_{ATFTA19}$ and $3(+T)_{AAFTA10}$ is 0.74 and 1.05 ns respectively, which is 0.36 to 0.67 ns longer than $3(+A)_{ATFAA15}$ (Figure 5-11, Table 5-2), which suggests the base identity may influence the magnitude of the change. We noticed the anisotropic decay for the quenched sequence CAFTC07ds with adjacent base-bulge is distinct compared to adjacent base-bulge in the enhanced sequences AAFTA10ds, ATFAA15ds, and ATFTAds (Figure 5-11, Table 5-2). As shown in Figure 5-11, there is a large shift in subpopulations on average 50% from $\beta_2$ to $\beta_1$ for the base-bulge in CAFTC07ds (Table 5-11, Table 5-2). The difference suggests 6-MI in the sequences ATFAA, AFTAA, and ATFTA has decreased susceptibility to local perturbations (Table 5-2). The single-base bulge adjacent to 6-MI increases solvent accessibility and its internal motion consistent with a more dynamic system. In our previous work we proposed when there is a correlation between the amount of conformation space 6-MI can sample and the degree of collisional quenching in the excited state.

DNA structural deformations adjacent to 6-MI alter its photophysical properties. For example, a base-bulge adjacent to the 6-MI causes a blue shift in the emission spectra in the
quenched and fluorescently enhanced DNA sequences (Figure 5-8). Accordant with solvent exposure and rotational data the extent to which the base-bulge altered 6-MI photophysical properties is dependent on the location of the bulge. For the fluorescently enhanced sequences with a base-bulge adjacent to 6-MI on the 5-prime side: 5(+T)_ATFAA15, 5(+A)_AAFTA10, and 5(+T)_ATFT1A19 the QY_{rel} is reduced by ~15% compared to dsDNA (Table 5-2, Figure 5-8). Interestingly, the magnitude of the change in QY_{rel} for the thymine 5-prime base-bulge is approximately the same as the adenine 5-prime base-bulge, indicating, the perturbation to 6-MI photophysics is independent of the identity of extra base. As shown in Figure 5-8, for the constructs with a base-bulge 3-prime to the 6-MI: 3(+A)_ATFAA15, 3(+T)_AAFT10, and 3(+T)_ATFT1A19, the reduction in QY_{rel} compared to the fluorescently enhanced duplexes is ~50%, which suggests altering the conformation of the 3-prime base has a greater effect on 6-MI photophysics compared to the 5-prime base-bulge (Figure 5-8, Table 5-2). Similar to the 5-prime base-bulges, the reduction in QY_{rel} is independent of extra base type. We were again surprised to not observe an enhancement in fluorescence with either a 5-prime or 3-prime base-bulge adjacent to 6-MI in the quenched duplex CAFTC07 ds (Figure 5-8, Table 5-2). Another indication that 6-MI has greater sensitivity to DNA deformations in the fluorescently enhanced sequence compared to sequences where 6-MI is quenched.

In addition to the changes in QY_{rel}, the adjacent base bulge also alter the 340 nm and 260 nm hypochromicity, consistent with a reduction in base stacking interactions. For example, the 340 nm hypochromicity of the base-bulge construct 3(+T)_AAFT10 is 0.12 ± 0.03 compared to 0.31 ± 0.02 for AAFT10 ds (Table 5-4), indicating a reduction in 6-MI stacking interactions with adjacent bases. It is worth noting the in the quenched sequence, the hypochromicity of the base-bulge construct 3(+T)_CAFTC07 is 0.15 ± 0.01 compared to 0.34 ± 0.06 for duplex
CAFT07Cds, again suggests a reduction in stacking interactions (Table 5-4). However, this reduction does not result in an increase in QY_{rel}.

The perturbations to the excited state lifetime of 6-MI are dependent on the location of the base-bulge. We noticed for the constructs with a 5-prime base-bulge: 5(+T)_ATFAA15 and 5(+A)_AAFTA10, the reduction in QY_{rel} is associated with the appearance of a subpopulation with a short lifetime τ_{1}~1.6 ns (Figure 5-10, Table 5-2). Thus, a 5-prime base-bulge increases the rate of non-radiative decay. Interestingly, the short lifetime is absent in the 5(+T)_ATFTA19 construct (Figure 5-10, Table 5-2). A possible explanation is the greater electronic overlap of 6-MI with adenine compared to thymine. We also noticed Similarly, there is a shift in the excited state decay to a shorter lifetime for the 5-prime bulge in quenched sequence 5(+A)_CAFTC07 (Figure 5-10, Table 5-2). We reported above the greatest reduction in QY_{rel} occurs with the base-bulge 3-prime to 6-MI (Table 5-2). The reduction in QY_{rel} for the 3-prime base-bulge constructs: 3(+A)_ATFAA15, 3(+T)_AAFTA10, and 3(+T)_ATFTA19 is associated with a large shift in the excited state decay to subpopulations with shorter lifetimes (Table 5-2, Figure 5-10). As shown in Figure 5-9, the excited state decay for 3-prime base-bulge shares a greater resemblance to the decay for single-stranded DNA than duplex DNA. In addition, we observed similar behavior for the quenched 3-prime base-bulge construct 3(+T)_CAFTC07, where the lifetime of τ_{1} decreases by 0.68 ± 0.06 ns, consistent with an increase in the non-radiative rate (Table 5-2, Figure 5-10). Thus, compared to the other types of bulges we examined, the 3-prime base-bulge enables 6-MI to sample a greater percentage of the conformations where non-radiative decay is favorable.

In summary the destabilization to the DNA conformation is independent of base-bulge location (Table 5-2). The photophysical properties of 6-MI are dependent on the location of the base-bulge (Table 5-2). We determined a base-bulge adjacent to 6-MI has a greater effect
on its solvent exposure and reduces ground state base stacking interactions. The base-bulge adjacent 6-MI reduces the QY with the largest reduction due to the 3-prime base-bulge (Table 5-2). The excited state decay subpopulations are shifted to faster lifetimes indicating an increase in the non-radiative rate (Table 5-2). Lastly, we note a correlation of increased 6-MI rotational motion and an increase in non-radiative decay (Table 5-2).

_The photophysical properties of the 6-MI: Adenine mismatch indicates proper H-bonding is necessary for enhancement of fluorescence_

We are interested in the relationship between conformational dynamics and QY. One way to increase 6-MI flexibility is to reduce H-bonding by a mismatched 6-MI/adenine base pair. When adenine is mismatched with 6-MI there is destabilization of the DNA conformation. As shown in Figure 8, there is a shift in 340 nm Tm with an adenine mismatch by ~2. ± 2 K, consistent with destabilizing 6-MI within the DNA (Table 5-4). The decrease in Tm highlights the importance of H-bonding to DNA stability. The destabilization of the DNA is coupled with an increase in solvent exposure (Table 5-2, Figure 5-7), indicating the conformation of 6-MI is altered compared to the duplex DNA.

The 6-MI rotational correlation time (θ_l) is faster when mismatched with adenine compared to the base pairing with cytosine, which suggests reducing the ability of 6-MI to correctly hydrogen allows for an increase in conformational flexibility. Interestingly, for the construct A(mis)_ATFAA15 6-MI θ_l is 0.25 ± 0.12 ns longer than 3(+A)_ATFAA15, which suggests a more significant role of the 3-prime adjacent base in hindering 6-MI conformational dynamics (Figure 11, Table 5-2). When 6-MI is mismatched with adenine in the quenched sequence A(mis)_CAFTC07, the β_1 component of the anisotropy decay increases by 230% (Figure 5-11,
Table 5-2). Again, for 6-MI in the quenched sequence CAFTC07, perturbing its conformational dynamics does not lead to enhancement of fluorescence (Table 5-2).

The 6-MI/adenine mismatch exhibits a blue shift in the emission spectra in the quenched and fluorescently enhanced constructs (Figure 5-8). The magnitude of the blue shift is similar for all perturbations, demonstrating the emission maximum is sensitive to slight changes to 6-MI conformation and dynamics. For the fluorescently enhanced sequence, when 6-MI is mispaired with adenine the QY$_{rel}$ is reduced by ~50%, which suggests hydrogen bonding is essential to limit 6-MI from adopting a favorable conformation for non-radiative decay (Figure 5-8). The introduction of a base-bulge or mismatch to the quenched sequence CAFTC07ds, has minimal effect on the QY (Figure 5-8 Table 5-2). The reduction in QY$_{rel}$ from the fluorescently enhanced duplex ATFAA15ds to the 6-MI/adenine mismatch A(mis)$_{ATFAA15}$ is associated with a shift in the excited state decay from largely a single long lifetime, $\tau_3$ (7 ns) = 94 ± 2%, of the population to multiple subpopulations with shorter lifetimes, indicating greater conformational heterogeneity for 6-MI and an increase in the non-radiative decay rate (Table 5-2). We also see the lifetime representing $\tau_1$ decreases by 0.75 ± 0.06 ns from the quenched CAFTC07 sequence to the 6-MI/adenine mismatch A(mis)$_{CAFTC07}$ (Figure 5-9), suggestive of an increase in the non-radiative rate. Interestingly, while $\theta_L$ for A(mis)$_{ATFAA15}$ is slightly larger compared to 3(+A)$_{ATFAA15}$, the molecular weighted mean lifetime $\tau_{(mol)}$ is 0.77 ± 0.06 ns shorter for the 6-MI/adenine mismatch (Table 5-2, Figure 5-10). One possible explanation is while the 6-MI/adenine mismatch does increase 6-MI conformational dynamics it alters the geometry of the adjacent bases to a lesser extent compared to the base-bulge. Thus, 6-MI can readily adopt a conformation favoring non-radiative decay pathways.
5.2.3 Molecular dynamics simulations indicate the conformation of 6-MI is dependent on sequence

Molecular dynamic simulations were carried out to examine the structural differences between fluorescently quenched and enhanced dsDNA in an attempt to unmask the key structural components. As shown in Figure 5-12, the RMSD from the initial structure suggests convergence of the simulations after ~1.5 ns. We did not observe any global structural difference, such as DNA bending, between the MD averaged structures CAFTC07ds and ATFAA07ds as indicated by the 0.3 Å RMSD between phosphate backbones. We parameterized the snapshots of the simulations (10ps/snapshot) for each base step using the program 3DNA.\textsuperscript{260,283} Using 3DNA the base step is defined by six parameters: rise, shift, slide, twist, tilt, and roll. We then binned the snapshot base step parameters to generate frequency of occurrence plots, which we then modeled as a Gaussian distribution. The Gaussian centers and FWHM of the curves provide a method to determine which structural parameters vary between simulations. We focused on the base steps 5-prime to 6-MI step (X-F step) and the 6-MI to 3-prime step (F-X). As a control we examined the T to G (15\textsuperscript{th} position) for CAFTC07 and ATFAA07 a good indicator of variations between simulations. As shown in Figure 5-13, there is excellent agreement between simulations at the 15\textsuperscript{th} position, which suggests any variation in the 6-MI steps is the result of the local sequence and not an artifact of the simulation. There are several differences in the base step parameters between sequences with and without 6-MI, which suggest incorporating 6-MI into
The RMSD from initial structure for MD snapshots indicates convergence of the simulation after ~1.5 ns. The deviation from the initial structure is similar for all simulations.
a construct slightly alters the local base structure compared to the guanine base (Table 5-5, Figures 5-14 through 5-19). We see the greatest difference in base step parameters between sequences with and without 6-MI in the X-F base step (Table 5-5, Figure 5-14 to 5-19). As shown in Figure 5-16 for constructs without 6-MI, CAGTC07 and ATGAA07, the center of the distribution for slide parameter is -0.49 and -0.13 Å respectively, consistent with a Slide of -0.22 Å for B-form DNA (Table 5-5 and 5-6). This is what we expected because the initial conformation of the constructs was B-form. In contrast, when the construct contains 6-MI, CAFTC07 and ATFAA07, the center of the distribution for slide parameter is -1.46 and -0.84 Å respectively, closer to the A-form Slide standard of -1.74 Å (Table 5-5 and 6-6). Thus, replacing guanine with 6-MI causes the X/F basepairs to slightly slide apart possible to accommodate the methyl group at the 6th position in 6-MI. In addition, while the center of the Shift parameter distributions for CAGTC07 and ATGAA07 is -0.7 and 0.16 Å, similar to the B-form standard of 0.05 Å (Table 5-5 and 6-6 Figure 5-15); there is a greater deviation from B-form for sequence with 6-MI. The center of the Shift distribution for CAFTC07 and ATFAA07 is 0.58 and 0.87 Å respectively (Table 5-5). We see the greatest difference between sequences with and without 6-MI is in the Twist parameter. For example there is an 11.45° difference between CAGTC07 and CAFTC07 (Table 5-5, Figure 5-20). The extent of undertwisting in CAFTC07 is an important observation that is discussed in greater detail below. Another interesting observation is the deviation in basetep parameters between sequences with and without 6-MI is greater for the quenched sequence CAFTC07 (Table 5-5, Figures 50-20). In addition, for most base-step parameter the spread in the distribution for the 6-MI containing sequence CAFTA07, is smaller compared to the sequence without 6-MI, CAGTC (Table 5-5, Figures 5-14 5-19). In contrast, for ATFAA07 and ATGAA07 the
### Table 5-5a. Gaussian fitted center of peak for parameters describing base/6-MI step

<table>
<thead>
<tr>
<th></th>
<th>Rise $\sigma$</th>
<th>Shift $\sigma$</th>
<th>Slide $\sigma$</th>
<th>Twist $\sigma$</th>
<th>Tilt $\sigma$</th>
<th>Roll $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAFTC07</td>
<td>3.34</td>
<td>0.74</td>
<td>0.58</td>
<td>1.60</td>
<td>-1.46</td>
<td>0.77</td>
</tr>
<tr>
<td>ATFAA07</td>
<td>3.45</td>
<td>0.90</td>
<td>0.87</td>
<td>1.30</td>
<td>-0.84</td>
<td>1.07</td>
</tr>
<tr>
<td>AAFTA10</td>
<td>3.44</td>
<td>0.78</td>
<td>0.48</td>
<td>1.32</td>
<td>-1.52</td>
<td>0.87</td>
</tr>
<tr>
<td>GTFTG32</td>
<td>3.39</td>
<td>0.78</td>
<td>0.71</td>
<td>1.33</td>
<td>-0.89</td>
<td>0.91</td>
</tr>
<tr>
<td>CTFAC07</td>
<td>3.32</td>
<td>0.74</td>
<td>0.44</td>
<td>1.52</td>
<td>-0.93</td>
<td>0.96</td>
</tr>
<tr>
<td>TTFCA18</td>
<td>3.26</td>
<td>0.70</td>
<td>0.61</td>
<td>1.34</td>
<td>-0.83</td>
<td>0.88</td>
</tr>
<tr>
<td>ATFTA19</td>
<td>3.47</td>
<td>0.81</td>
<td>0.86</td>
<td>1.39</td>
<td>-0.82</td>
<td>1.03</td>
</tr>
<tr>
<td>CAGTC07</td>
<td>3.47</td>
<td>0.75</td>
<td>-0.70</td>
<td>1.87</td>
<td>-0.49</td>
<td>1.57</td>
</tr>
<tr>
<td>ATGAA07</td>
<td>3.24</td>
<td>0.85</td>
<td>0.16</td>
<td>1.76</td>
<td>-0.13</td>
<td>1.85</td>
</tr>
<tr>
<td>A(mis)_ATFAA07</td>
<td>3.48</td>
<td>0.85</td>
<td>-0.05</td>
<td>1.22</td>
<td>-1.46</td>
<td>0.96</td>
</tr>
<tr>
<td>A(mis)_ATFAA07</td>
<td>1.61</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 5-5b. Gaussian fitted center of peak for parameters describing 6-MI/base step

<table>
<thead>
<tr>
<th></th>
<th>Rise $\sigma$</th>
<th>Shift $\sigma$</th>
<th>Slide $\sigma$</th>
<th>Twist $\sigma$</th>
<th>Tilt $\sigma$</th>
<th>Roll $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAFTC07</td>
<td>3.30</td>
<td>0.62</td>
<td>-0.53</td>
<td>1.40</td>
<td>-0.79</td>
<td>1.14</td>
</tr>
<tr>
<td>ATFAA07</td>
<td>3.23</td>
<td>0.65</td>
<td>-0.63</td>
<td>1.32</td>
<td>-0.11</td>
<td>1.77</td>
</tr>
<tr>
<td>AAFTA10</td>
<td>3.31</td>
<td>0.62</td>
<td>-0.67</td>
<td>1.49</td>
<td>-0.63</td>
<td>1.22</td>
</tr>
<tr>
<td>GTFTG32</td>
<td>3.38</td>
<td>0.69</td>
<td>0.02</td>
<td>1.57</td>
<td>-0.89</td>
<td>1.09</td>
</tr>
<tr>
<td>CTFAC07</td>
<td>3.19</td>
<td>0.67</td>
<td>-0.55</td>
<td>1.34</td>
<td>-0.15</td>
<td>1.37</td>
</tr>
<tr>
<td>TTFCA18</td>
<td>3.15</td>
<td>0.47</td>
<td>-0.33</td>
<td>1.37</td>
<td>-0.20</td>
<td>1.16</td>
</tr>
<tr>
<td>ATFTA19</td>
<td>3.28</td>
<td>0.62</td>
<td>-0.66</td>
<td>1.44</td>
<td>-0.49</td>
<td>1.18</td>
</tr>
<tr>
<td>CAGTC07</td>
<td>3.41</td>
<td>0.71</td>
<td>-0.19</td>
<td>1.76</td>
<td>-0.84</td>
<td>0.94</td>
</tr>
<tr>
<td>ATGAA07</td>
<td>3.42</td>
<td>0.79</td>
<td>-0.52</td>
<td>1.67</td>
<td>-0.13</td>
<td>1.76</td>
</tr>
<tr>
<td>A(mis)_ATFAA07</td>
<td>3.16</td>
<td>0.66</td>
<td>0.05</td>
<td>1.67</td>
<td>-0.81</td>
<td>1.02</td>
</tr>
<tr>
<td>A(mis)_ATFAA07</td>
<td>0.57</td>
<td>1.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9000 snapshots (10ps/snapshot) base-steps, either X / 6-MI or 6-MI / X base-step, were parameterized based on 6 parameters Rise, Shift, Slide, Twist, Tilt, and Roll using the program 3DNA. The ranges of values for each parameter were binned and the histograms were fit to a Gaussian function. The center value of the Gaussian peak is represented in the above table. The only parameter where there is a deviation between AFT07 and TFA07 followed by a decrease in the deviation after changing the sequence to CAFTC07 was Twist for the X/6-MI base step $\sigma$ = FWHM for Gaussian distribution.
Table 5-6 Base step parameters for B-form and A-form DNA

<table>
<thead>
<tr>
<th>Base steps</th>
<th>Shift (Å)</th>
<th>Slide (Å)</th>
<th>Rise (Å)</th>
<th>Tilt (º)</th>
<th>Roll (º)</th>
<th>Twist (º)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG/CT A-form</td>
<td>0.05</td>
<td>-1.74</td>
<td>3.28</td>
<td>-0.13</td>
<td>10.75</td>
<td>29.88</td>
</tr>
<tr>
<td>AG/CT B-form</td>
<td>0.05</td>
<td>-0.22</td>
<td>3.37</td>
<td>0.04</td>
<td>-2.81</td>
<td>34.96</td>
</tr>
<tr>
<td>TG/CA A-form</td>
<td>0.06</td>
<td>-1.72</td>
<td>3.28</td>
<td>-0.15</td>
<td>10.73</td>
<td>30</td>
</tr>
<tr>
<td>TG/CA B-form</td>
<td>0.06</td>
<td>-0.22</td>
<td>3.37</td>
<td>0.03</td>
<td>-2.82</td>
<td>34.79</td>
</tr>
<tr>
<td>GT/AC A-form</td>
<td>0</td>
<td>-1.68</td>
<td>3.27</td>
<td>-0.04</td>
<td>10.78</td>
<td>31.85</td>
</tr>
<tr>
<td>GT/AC B-form</td>
<td>-0.03</td>
<td>-0.16</td>
<td>3.38</td>
<td>0.02</td>
<td>-2.79</td>
<td>37</td>
</tr>
<tr>
<td>GA/TC A-form</td>
<td>-0.01</td>
<td>-1.66</td>
<td>3.27</td>
<td>-0.03</td>
<td>10.76</td>
<td>31.97</td>
</tr>
<tr>
<td>GA/TC B-form</td>
<td>-0.03</td>
<td>-0.16</td>
<td>3.38</td>
<td>0.03</td>
<td>-2.8</td>
<td>36.83</td>
</tr>
</tbody>
</table>

Table 5-6 Standard values for A and B form for each of the four step
A comparison of CAFTC07ds (■) and ATFAA07ds (●) base-step parameters for the 14 to 15 step. There is excellent agreement in base-step parameters between the two simulations, which suggests any variation in the 6-MI steps is the result of the local sequence and not an artifact of the simulation.
spread in the distribution for most of the base-step parameters are similar (Table 5.0 and Figures 5-14 5-19). There are several differences in the base-step parameters between the CAFTC07 and ATFAA07 simulations, notably in Slide, Twist, Tilt, and Roll (Table 5-5). For example, there is a -0.62 Å difference between the center of the distributions for CAFTC07 and ATFAA07 in Slide for the X-F step. Similarly, the difference between the two in the F-X step Slide parameter is-0.68 Å (Table 5-5 Figure 5-16). In addition, the difference in the Twist parameter X-F step center of the distribution between CAFTC07 and ATFAA07 is ~2.9° (Figure 5-20, Table 5-5). We also noticed there is a difference of ~2.8° and ~2.0° in the X-F and F-X step Tilt center of distribution between CAFTC07 and ATFAA07 (Table 5-5, Figure 5-18). Finally, the greatest difference between CAFTC07 and ATFAA07 base-step parameters is the in base-step Roll, for the X-F step the difference is 7.3° and for the F-X step the difference is 3.1° (Table 5-5, Figure 5-20). To distinguish which parameters best describe the structural difference between fluorescently quenched and unquenched sequences a larger dataset was necessary, simulations of AAFTA10 (sequence with enhanced fluorescence), GTFTG32 (sequence reported by Knutson and coworkers with enhanced fluorescence), CTFAC07 (changed the sequence of ATFAA07, expected no enhancement of fluorescence), and the fluorescently enhanced ATFTA19 were performed to parse out the which base step parameters are significant.

Analysis of the parameterized base-steps for AAFTA10 revealed a several similarities to CAFTC07. For example, while we reported a difference in the X-F base-step Roll between CAFTC07 and ATFAA07; the X-F base step Roll distribution for AAFTA10 is almost identical to CAFTC07 (Figure 5-19). We determined the Twist parameter for X-F step is the only parameter where the AAFTA10 distribution is comparable to ATFAA07 (Table 5-5, Figure 5-20). The center of the Twist X-F Gaussian distribution for AAFTA10 is 26.9°, which deviates
from the 24.3° Gaussian center for CAFTC07 and is closer in value to the 27.9° Gaussian center for ATFAA07 (Figure 5-20, Table 5-5). In addition, analysis of simulations for the fluorescently enhanced GTFTG32 and ATFTA19 revealed the Gaussian center for the X-F Twist parameter distribution is 27.7° and 29.74°, consistent with the distribution for ATFAA07 (Figure 5-20, Table 5-5). We previously had shown the identity of n ± 2 bases adjacent to 6-MI are critical for enhancement of fluorescence. MD simulations allowed us to model the effect of these bases on 6-MI base-step conformation. When the sequence of ATFAA07 is mutated to CTFAC07, which is the same n±2 bases in CAFTC07 (Table 5-1), the Gaussian center for the twist parameter distribution is shifted to 25.3° (Figure 5-20, Table 5-5), comparable to the distribution for CAFTC07. Finally, we performed a MD simulation with an unrelated quenched sequence TTFCA18. Analysis of the parameterized base-steps revealed the Gaussian center for the X-F Twist distribution is 25.24° (Table 5-5), which suggests an untwisting of the X-F step in sequences with high degree of fluorescence quenching. The importance of correct H-bonding was also examined by simulating the 6-MI/adenine mismatch in the ATFAA07 sequence.

Crystal and N.M.R. structures of duplexes with a G:A mismatch have minimal perturbation to the global structure with local disruption of base stacking interaction as the result of an increase in base roll and propeller twist. We elected to use the 6-MI/adenine mismatch for simulating the effect of structural perturbation on 6-MI geometry because of the absence of a major kink in the DNA construct. The MD simulation of the 6-MI/adenine mismatch revealed several structural deformations from the guanine cytosine basepair (Table 5-5). We hypothesized the conformational changes are necessary to accommodate the larger basepair. For example, there is a large change in the X-F step Tilt distribution, the Gaussian center is located at 7.5° compared to 2.5° for ATFAA07. We also observed a sharp change in
the Tilt distribution for the F-X step, the sharp Tilt are a possible indicator of decreased stacking interactions between 6-MI and adjacent bases, consistent with photophysical data (Table 5-5, Figure 5-18). For the 6-MI/adenine mismatch base step parameters Slide and Shift, we observed two conformational distributions (Table 5-5 Figures 5-16 and 5-15). It is important to note, the simulation may need to run longer to overcome any energy barriers to converge to a single population. The important observation was that the mismatch leads to untwisting of the X-F step. As shown in Figure 5-20, the Gaussian center for the Twist X-F base step distribution is 24.8°, consistent with the distributions for the quenched constructs (Table 5-4, Figure 5-7). The untwisting of X-F step in the 6-MI/adenine mismatch DNA suggests fluorescence quenching is the result of a slightly less ordered conformation. The MD simulations data indicate specific conformation may lead to enhancement of fluorescence. It is interesting that the difference in conformation between the sequences is minimal.
For constructs containing 6-MI, the distance between 5-prime base and 6-MI is on average 3.38 Å with an average FWHM (σ) of 0.79 Å. Interestingly, the spread in the distribution is smaller for CAFTC07 and CAGTC07 compared to ATFAA07 and ATGAA07 (C-D): 6-MI/X step The distance between 6-MI and 3’base is on average 3.28 Å with a spread in the distribution of 0.65 Å. Interestingly, for the 6-MI/adenine mismatch the average base step distance is ~3.16 Å, indicating a slight compaction of the 6-MI/adenine step. For sequences without 6-MI, CAGTC and ATGAA the distance between basepairs is ~3.41 Å with σ of 0.75 Å, which suggests the 3’ prime basepair is closer to 6-MI than guanine.

Figure 5-15 Rise Base step
The shift between the 5-prime base and 6-MI is an average 0.36 Å for all 6-MI simulations with the exception of the 6-MI/adenine mismatch. Introduction of the adenine mismatch leads to 2 populations with the greater population presenting a smaller average shift of -.047 Å. The average shift for sequences without 6-MI (CAGTC and ATGAA) is less than that for 6-MI containing sequences. In addition, there is a wider distribution. For the 6-MI to 3-prime base step the average shift is similar for CATFC07 and ATFAA07. Interestingly, the GTFTG32 construct and the 6-MI/adenine mismatch exhibit a greater shift between 6-MI and 3’ adjacent base-step compared to the other 6-MI constructs.

Figure 5-16 Slide base step

(A-B) X/F base step

(C-D) F/X base step
The amount of slide between the 5-prime base and 6-MI can be separated into two groups: 1) is centered at ~1.4 Å consisting of CAFTC07, AAFTA10 and A(mis)_ATFAA07 and 2) consisting of the remaining sequences centered at ~0.86Å. The plot comparing sequences with and without 6-MI suggests the presence of 6-MI requires greater staggering of the basepairs. Also, there is a greater distribution in slide for sequences without 6-MI (C-D) Slide F/X base step. The slide parameter for the 6-MI to 3-prime base step in ATFAA07 is centered at -0.11 Å smaller compared to CAFTC07 which is centered at -0.79Å, which at first suggest this parameter is important for enhancement of fluorescence. However, for CTFAC07 slide is centered at -0.15Å, similar to ATFAA07. The 6-MI/X slide for adenine mismatch A(mis)_ATFAA07 again indicates 2 populations with one conformation similar to ATFAA07 and the other closer to CAFTC07. For sequences without 6-MI the slide parameter is consistent with 6-MI containing sequences.
The twist between the 5-prime base and 6-MI is the one parameter consistent with fluorescent data. For example, the twist in ATFAA07 is centered at 27.3°, while in CAFTC07 it is centered at 24.4°. In addition, altering ATFAA07 n+2 to cytosine shifts the center of the twist to 25.6°, similar to CAFTC07. Also, the twist distribution for the 6-MI/adenine mismatch A(mis)_ATFAA07 resembles CAFTC07. Finally, the distribution in twist is similar for all fluorescently enhanced sequences (ATFAA07, AAFTA10, GFITGG32, and ATFTTA19). Interestingly, the distribution in CAGTC twist is centered at 35.8°, which suggest under twisting of the duplex when 6-MI is inserted into quenched duplex (C-D) Twist F/X base step he distribution in twist between the 6-MI and 3-prime base is very similar for all sequences, with the exception of the 6-MI/adenine mismatch. The distribution for the 6-MI/adenine mismatch A(mis)_ATFAA07 indicates untwisting of the DNA duplex possibly to accommodate the mispaired base. The disparity between WT sequences guanine – thymine (33.0) and guanine- adenine (36.7) is
The tilt between the 5-prime base and 6-MI ranges from TFA07 (2.51°) to AFT07 (-0.76°), with no consistent trend with experimental data. The adenine mismatch requires tilting of the 5-prime base (7.5°) to accommodate the mispaired adenine. The tilt observed in the WT sequence is similar to sequences containing 6-MI.

The tilt between the 6-MI and 3-prime base spans a range of 4.5° for CTFAC07 (-2.13°) and TFT32 (2.53°), again with no consistent trend observed in fluorescence data. The adenine mismatch requires tilting of the 3-prime base (-5.62°) in the opposite direction to accommodate the mispaired adenine. The WT data is consistent with 6-MI containing sequences.

Figure 5.19 Tilt base step

(A-B) Tilt X/F base step. The tilt between the 5-prime base and 6-MI ranges from TFA07 (2.51°) to AFT07 (-0.76°), with no consistent trend with experimental data. The adenine mismatch requires tilting of the 5-prime base (7.5°) to accommodate the mispaired adenine. The tilt observed in the WT sequence is similar to sequences containing 6-MI.

(C-D) Tilt F/X base step. The tilt between the 6-MI and 3-prime base spans a range of 4.5° for CTFAC07 (-2.13°) and TFT32 (2.53°), again with no consistent trend observed in fluorescence data. The adenine mismatch requires tilting of the 3-prime base (-5.62°) in the opposite direction to accommodate the mispaired adenine. The WT data is consistent with 6-MI containing sequences.
Figure 5-20 Roll base step

(A-B) Roll X/F base step. The roll between the 5-prime base and 6-MI can be divided into two groups one centered at ~3.0°, AFT07 and AFT10 and the other group at ~10° made up of the remaining sequences. The roll observed in the WT sequence is similar to sequences containing 6-MI (C-D) Roll F/X base step. The Roll between the 6-MI and 3-prime base spans a range of ~5° with TFT32 (0.22°) and TFA07 (5.0°), again with no reliable trend to fluorescence data. The adenine mismatch leads to increased roll. WT data is consistent with 6-MI containing sequences.
Figure 5-21: The twist between the 5-prime base and 6-MI is the one parameter where there is a difference between ATFAA07 (red ●) the twist is 27.3° FWHM (σ) = 10.6° and CAFTC07 (black ■) the twist is 24.4 σ = 8.3°. When the adenine n+2 bases are replaced with cytosine in the ATFAA07 construct CTFAC07 (cyan ◄) the twist angle decreases to 25.6 σ = 10.38°, similar to CAFTC07. In addition, a conformational change occurs with the introduction of the 6-MI/adenine mismatch A(mis)_ATFAA07 (orange ►) the twist angle decrease to 24.8 σ = 8.32°, consistent with the CAFTC07 construct. Finally twist angle for the fluorescently enhanced GTFTG32 (blue ●), CAFTC10 (green ▲), and ATFTA19 (magenta ▼) is 27.74 σ = 9.45°, 26.98 σ = 8.96°, and 29.74 σ = 11.15°, respectively. Thus, the degree of twist in the fluorescently enhanced sequences is greater compared to the quenched sequences.
5.3 Discussion

In our previous study, we reported on the enhanced fluorescence for 6-MI in the pentamer sequence ATFAA/AAFTA dsDNA, and proposed a model where 6-MI enhanced fluorescence was the result of a sterically constrained conformation that minimized collisional quenching of the excited state. Consistent with this model, we show that the internal motion of 6-MI is decreased within the sequence ATFAA/AAFTA. Earlier work focusing on the n+1 flanking sequence role in fluorescence quenching of 6-MI and 3-MI (6-MI homolog) indicates the following order of increasing QY T ~ C >> G ~ A.\(^{31,42}\) We noticed the pentamer ATFAA/AAFTA are outliers to this model. Thus, we are interested in the cause of these outliers. In this work we presented results on QY of sequences where two adenines were 5’ to 6-MI and the 3’ side was occupied by C, T or G (Table 1) We reported 6-MI QY order ranking in AAFX sequences, including the pentamer sequence (AAFTA), which indicates the order of increasing QY with T >> G = C = A (Table 2 and 3). Interestingly, the order of 6-MI internal motion in the AAFX series fast to slow is C < G < T (Table 5-2 and 5-3). We also examined several other sequences and culled the literature but were unable to determine a clear sequence dependent pattern for quenching or enhancement of fluorescence in duplex DNA (Table 5-2). A larger dataset containing all pentamer sequence permutation would be necessary to generate a complete model.

Based on our previous work and this data, we hypothesized a certain degree of flexibility was necessary for 6-MI to adopt a conformation favorable for fluorescence quenching. To gain further insight into this relationship and the general mechanism of fluorescence quenching systematic base-bulges were introduced in the quenched and enhanced sequences reducing specific base interactions with 6-MI to increase conformational flexibility. The systematic perturbations afforded a gradient of destabilization to 6-MI structure and dynamics. In the
discussion we aim to emphasize the following conclusions related 6-MI fluorescence: i. the observed changes to fluorescence reflect the relative position of the perturbation to 6-MI  ii. reducing ground state base stacking does not lead to enhancement of fluorescence iii. conformational flexibility is correlated with dynamic quenching of 6-MI excited state iv. the 3-prime base-bulge adjacent to 6-MI and adenine mismatch relieve constraints to 6-MI dynamics v. molecular modeling indicates the conformation of 6-MI is sequence dependent

5.3.1 *The observed changes to fluorescence reflects the position of 6-MI relative to the perturbation*

The three types of perturbations all lead to a localized DNA instability with varying effects on 6-MI conformational dynamics and photophysics. Our results demonstrate the change in 6-MI fluorescence is greatest when the base-bulge is adjacent to 6-MI on the 3’ side or 6-MI is mispaired to adenine. We know from crystal and N.M.R. structures of base-bulges, the distortion to the DNA is highly localized with a location dependent effect on the conformation of surrounding bases.\textsuperscript{264,285-288} For example, an extra purine has been shown to insert into one strand, which in turn forces bases on the complementary strand apart in a wedge conformation disrupting base stacking interactions.\textsuperscript{285,288} To accommodate the extra purine there is stretching in the complementary strand rather than compression of the strand containing the bulge.\textsuperscript{288} Similar to the base-bulge, the perturbation from a G:A mismatch is highly localized. The G:A mismatch is known to increase base roll and propeller twist at the site of the mismatch, leading to local disruption of base stacking.\textsuperscript{284} The base-bulge and mismatch structural data suggests there should be a change in 6-MI conformation and the magnitude of the change will depend on the relative location of the perturbation to 6-MI. We anticipated that local DNA structural deformation would lead to significant changes in 6-MI conformational flexibility and
fluorescence, which we could use to gauge the relationship between the two. In fact we found that 6-MI fluorescence is surprisingly insensitive to most of the perturbations. This is unexpected given how sensitive the quantum yields are to primary sequence, particularly the special AFTAA pentamer. Alterations in this sequence lead to dramatic changes in the quantum yield, yet introduction of structural perturbations in the context of AFTAA as presented above often led to only small changes in quantum yield. Similarly, many of the changes introduced to quenched sequences did not significantly increase the quantum yield. We know from monitoring the melting temperature of our duplexes (ΔTm, Table 5-4), that the base-bulge reduces the stability of the dsDNA similar to previous literature results (Table 5-4, Figure 5-6). In addition, the magnitude of destabilization between the quenched and fluorescently enhanced sequence are comparable signifying no obvious thermodynamic benefit from the pentamer sequence ATFAA or AFTAA. We see from Table 5-4, the base-bulges lead to equivalent levels of DNA destabilization independent of their location within the DNA. We do know from our results 6-MI fluorescence is influenced by the location of these perturbations (Table 5-2, Figure 5-8).

As shown in Table 5-2 the base-bulge on the complementary strand increases 6-MI solvent exposure by 80% but there is only a 12% reduction in quantum yield, while both the 3-prime adjacent base-bulge to 6-MI and the adenine mismatch result in a 160% increase in solvent exposure and a 50% reduction in QY (Table 5-2). This underscores the complexity between 6-MI conformational dynamics and fluorescence; since solvent accessibility is imperfectly correlated with QY other factors must be considered. One can envision the extra base on the complementary strand leads to a restrained wedge conformation of 6-MI maintaining enhanced fluorescence while increasing solvent exposure. In contrast, the solvent exposure increases from the base-bulge adjacent to 6-MI and adenine mismatch, are due to 6-MI conformational
flexibility. We also noted the solvent exposure is greatest when the base-bulge is adjacent to 6-MI in the quenched sequence CAFTC07, another indication of the differences in conformational flexibility between the quenched and fluorescently enhanced sequences (Table 5-2). Interestingly, destabilizing the quenched sequence AFT07ds yields an increase in solvent exposure with insignificant change to the quantum yield, which suggests 6-MI is not extrahelical and dynamic quenching is maintained. The comparison of 6-MI solvent exposure with Q.Y. indicates destabilizing the surrounding bases is not enough, but increased 6-MI flexibility is necessary for dynamic quenching. We hypothesized the perturbations would reduce ground state base-stacking interactions and were interested in their relationship to 6-MI fluorescence. An indicator of ground state base-stacking interactions and their relationship with fluorescence quenching is hypochromicity.

Reducing ground state base stacking does not lead to enhancement of fluorescence

The hypochromic effect is the result of coupling between a particular electronic excited state of the nucleoside analog and different excited state of surrounding bases; this leads to a smaller transition dipole and decreased absorption and is a powerful indicator of stacking between bases.\textsuperscript{18,290,294} We were surprised to observe the similarity in 6-MI hypochromicity for the fluorescently enhanced AAFTA10ds which is 31 ± 2% and the 34 ± 6% hypochromicity in the quenched CAFTC07ds sequences, indicating enhancement of fluorescence is independent of ground state base stacking interactions (Table 5-4). As we hypothesized the base-bulge and mismatch reduced the hypochromicity, reflecting diminished 6-MI ground state base-stacking interactions. The reduction in 6-MI hypochromicity for both the quenched and fluorescently enhanced sequence is correlated with the relative position of the base-bulge or mismatch to 6-MI (Table 5-4). This further supports the idea that DNA destabilization is highly localized.
Interestingly, enhanced fluorescence was not observed upon reduction in hypochromicity for the quenched sequence CAFTC07, indicating dynamics in the excited state effectively quenches 6-MI fluorescence (Table 3, 4). More specifically, the reduction in hypochromicity was greatest for the base-bulge adjacent to 6-MI and the mismatched adenine, where we observe a significant decrease in quantum yield. This signifies that fluorescence quenching is related to excited state dynamics (Table 5-2 and 5-4). We determined the observed alteration in 6-MI hypochromicity are dependent on the relative position of 6-MI to the perturbation.

In summary there is clear evidence that the perturbations introduced by mismatched base and base-bulge are influencing the local structural environment of the 6-MI probe. Changes in hypochromicity are observed, which directly reflect changes in dipole interactions. However, these changes are not closely correlated with the quantum yield as would be expected if ground state interactions were responsible for quenching 6-MI within duplex DNA. This suggests that the varying quenching rates are not related to static structures but are more likely due to changes in excited state dynamics, occurring subsequent to absorption.

*6-MI fluorescence is directly correlated with conformational flexibility and dynamic quenching of the excited state*

In our previous paper we proposed a model where 6-MI enhanced fluorescence is the result of a sterically restrained conformation that minimizes collisional quenching of the excited state. We theorized destabilizing the DNA structure would lead to increased 6-MI conformational flexibility and dynamic quenching of the excited state. The TR-anisotropy decay data is consistent with a correlation between 6-MI rotational motion and dynamic quenching of the excited state (Table 5-2). In the results we highlighted when the base-bulge is on the complementary strand there is minimal increase in 6-MI flexibility and little change to excited
state decay (Table 5-2). We had expected any local distortion to the DNA would lead to an increase in 6-MI flexibility and increased dynamic quenching of fluorescently enhanced sequences ATFAA, AAFTA, and ATFTA. Examination of fluorescently enhanced 6-MI TR-anisotropy reveals the constrained conformation is maintained when the base-bulge is on the complementary strand (Figure 5-11). The increased access to external quencher and hindered 6-MI local motion signifies the DNA structure is able to accommodate the base-bulge with localized destabilization to DNA structure. The lack of change in excited state decay for the bulged-base on the complementary strand supports this model (Table 5-2, Figure 5-10).

Previous TR-anisotropy work on 2-AP and base-bulge illustrate increased motion only occurs when 2-AP is adjacent to the bulge, due to the base-bulge reducing the steric constrains only on the adjacent bases.\textsuperscript{295} Similar to this 2-AP finding, when the base-bulge is adjacent to 6-MI the constraints to flexibility are relieved (Table 5-2, Figure 5-11).

A base-bulge adjacent to the 6-MI destabilizes the DNA with an increase in local flexibility of adjacent bases. The increased motion of 6-MI (Table 5-2) as a result of the perturbation is correlated with a reduction in quantum yield, which supports a model where enhanced fluorescence results from a rigid conformation of 6-MI within duplex DNA. The increase in flexibility is greatest for the 3-prime base-bulge adjacent to 6-MI, which corresponds to the maximum dynamic quenching (Table 5-2). When the base-bulge is 5 prime to 6-MI the change in excited state lifetime and rotational correlation time is not as significant, signifying the base-bulge has different effects on the 5-prime and 3-prime side of 6-MI (Table 5-2). The difference in excited state lifetime for 3-prime and 5-prime base-bulge are correlated with the local motion of 6-MI (Table 5-2). The difference in fluorescence quenching based on the location of the bulge, either 5-prime or 3-prime, indicates the bulge has an asymmetric influence on adjacent base conformation. One possible explanation is the bulging base has been shown to stack with
the 5-prime base adjacent to the bulge\textsuperscript{287,288,296} (in this case 6-MI in 3(+A)\_ATFAA155 and 3(+T)-AAFAT10) and the increased internal motion of 6-MI may lead to a conformation favorable for dynamic quenching with the bulging base. The kink at the bulge site would prevent the 3-prime 6-MI, in this case 5(+T)\_ATFAA15 and 5(+A)\_AAFTA10, from stacking with the bulged base thereby maintaining the conformation with enhanced fluorescence. We theorized a base-bulge in the quenched sequence would minimize 6-MI interactions with adjacent bases leading to enhancement of fluorescence, which we did not observe (Table 5-2).

Interestingly, in the quenched sequence AFT07 the adjacent base-bulge leads to a 56% increase in $\beta_1$ population for the rotational correlation time, while maintaining dynamic quenching (Figure 5-11, Table 5-2). It appears that restraining the local motion of the fluorophore is necessary for the enhancement of fluorescence. Interestingly, previous work with DNA containing 2-AP indicated freezing has an effect on the excited state decay and quantum yield; at 77K there is an absence of shortest decay $\tau_1$ and increased QY implying thermal fluctuations are needed to access the dynamically quenched state.\textsuperscript{37} We observed a similar trend where distorting the fluorescently enhanced duplexes had little effect on lifetimes, $\tau_2$ and $\tau_3$ but leads to the appearance of a faster component ($\tau_1$) a new conformation of 6-MI that is dynamically quenched (Figure 5-10 Table 5-2). Based on hypochromicity and TR-anisotropy data the excited state decay component $\tau_1$ corresponds to an unstacked weakly associated 6-MI conformation where dynamic quenching is favorable. Nordlund and Wu also determined from the temperature dependence of the decay amplitudes for 2-AP in DNA, the component $\tau_1 = ~0.3$ ns is related an unstacked conformation.\textsuperscript{221} The quenching of fluorescence for 2-AP and 6-MI in DNA requires a dynamic system either to access a conformation favorable for quenching or collisional quenching with adjacent bases. Destabilizing the DNA structure
by mismatched adenine base leads to increased rotational motion, and dynamic quenching of the excited state.

We have shown hydrogen bonding between 6-MI and cytosine is necessary to prevent dynamic quenching of the excited state. The three hydrogen bonds between 6-MI and cytosine limit the internal motion of 6-MI preventing dynamic quenching of 6-MI fluorescence. The adenine mismatch reduces H-bonding between the bases, decreasing stability of the base pair and leading to greater internal motion of the 6-MI, where $\theta$ is decreased by 0.97 ns for A(mis)_ATFAA15 (Figure 5-11, Table 5-2). It is worth noting the rotational time is similar between A(mis)_ATFAA15 and A(mis)_CAFTC07 and the difference is in the $\beta_1$ term (Table 3). The TR-fluorescence measurements for the mismatch indicate the increased local motion of 6-MI leads to collisional quenching of the excited state (Figure 5-9, Table 5-2). The excited-state decay is shifted in the mismatch compared to dsDNA indicating the quenched population ($\tau_1$) corresponds to a 6-MI conformation with increased motion (Table 5-2). Investigation into the dynamics of 2-AP▪X mismatch revealed hydrogen bonding influenced the internal motion of 2-AP. The most stable base-pair 2AP▪T,297,298 limits fluorescence quenching due to favorable hydrogen bonding between 2-AP and T preventing 2-AP from conformational flexibility and collisional quenching with its adjacent bases.34 The adenine mismatch allows for greater flexibility of the 6-MI, which in turn leads to dynamic quenching of the 6-MI.

In summary, the correlation between the rotational correlation time of 6-MI and excited state lifetime reveals the relationship between conformational flexibility and dynamic quenching of the excited state. Introduction of a 3-prime base-bulge adjacent to 6-MI or an adenine mismatch reduces steric constraints to 6-MI’s conformation allowing dynamic quenching. From these results we can assign the specific lifetimes to 6-MI motion: $\tau_1$ corresponds to the
unstacked or weakly associated population, $\tau_3$ corresponds to the rigid conformation, and $\tau_2$ is an intermediate of the two conformations.

5.3.2 Molecular modeling indicates the conformation of 6-MI is sequence dependent

In the previous section we emphasized the correlation between conformational flexibility and dynamic quenching of the excited state. We proposed three distinct conformational subpopulations that give rise to the observed excited state decay of 6-MI in dsDNA. In this section our aim is to elucidate the conformation of 6-MI that limits dynamic quenching of the excited state. Although 2-AP and to a lesser extent 6-MI have been used extensively to investigate DNA dynamics, only a few studies have been conducted to model probe conformation in DNA using molecular dynamics simulations, and compare those results with experimental fluorescence measurements.$^{39,299,300}$ Nordlund and co-workers examined the dynamics of a decamer containing 2-AP, which revealed asymmetric displacement of 2-AP about the central axis of the helix corresponding to varying degrees of interbase stacking, to which they connected with the subpopulations of the excited state decay.$^{39}$ We found the introduction of 6-MI minimally perturbs the global DNA structure as shown by the slight deviation (<1 Å) between the DNA backbones of WT and 6-MI sequence, in agreement with thermodynamic data (Table 5-4). There are several minor differences between WT and 6-MI sequences reflected in the base-pair parameters (Table 5-4). This is not surprising since 6-MI is slightly larger than guanosine (with two 6-member rings) and contains an $\alpha$-carbonyl and $\beta$-methyl groups adjacent to the glycosidic nitrogen. A previous molecular dynamics study indicates 6-MI has minimal effect on overall DNA structure; the simulations however, revealed a higher propensity of opening for a 6-MI/uracil base-pair than guanine/uracil base-pair.$^{299}$ For MD simulations of 6-MI enhanced or quenched fluorescence dsDNA, analysis of the six
parameters describing a base-pair step reveals a correlation between the twist parameter and the experimentally observed results indicating twisting of the duplex is linked with enhancement of fluorescence (Table 5-5, Figure 5-20). Twist refers to the rotation around the helical axis between adjacent base-pairs. The quenched sequences exhibit an under-twisting of the DNA (Table 5-5). The under-twisting of the DNA may produce orthogonal transition dipole moments for 6-MI and adjacent bases leading to a favorable conformation for dynamic quenching. It is evident from the A(mis)_ATFAA07 MD simulation there is greater conformational flexibility not only for 6-MI but with the overall duplex DNA. The adenine mismatch simulation ( A(mis)_ATFAA07 ) reveals untwisting at the X-F base step observed suggesting there is a link between this 6-MI conformation and fluorescence quenching (Table 5-5, Figure 5-20). The mismatch simulation also exhibits increased tilt for the X-F step, decreased tilt for the F-X step, and increased slide to allow for the larger 6-MI:Adenine basepair, which has also been observed in the NMR structure. The structural changes observed in the adenine mismatch A(mis)_ATFAA07) MD simulation indicate greater conformational flexibility of the 6-MI base compared to TFA07 MD simulation and agree with increased rotational motion observed in TR-anisotropy (Table 5-2). Previous work examining electronic coupling between bases found significant correlation between the twist angle and electronic coupling for adenine-adenine interactions and the slide and shift parameters for thymine-thymine interactions. A change of twist angle was proposed to increase the overlap between adenine 5-membered rings, thus boosting the coupling between states. Analysis of the twist angle vs. time reveals random fluctuations around an average structure. We only observe a single 6-MI Twist population which may have to do with the short length of the simulation or starting from one initial structure. Further data analysis, beyond the scope of this
project, would be necessary to determine if there is correlated motion between the observed conformations.

5.3.3 Summary

We have previously reported on the sequence dependent enhancement of 6-MI fluorescence in duplex DNA. We hypothesized certain conformational constraints on 6-MI resulted in enhancement of fluorescence. For quenched sequences, conformational flexibility could either generate favorable ground-state interactions or lead to dynamic quenching of the excited state. To explore the relationship between 6-MI conformational flexibility and dynamic quenching we systematically introduced a base-bulge at different positions relative to 6-MI in addition to introducing an adenine:6-MI mismatch. We then characterized their effect on 6-MI photophysics in the quenched and fluorescently enhanced sequences. We discovered 6-MI photophysics was insensitive to base-bulge on the complementary strand. We found the 3-prime base-bulge adjacent 6-MI and adenine mismatch relaxed steric constraints on 6-MI in the fluorescently enhanced sequence leading to dynamic quenching and the appearance of a fast component in the excited state decay. The systematic distortions to DNA structure revealed a direct link between 6-MI conformational flexibility and dynamic quenching of the excited state. We also sought to create a better model establishing the behavior of 6-MI within duplex DNA and fluorescent properties using molecular dynamic simulations. Comparison between MD simulations of quenched DNA constructs and constructs with enhanced fluorescence suggested the under-twisting of the DNA leads to dynamic quenching of 6-MI.
6 Biophysical characterization of the HU-overhang complex

6.1 Introduction

Nucleoid-associated proteins are essential for regulating the structure and dynamics of the bacterial genome. Nucleoid-associated proteins bend, wrap and bridge the DNA to alter the local and global chromosomal geometry. Perturbations to chromosomal geometry affects multiple cellular processes including transcriptional regulation.  

HU or “histone-like” protein, is one of the most conserved nucleoid-associated proteins across bacterial genome. Cellular functions of HU include transcriptional regulation and structural maintenance of the bacterial chromosome. More specifically, HU influences DNA recombination and regulates the expression of a diverse array of genes, such as ones involved in cellular metabolism, respiration, translation, and ribosome biogenesis. HU also maintains negative supercoiling of plasmid DNA even when relaxed by topoisomerase I. Interestingly, the binding profile of HU to the bacterial genome is non-specific with a notable preference for A/T-rich regions.

Currently, there are no known topology-independent sequences bound specifically by HU. Instead, local structural features of the nucleic acid polymer dictate HU affinity. HU employs two separate binding modes when interacting with DNA depending on the conformational dynamics of the polymer, called structurally specific and nonspecific. The defining structural feature of the non-specific mode is a uniform nucleic acid polymer, that is either single- or double-stranded. Characteristics of the non-specific binding mode are population heterogeneity and sensitivity to component concentrations, including ionic strength. The specific binding mode of HU recognizes certain distortions to the nucleic
acid polymer that increase conformational dynamics, including ss/ds junctions (for simplicity from now on will be referred to as junction) as well as nicked, gapped, and branched DNA. In contrast to the nonspecific binding, the specific binding is between 1 and 4 orders of magnitude tighter and is insensitive to the concentration of salt. In both cases, HU functions as an architectural element introducing or stabilizing bends in the DNA. The protein-induced distortion facilitates the formation of higher-order functional complexes.

In *E. coli* HU exists as a heterodimer during stationary phase, composed of ~9kDa subunits (HUα and HUβ). Currently, there is no crystal or solution structure for *E. coli* HUαβ bound to a physiologically relevant nonspecific or specific construct, possibly due to the conformational heterogeneity of the HU-DNA complex. However, there is a crystal structure of *Anabaena* HU bound to a pseudo-self-complementary 17-mer duplex with three mismatched T:T and four unpaired T (Figure 1.0). Despite the peculiar substrate, the structure has provided insight into the general structure of HU-DNA complexes. For example, in the crystal structure the β-ribbon arms from each monomer follow the minor groove of the DNA. Two conserved proline residues at the tip of each arm insert between DNA base pairs. There are 9 basepairs (bp) between the intercalating prolines. The prolines kink the 5′-terminal and 3′-terminal flanking regions towards the α-helical body region of HU (Figure 6-1). In the crystal structures the calculated HU-induced bend angle ranges from 105 to 139°.

Additional structural information is gained from HU-DNA footprinting data. Three site-specific cysteine HU mutants were chemically converted into nucleases and individually complexed with cruciform and nicked DNA to generate cleavage patterns indicative of protein binding sites. The footprinting data suggests that the HUβ arm interacts with the
Figure 6-1 HU crystal structure

The crystal structure for homodimer HUαα from cyanobacterium Anabaena (PDB ID = 1P71) bound to a pseudo-self-complementary double-stranded DNA molecule. The crystal structure reveals the two intercalating prolines (dark blue) stack with unpaired thymines (green). The spheres represent the position of the HU site specific mutations using the following colors, green for α-HU Q43C, yellow for α-HU A78C, magenta for α-HU K90C, red for β-HU Q64C, and orange for β-HU T70C.
bases at the ss/ds junction (referred to herein as junction), while the HUα arm follows the minor groove 3’ of the nick on the continuous strand. Obtaining this data required first isolating HU-DNA complexes from a native gel and activating HU nuclease activity to generate cleavage patterns. This method could be problematic if the in-gel and solution complexes differ. Also, only three site-specific mutants were used in this study to generate a model of the complex. A larger data set is preferred for constructing an accurate model. Concerns with how the crystal structure translates to biologically relevant ligands and the limitations of the footprinting model prompt further investigation of HU’s specific binding mode.

HU binds with high affinity to an array of biologically relevant nucleic acid structures. One such construct is a DNA overhang, composed of two DNA strands where one is longer than the other. Compared to nicked, gapped, and cruciform DNA which require at a minimum 3 oligomers for assembly the overhang is a relatively simpler construct, useful as a model system for investigating HU’s structurally-specific binding mode. In addition, the overhang is an important intermediate in several biological processes that require HU. For example, the overhang motif is an important intermediate during replication, priming repair of the double-stranded breaks, and priA replication. It has been reported that HU exhibits binding polarity, favoring the 3’ overhang (single-strand portion is 3’ from the ss/ds junction) over the 5’ overhang. This binding polarity may extend to all high affinity structures, such that HU binds in a particular orientation. Overhang constructs are good systems for exploring HU binding polarity because of their inherent directionality, the 5’ end differs from the 3’ end. In addition, as a model system it is possible to apply the HU-overhang binding orientation to more complex systems such as nicked, gapped, and branched DNA-HU complexes.
It is known that DNA conformational dynamics are an important factor in HU specificity. However, the structural elements that lead to specificity in the HU-DNA binding interaction have yet to be determined. An important step forward is to identify the solution state geometry of the HU-overhang complex and understand the role of nucleic acid topology in HU specificity. The objective of this study is to improve the model of HU binding to DNA overhangs. A combination of techniques are employed to this end, with an emphasis on using the fluorescent base analog 6-methylisoxanthopterin (6-MI) to examine local HU-induced perturbation to the DNA structure and fluorescence resonance energy transfer to investigate the global distortion of the DNA structure by generating distance constraints for the HU-overhang complex in solution.

This work examines DNA conformational dynamics at single base resolution. We achieve this level of resolution by incorporating the fluorescent base analog 6-MI at specific positions within an oligomer and measuring its dynamics with time correlated single photon counting (TCSPC). The measurement of the excited state lifetime decay of 6-MI, yields information on the local 6-MI environment and heterogeneity of the sample. This study uses a specific pentamer sequence ATFAA (where F represent 6-MI) in a 34 bp oligomer (Table 6-1.0). The relative quantum yield of 6-MI in the ATFAA sequence increases from 0.2 to 0.85 upon duplex formation. This results from 6-MI adopting a constrained conformation in ATFAA dsDNA, characterized by a single excited state decay compared to the multi-exponential decays observed for ssDNA ATFAA.

Time-resolved fluorescence experiments can also monitor the reorientation of the emission dipole during the excited-state lifetime to yield an anisotropy decay, which reflects both local motion ($\theta_L$) and the global tumbling ($\theta_R$) of the oligomer. The excited state decay of 6-MI in ATFAA dsDNA indicates a single conformation within the duplex. The
rotational correlation time $\theta_r$ for this unique conformation is $\sim 1.5$ ns, longer than the rotational correlation time of 6-MI in ssDNA and in other dsDNA sequences. The relatively long rotational correlation time is consistent with a constrained 6-MI conformation as predicted from the lifetime data, suggesting that there is limited dynamic quenching with adjacent bases. The TCSPC data reports on the rigidity and polarity of the microenvironment surrounding 6-MI at a specific base site. A method to map HU-induced distortions to the DNA at a single base level requires incorporation of 6-MI at different positions within the overhang construct and measuring its conformational dynamics.

The role of DNA structure and dynamics in HU’s structurally-specific binding mode as well as the extent to which HU-induced distortions of DNA structure influences affinity such as in the overhang construct remains a topic for study. Fluorescence resonance energy transfer (FRET) spectroscopic methods have successfully investigated DNA topology and protein-induced distortions. This technique measures the extent of nonradiative energy transfer between a “donor” and “acceptor” molecule and is often described as a “molecular ruler”, since it is straightforward to mathematically convert energy transfer into the distance between donor and acceptor molecules. We used the donor fluorescein (FAM) and acceptor tetramethylrhodamine (TAMRA) to label DNA constructs at their ends (Table 6-1). Measurement of the distance dependent energy transfer between dyes enables prediction of DNA geometry. In this case, 6-MI acts as the donor fluorophore and its placement at unique positions relative to the junction generates a series of donor constructs. We labeled six site-specific HU mutants with FAM to act as the 6-MI acceptor (Figure 6-1). Distances between donors and acceptors were measured using FRET. These experimentally obtained intermolecular FRET determined distances and structural restraints informed our model
building of HU-DNA complexes.\textsuperscript{311-313} FRET distances were compared to theoretical
distances as a means for evaluating a set of molecular models.

This work uses a DNA sequence based upon the H1 binding site of integration host factor
(IHF), a structural homolog of HU to generate all overhang constructs (Table 6-1). We
choose to use this sequence for several reasons. First, in the presence of the recombinase
protein integrase (Int), HU is known to bind this sequence with high affinity.\textsuperscript{314} In addition,
since the IHF-induced bending of the H1 sequence is well characterized, it provides a known
structure for comparison to HU-induced conformational changes. We used a series of 3’ and
5’ overhangs, as well as nicked DNA, to investigate HU specific binding. Single- and
double-stranded forms of the H1 DNA sequence were used for comparison to the structurally-
specific HU binding. HU represents a model system to investigate structure and dynamic
aspects of larger nucleoprotein complexes containing significant bending or wrapping of
DNA around the protein surface, such as the nucleosome.

6.2 Results

6.2.1 HU-DNA stoichiometry

Several studies have investigated the HU to DNA stoichiometric ratio for a number
of different constructs and found the ratio was dependent on binding mode and length of the
DNA construct.\textsuperscript{60,175,191,196} Interestingly, the stoichiometric ratios for the ssDNA-HU and
overhang-HU complexes were not determined. To investigate these stoichiometric ratios, this
work monitored fluorescence anisotropy of 6-MI-containing substrates as a function of HU
concentration.
Table 6-1 HU DNA constructs

<table>
<thead>
<tr>
<th>(A) HU affinity and mapping base microenvironment and dynamics</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFA07 5’ TAT GAT F AA ACT ATG AAT CAA CTA CTT AGA TGTT</td>
</tr>
<tr>
<td>TFA15 5’ TAT GCA GTC ACT AT F AAT CAA CTA CTT AGA TGTT</td>
</tr>
<tr>
<td>(B) FRET construct</td>
</tr>
<tr>
<td>H34ds 5’ FAM-TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT</td>
</tr>
<tr>
<td>3’ TAMRA-ATA CGT CAG TGA TAC TTA GAT GAA TCT ACCA-TAMRA</td>
</tr>
<tr>
<td>3prime_TOTAL 5’ TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT-FAM</td>
</tr>
<tr>
<td>3’ TAMRA-ATA CGT CAG TGA TAC TTA-TAMRA</td>
</tr>
<tr>
<td>3prime_DS 5’ FAM-TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT</td>
</tr>
<tr>
<td>3’ TAMRA-ATA CGT CAG TGA TAC TTA-TAMRA</td>
</tr>
<tr>
<td>3prime_SS 5’ TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT</td>
</tr>
<tr>
<td>3’ TAMRA-ATA CGT CAG TGA TAC TTA-TAMRA</td>
</tr>
<tr>
<td>5prime_TOTAL 5’ FAM-TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT-FAM</td>
</tr>
<tr>
<td>3’ TAMRA-ATA GTT GAA TCT ACCA-TAMRA</td>
</tr>
<tr>
<td>5prime_DS 5’ TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT</td>
</tr>
<tr>
<td>3’ TAMRA-ATA GTT GAA TCT ACCA</td>
</tr>
<tr>
<td>5prime_SS 5’ FAM-TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT</td>
</tr>
<tr>
<td>3’ TAMRA-ATA GTT GAA TCT ACCA</td>
</tr>
<tr>
<td>3prime_FAM-nic_total 5’ TAT GCA GTC ACT ATG A AT CAA CTA CTT AGA TGTT-FAM</td>
</tr>
<tr>
<td>3’ TAMRA-ATA CGT CAG TGA TAC T TA GTT GAA TCT ACCA-FAM</td>
</tr>
<tr>
<td>5prime_FAM-nic_total 5’ FAM-TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT</td>
</tr>
<tr>
<td>3’ TAMRA-ATA GTT GAA TCT ACCA</td>
</tr>
<tr>
<td>3prime_FAM-nic_DS 5’ TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT-FAM</td>
</tr>
<tr>
<td>3’ TAMRA-ATA GTT GAA TCT ACCA</td>
</tr>
<tr>
<td>5prime_FAM-nic_DS 5’ FAM-TAT GCA GTC ACT ATG AAT CAA CTA CTT AGA TGTT</td>
</tr>
<tr>
<td>3’ TAMRA-ATA GTT GAA TCT ACCA</td>
</tr>
</tbody>
</table>

Table 6-1(A) We used the first two sequences to characterize HU affinity and map HU-induced perturbation to base microenvironment and dynamics. A series of complementary strands at different lengths altered the position of 6-MI relative to the junction. The red (F) indicates the position of 6-MI. Nomenclature is first three letters represent 6-MI (F) and adjacent bases followed by number for position in oligo. This is followed by underscore ( _ ), number to represent length of complement, and the number after ( bp ) represents position of complement either 5 of 3 prime. B) FRET constructs to examine conformation of constructs in the presence and absence of protein. The FAM represents position of donor fluorescein, the TAMRA represents the position of acceptor Carboxytetramethylrhodamine, and the yellow highlighted ( / ) is the location of the nick in the DNA.
### (C) FRET DNA to protein constructs

<table>
<thead>
<tr>
<th>Construct</th>
<th>5' Sequence</th>
<th>3' Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFA15_10bp5</td>
<td>TAT GCA GTC ACT AT F AAT CAA CTA CTT AGA TGGT</td>
<td>ATA CGT CAG T</td>
</tr>
<tr>
<td>TFA15_15bp5</td>
<td>TAT GCA GTC ACT AT F AAT CAA CTA CTT AGA TGGT</td>
<td>ATA CGT CAG TGA TAC</td>
</tr>
<tr>
<td>TFA15_18bp5</td>
<td>TAT GCA GTC ACT AT F AAT CAA CTA CTT AGA TGGT</td>
<td>ATA CGT CAG TGA TAC TTA</td>
</tr>
<tr>
<td>TFA07_18bp5</td>
<td>TAT GAT FAA ACT ATG AAT CAA CTA CTT AGA TGGT</td>
<td>ATA CTA CTT TGA TAC TTA</td>
</tr>
<tr>
<td>TFA15_23bp3</td>
<td>TAT GCA GTC ACT AT F AAT CAA CTA CTT AGA TGGT</td>
<td>A TAC TTA GTT GAT GAA TCT ACCA</td>
</tr>
<tr>
<td>TFA15_20bp3</td>
<td>TAT GCA GTC ACT AT F AAT CAA CTA CTT AGA TGGT</td>
<td>C TTA GTT GAT GAA TCT ACCA</td>
</tr>
<tr>
<td>TFA07_19bp3</td>
<td>TAT GAT FAA ACT ATG AAT CAA CTA CTT AGA TGGT</td>
<td>TA GTT GAT GAA TCT ACCA</td>
</tr>
<tr>
<td>TFA07_12bp3</td>
<td>TAT GAT FAA ACT ATG AAT CAA CTA CTT AGA TGGT</td>
<td>AT GAA TCT ACCA</td>
</tr>
</tbody>
</table>

Table 6-1(C): The DNA constructs contain 6-MI (F) at either the 7th or 15th position. To alter the position of 6-MI relative to the junction, we varied the length and position (either 5’ or 3’) of the complement strand. These constructs were used to measure the distance dependent FRET between 6-MI and HU mutants labeled with FAM. The distance constraints were then used to evaluate models of HU binding to overhangs. Nomenclature follows the method outlined for constructs in (A)
Fluorescence anisotropy directly measures the formation of protein-DNA complexes as protein binding causes a decrease in rotational diffusion of the DNA.\textsuperscript{214,315} A strength of this technique is the ability to measure both weak and strong interactions between protein and DNA. Solution-based techniques are preferred to the more conventional electrophoresis or GMSA, which can slow protein-nucleic acid dissociation in the gel skewing the analysis.\textsuperscript{316} In addition, incorporation of the fluorescent base analog 6-MI into the DNA sequence means anisotropy values correlate directly to the global DNA motion.\textsuperscript{317} The fluorescence anisotropy of 6-MI incorporated into a DNA oligomer depends on the size and shape of the oligomer. Therefore, HU binding leads to an increase in molecular size and shape, which is associated with an increase in fluorescence anisotropy.

The HU to DNA stoichiometric ratio was determined by titrating HU into a solution containing either the low affinity single-strand TFA15 ss or the high affinity overhang at a concentration 10-fold greater than the $K_D$. At this high concentration the added HU will complex with DNA and plotting the anisotropy as a function of the HU:DNA molar ratio yields an inflection point indicative of the binding stoichiometric ratio.\textsuperscript{315} Based on GMSA data, it was hypothesized that the stoichiometric ratio of the HU structurally-specific binding mode was 1:1.\textsuperscript{181} For the high affinity 3'-overhang TFA15_18bp5 an inflection point occurs at a stoichiometric ratio of 1 HU to 1 TFA15_18bp5 overhang (Figure 6-2). In addition, 1:1 stoichiometric ratios were also calculated for the high affinity 5’ overhang (TFA15_18bp3) and nicked DNA (TFA15_NIC_18bp3) substrates (Table 6-2). A stoichiometric ratio of 1:1 was also measured using overhang DNA labeled with the external fluorescent probe rhodamine (H34a_18bpaL5’_RHO Table 6-2), confirming that 6-MI does not influence binding stoichiometry.
6-MI anisotropy as a function of HU concentration indicates the stoichiometry is 1:1 between the protein and the high affinity DNA construct TFA\textsubscript{18bp5} and a 3:1 stoichiometric ratio between HU and low affinity single-stranded DNA construct TFA15 ss. HU was titrated into a sample containing a fixed DNA concentration, 6-MI was excited at 340 nm and the anisotropy as a function of HU concentration was recorded at 440 nm. The stoichiometric ratio was determined using the method outlined in Chapter 3.
Table 6-2 HU equilibrium association constants for specific and non-specific binding modes

<table>
<thead>
<tr>
<th>DNA construct</th>
<th>Salt [K] (mM)</th>
<th>$K_A$ (nM)$^{-1}$</th>
<th>Stoichiometric Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFA15_18bp5</td>
<td>60</td>
<td>3.8 ± 1.5</td>
<td>1:1</td>
</tr>
<tr>
<td>TFA15_18bp3</td>
<td>60</td>
<td>4.1 ± 1.5</td>
<td>1:1</td>
</tr>
<tr>
<td>TFA15_23bp3</td>
<td>60</td>
<td>3.8 ± 5.0</td>
<td></td>
</tr>
<tr>
<td>TFA07_10bp5</td>
<td>60</td>
<td>3.1 ± 0.8</td>
<td></td>
</tr>
<tr>
<td>TFA15_NIC_18bp3</td>
<td>60</td>
<td>2.3 ± 1.2</td>
<td>1:1</td>
</tr>
<tr>
<td>H34a_18bpAL5’_RHO</td>
<td>60</td>
<td>3.2 ± 0.7</td>
<td>1:1</td>
</tr>
<tr>
<td>H34a_18bpAL5’_RHO + HU_Q43C</td>
<td>60</td>
<td>2.5 ± 1.0</td>
<td></td>
</tr>
<tr>
<td>H34a_18bpAL5’_RHO + HU_K90C</td>
<td>60</td>
<td>2.7 ± 1.8</td>
<td></td>
</tr>
<tr>
<td>H34a_18bpAL5’_RHO + HU_A78C</td>
<td>60</td>
<td>2.4 ± 1.8</td>
<td></td>
</tr>
<tr>
<td>H34a_18bpAL5’_RHO + HU_Q64C</td>
<td>60</td>
<td>1.7 ± 1.0</td>
<td></td>
</tr>
<tr>
<td>H34a_18bpAL5’_RHO + HU_T70C</td>
<td>60</td>
<td>1.7 ± 1.0</td>
<td></td>
</tr>
<tr>
<td>TFA15 ss</td>
<td>50</td>
<td>3.5 ± 0.8</td>
<td>3:1</td>
</tr>
<tr>
<td>TFA15 ss</td>
<td>200</td>
<td>0.5 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>TFA15 ds</td>
<td>50</td>
<td>1.8 ± 0.2</td>
<td>3:1</td>
</tr>
<tr>
<td>TFA15 ds</td>
<td>200</td>
<td>0.2 ± 0.1</td>
<td></td>
</tr>
</tbody>
</table>

Table 6-2 Above is shown the equilibrium binding parameters for the two modes of HU binding. HU was titrated into 6-MI incorporated or TAMRA 5-prime labeled DNA at a fixed concentration of 300 pM to measure the specific binding association constants (rows 1-11). Based on the 1:1 stoichiometric ratio, a single-site non-interacting model was used to analyze the plot of intensity vs. HU concentration. The last four rows show the macroscopic ($k$) association constant and the three corresponding microscopic binding constants ($K_1, K_2, K_3$) for the non-specific binding mode at two salt concentrations 50 and 200 mM NaCl. The stoichiometric ratio for HU non-specific binding to 34-mer single-strand (TFA15 ss) and 34-mer double-strand (TFA15 ds) is 3:1.
HU was shown to require a minimum length of 24 nt to bind ssDNA, while an additional 12 nt is required for each additional HU; under low stringency conditions (10 mM NaCl), the stoichiometric ratio is 2:1 for HU binding to ssDNA of 28 to 36 nucleotides (nt) in length.\(^{181}\) However, the inflection point determined by anisotropy in the plot of HU binding to ssDNA (TFA15ss) occurs at approximately 3 HU to 1 single-strand DNA (Figure 6-2), indicating a stoichiometric ratio of 3:1 and suggests a binding site size of ~11 bases. Previously measured stoichiometry ratios were based on GMSA data, which has been shown to depend on both protein-DNA ratio and DNA compaction adding additional complexity to data interpretation.\(^{60}\) Interestingly, the stoichiometric ratio between HU and the H1 34-mer dsDNA was measured to be 3:1, consistent with a binding site size of ~11 bp.\(^{60}\) The similarities between the two non-specific binding constructs highlight the dependence of binding site size on just polymer length and not structure, while the 1:1 ratio for the structurally-specific mode suggests HU binding occludes binding of additional proteins.

6.2.2  *HU affinity is dependent on DNA structure*

The structurally-specific binding mode of HU was reported to exhibit preferential binding to 3' overhang with an apparent association constant of 63 μM\(^{-1}\) compared to 1.0 μM\(^{-1}\) for the 5' overhang.\(^{103}\) Interestingly, in a RNA-DNA hybrid overhang the binding polarity is absent and the association constant for both 5 and 3' overhangs is ~ 0.063 nM\(^{-1}\).\(^{80}\) For this reason, it was hypothesized that the observed binding polarity was an artifact of the gel matrix. To investigate this possibility, HU binding to the 5' and 3' overhang was examined in solution monitoring fluorescence intensity as a function of protein concentration. 6-MI intensity is sensitive to stacking interactions with adjacent bases making it an excellent reporter of protein-induced perturbation to DNA conformation.\(^{55}\) Importantly, HU binding is
Figure 6-3 HU binding affinity for 5' and 3' overhang

Plotting the intensity change as a function of HU concentration reveals protein affinity is similar for both the 5' overhang and 3' overhang. HU was titrated into a sample containing either 300 pM 3-prime overhang (TFA15_18bp5, black squares) or 5-prime overhang (TFA15_18bp3, white circles). The binding of HU to the structural specific overhang was modeled using 1:1 binding as given in Section 3.6 equation 3.9.
known to distort DNA conformation, which will change 6-MI local environment leading to a change in fluorescent intensity.

For the 3’ overhang TFA15_18bp5 an increase in fluorescence intensity was observed as a function of HU concentration and a saturable binding curve was obtained (Figure 6-3, Table 6-2). A single site binding model was used for analysis yielding an association constant of $3.8 \pm 1.5 \text{ nM}^{-1}$, which is 2 orders of magnitude tighter than the previously reported value of $0.063 \text{ nM}^{-1}$. One possible source for the discrepancy in the measured affinities is that GMSA experiments used 2 nM of overhang DNA, which is well above the $0.3 \text{ nM} K_D^{103}$. This study used 0.3 nM for the solution experiments. In addition, we also detected binding between HU and the 5’ overhang, consistent with the absence of an observed HU binding preference for the RNA/DNA hybrid overhang. HU affinity for the 5’ overhang is $4.1 \pm 1.5 \text{ nM}^{-1}$, comparable to the 3’ overhang $K_A$ but 60-fold tighter than the detected affinity between HU and the RNA/DNA hybrid (Figure 6-3, Table 6-2). Previous studies have shown HU and the homolog IHF to bind DNA constructs with similar affinity demonstrating the observed values are reasonable. First, the reported association constant for HU binding to a series of crystal structure related constructs range from $250 \text{ nM}^{-1}$ to $0.16 \text{ nM}^{-1}$ comparable to the overhang values. In addition, Sugimura and Crothers have reported a $K_D \approx 0.025 \text{ nM}$ for the IHF-H’ complex obtained with direct stop-flow measurements of the association and dissociation rate constants, consistent with the reported values for HU specific binding.$^{318}$

In this work a series of overhang constructs were used to characterize the HU-DNA interaction. For each construct the overall length of oligomer was the same (34 bases) while the length of the double-stranded DNA portion was varied. HU binding to the shortest and
The above plots (A-C) show the intensity change of 6-MI labeled DNA as a function of HU concentration. The similar binding curves for the overhang constructs with shortest duplex portion TFA07_10bp5' (A) and with the longest duplex portion TFA15_23bp3' (B) demonstrate the duplex length used in this study does not affect HU affinity. The plot for construct containing a nick at the 18th bp (C) indicates HU has similar affinity for the DNA with reduced conformational flexibility. The binding of HU to the structural specific overhang was modeled using 1:1 binding as given in Section 3.6 equation 3.9.
longest duplex lengths was investigated by fluorescence intensity as a function of HU concentration to determine if duplex length affected affinity. The observed binding curves were analyzed using a single site binding model and association constants of $3.1 \pm 0.8$ and $3.8 \pm 5.0 \text{nM}^{-1}$ were determined for the 10 and 23 bp overhang. This finding confirms that HU has high affinity for both overhang constructs regardless of duplex length (Figure 6-4, Table 6-2). HU binding to nicked DNA (containing a single-strand break) was examined in addition to examining binding to an array of overhangs. Previous studies had indicated HU binding affinity for nicked DNA is $0.125 \text{nM}^{-1}$. Consistent with the overhang intensity data, HU binding affinity for the nicked DNA calculated using a single site model is $2.3 \pm 1.2 \text{nM}^{-1}$, an order of magnitude tighter than reported (Figure 6-4, Table 6-2).

A general concern with using any fluorescent base analog (FBA) is the possibility of influencing the binding interaction; in this case HU affinity for 6-MI containing overhang constructs. To rule out this possibility, we measured the binding affinity of an externally-labelled construct 3prime_18bp_RHO in which the fluorophore Rhodamine was attached to the 5’ end of the duplex (Figure 6-5, Table 6-2). For this construct, the fluorescence intensity increased as a function of HU concentration and saturable binding was observed. Analysis of the binding curve using a single binding site model yields an association constant of $3.2 \pm 0.7 \text{nM}^{-1}$, in good agreement with 6-MI-determined affinities.

**HU affinity for single-stranded DNA is dependent on [NaCl]**

Previous reports indicate that HU binds single-stranded DNA non-cooperatively with low affinity in the gel. The authors note under low stringency conditions (50 mM NaCl), HU affinity for ssDNA is ~3X weaker than it is for dsDNA. Given the structural nature of the
Figure 6-5 Binding affinity of HU mutants

The plots (A-C) show the change in TAMRA fluorescent intensity for an overhang containing the external dye rhodamine (3prime_18bp_RHO) versus [HU]. The plots are with wild type HU (HU_WT = A), mutant Q43C (HU_Q43C = B), and mutant Q64C (HU_Q64C = C). Using the above plots to calculate the association constant for WT_HU, mutants HU_Q43C, and HU_Q64C as 3.2 nM$^{-1}$, 2.5 nM$^{-1}$, and 1.7 nM$^{-1}$ respectively confirms labeling protein with TAMRA or labeling DNA with 6-MI does not affect protein affinity. The binding of HU to the structural specific overhang was modeled using 1:1 binding as given in Section 3.6 equation 3.9.
overhang constructs, we were interested in evaluating the affinity of HU for ssDNA in solution. In previous studies, HU affinity for single-stranded DNA was found to be dependent on Na\(^+\) concentration in gel, indicating that the interaction between HU and ssDNA is mainly electrostatic in nature, similar to the HU- dsDNA interaction. \(^1\)

In this work HU was titrated into a solution containing either single- or double-stranded DNA at a fixed concentration in the presence of either 50 or 200 mM NaCl. Figure 6 depicts 6-MI anisotropy as a function of HU concentration. In the absence of HU, the increased flexibility of the single-stranded DNA is reflected in a lower initial anisotropy of 0.11 relative to 0.18 for double stranded DNA.

For single- and double-stranded DNA, protein binding leads to an increase in anisotropy, which plateaus at a protein concentration greater than 3 μM. As expected, the change in anisotropy as a consequence of protein binding for TFA15 ss (Δr ~ 1.0), is larger than that for TFA15 ds (Δr ~ 0.7) (Figure 6-6). In Figure 6-6 the binding curves are fit to a model of non-interacting binding sites, the number of binding sites was determined independently from stoichiometric measurements (Figure 6-2).\(^2\) In contrast to the literature, under low stringency conditions (50 mM NaCl) the microscopic association constant (\(k\)) for TFA15ss is 3.5 ± 0.78 μM\(^{-1}\), comparable to TFA15ds, which is 1.8 ± 0.16 μM\(^{-1}\) (Table 6-2). This value is in good agreement with the microscopic association constant of 2.1 μM\(^{-1}\) previously determined for HU binding H’ 34-mer dsDNA in solution (a related IHF consensus sequence).\(^3\) The microscopic binding constants are related to the stepwise binding constants, where \(K_1 = 3k\), \(K_2 = k\), and \(K_3 = k/3\) (Table 6-2). We find that the affinity of HU for ssDNA is dependent on salt concentration (Figure 6-6, Table 6-2). Specifically, in the presence of 50 mM NaCl HU affinity for single-stranded DNA is 3.5 ± 0.78 μM\(^{-1}\) and

275
Figure 6-6 HU structural non-specific affinity

The above plots show the anisotropy of 6-MI incorporated into DNA as a function of HU concentration. Plot (A) represents binding under low stringency (50 mM NaCl) to single (TFA15 ss, circle) and double (TFA15 ds, square) strand DNA. The plots reveal HU has similar affinity for single and double-stranded DNA. Plots (B) shows binding under high stringency (200 mM NaCl) for the single (TFA15 ss, circle) and double (TFA15 ds, square) stranded DNA. The 200 mM NaCl plot demonstrates weaker affinity of HU under stringent conditions. The binding of HU to the structural non-specific ss and dsDNA was modeled as 3:1 binding as given in Section 3.6 equation 3.9
increasing the salt concentration to 200 mM NaCl reduces HU affinity ~7-fold (Table 6-2). Interestingly, it does not appear that HU structure-specific binding is strictly dependent on DNA flexibility as HU affinity for single-stranded DNA is expected to be highest, high affinity however, requires either a single-strand break or ss/ds DNA junction.

6.2.3 **HU influences the excited state of 6-MI incorporated within the overhang**

The excited state of 6-MI is sensitive to the local environment and therefore an excellent reporter of perturbations to base conformation. We use this reporter to probe the extent of DNA distortion away from the site of intercalation in the HU-overhang complex and draw comparisons to the HU-induced kinks observed in the co-crystal structure. One measure for assessing this is the amplitude weighted mean lifetime (τ_{mol}), which is proportional to the integrated area of the excited state decay and is used to calculate the relative quantum yield. The value represents the average lifetime of 6-MI (the inverse of 6-MI decay rate) and is a good indicator of the HU-induced perturbations to 6-MI microenvironment. A representative excited state decay and residual plot for 6-MI in the presence and absence of HU are shown in Figure 6-7. In this case HU binding reduces dynamic quenching at earlier times leading to a longer average lifetime.

Based on work with 6-MI in the ATFAA sequence, three excited state populations were presumed for 6-MI in the single strand portion of the overhang; while the excited state decay of 6-MI in the duplex portion would be largely single-exponential with a lifetime of ~6.5 ns similar to monomer 6-MI.\textsuperscript{55} For all excited state decays the $\chi^2$ for the models were in an acceptable range between 1.01 and 1.31. In addition no systematic noise was observed in the residual (Figure 6-7). In the absence of HU the mean lifetime of 6-MI increases as its
The plot (A) shows the excited state decay of 6-MI incorporated into an overhang in the presence (red) and absence (blue) of HU. The excited state decay is multi-exponential and in the presence of HU the lifetime of 6-MI increases. The residual of the model is shown below, systematic noise in the residual is indicative of a poor model.

Figure 6-7 6-MI overhang excited state decay
Table 6-3 Mapping HU-induced perturbations to base conformation using the position dependent excited state decay of 6-MI within the overhang construct

<table>
<thead>
<tr>
<th>DNA</th>
<th>DNA + HU</th>
</tr>
</thead>
<tbody>
<tr>
<td>5' OH, SS Bases 6-MI to Junc.</td>
<td>5' OH, DS bp 6-MI to Junc.</td>
</tr>
<tr>
<td>α(1) ± 0.006</td>
<td>α(1) ± 0.005</td>
</tr>
<tr>
<td>0</td>
<td>0.20</td>
</tr>
<tr>
<td>1</td>
<td>0.35</td>
</tr>
<tr>
<td>2</td>
<td>0.43</td>
</tr>
<tr>
<td>4</td>
<td>0.36</td>
</tr>
<tr>
<td>5</td>
<td>0.43</td>
</tr>
<tr>
<td>6</td>
<td>0.46</td>
</tr>
<tr>
<td>7</td>
<td>0.52</td>
</tr>
<tr>
<td>8</td>
<td>0.56</td>
</tr>
<tr>
<td>9</td>
<td>0.60</td>
</tr>
<tr>
<td>10</td>
<td>0.62</td>
</tr>
<tr>
<td>12</td>
<td>0.55</td>
</tr>
<tr>
<td>13</td>
<td>0.54</td>
</tr>
<tr>
<td>14</td>
<td>0.54</td>
</tr>
<tr>
<td>16</td>
<td>0.53</td>
</tr>
<tr>
<td>18</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Table 6-3A The excited state decay of 6-MI in the presence and absence of HU is dependent on its position relative to the junction in the overhang. The data is separated by duplex (DS) and single-strand (SS) portions of the 5-prime overhang (5' OH.)
Table 6-3-A Mapping HU-induced perturbations to base conformation using the position dependent excited state decay of 6-MI within the overhang construct

<table>
<thead>
<tr>
<th>3' OH. SS Bases 6-MI to Junc.</th>
<th>DNA</th>
<th>DNA + HU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha_1$ (±)</td>
<td>$\tau_1$ (ns) ±</td>
</tr>
<tr>
<td>0</td>
<td>0.11 ± 0.01</td>
<td>0.03 ± 0.004</td>
</tr>
<tr>
<td>1</td>
<td>0.21 ± 0.01</td>
<td>0.13 ± 0.005</td>
</tr>
<tr>
<td>3</td>
<td>0.40 ± 0.01</td>
<td>0.15 ± 0.005</td>
</tr>
<tr>
<td>5</td>
<td>0.38 ± 0.01</td>
<td>0.18 ± 0.005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3' OH. DS Bp 6-MI to Junc.</th>
<th>DNA</th>
<th>DNA + HU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha_1$ (±)</td>
<td>$\tau_1$ (ns) ±</td>
</tr>
<tr>
<td>0</td>
<td>0.11 ± 0.00</td>
<td>0.03 ± 0.001</td>
</tr>
<tr>
<td>1</td>
<td>0.00 ± 0.00</td>
<td>0.06 ± 0.001</td>
</tr>
<tr>
<td>3</td>
<td>0.00 ± 0.00</td>
<td>0.02 ± 0.001</td>
</tr>
<tr>
<td>5</td>
<td>0.00 ± 0.00</td>
<td>0.04 ± 0.001</td>
</tr>
<tr>
<td>7</td>
<td>0.00 ± 0.00</td>
<td>0.03 ± 0.001</td>
</tr>
<tr>
<td>9</td>
<td>0.00 ± 0.00</td>
<td>0.05 ± 0.001</td>
</tr>
</tbody>
</table>

Table 6-3-A The excited state decay of 6-MI in the presence and absence of HU is dependent on its position relative to the junction in the overhang. The data is separated by duplex (DS) and single-strand (SS) portions of the 3-prime overhang (3’ OH.)
(A) 
**Right:** A plot of 6-MI $\tau_{\text{mol}}$ versus its position in the duplex portion of the 3’ overhang. As the position of 6-MI is moved from the junction and into the duplex portion $\tau_{\text{mol}}$ increases to ~6.5 ns, consistent with a restrained conformation in the duplex. **Left:** the change in $\tau_{\text{mol}}$ upon HU binding versus 6-MI position in the duplex portion. Upon HU binding $\tau_{\text{mol}}$ for 6-MI at the junction increases by ~0.9 ns, indicating a reduction in phase-space sampling for 6-MI as a consequence of protein binding. In addition, $\tau_{\text{mol}}$ for 6-MI positioned 1 to 7 from the junction in the duplex portion decreases by ~0.30 ns with HU binding. The decrease in $\tau_{\text{mol}}$ is an indicator that HU relaxes the restraint on 6-MI conformation in the duplex portion. All $\tau_{\text{mol}}$ values are given in Table 6-3 and are calculated from the lifetime data as indicated in Chapter 3.

(B) 
**Right:** A plot of 6-MI $\tau_{\text{mol}}$ versus its position in the 3’ overhang. As the position of 6-MI is moved from the junction and into the single-strand portion of the 3’ overhang, plateauing at approximately 2.1 ns. **Left:** Percentage change in $\tau_{\text{mol}}$ upon HU binding as a function of distance from the junction. The molecular weighted lifetime of 6-MI positioned at the junction increases ~25% upon HU binding, which is attributed to increased stability of the base as a consequence of protein binding. All $\tau_{\text{mol}}$ values are given in Table 6-3 and are calculated from the lifetime data as indicated in Chapter 3.

Figure 6-8  $\tau_{\text{mol}}$ for 6-MI in the 3’ overhang
Right: A plot of 6-MI $\tau_{\text{mol}}$ versus its position in the duplex portion of the 5' overhang. As the position of 6-MI is moved from the junction and into the duplex portion $\tau_{\text{mol}}$ increases to ~6.5 ns, consistent with 3' overhang data representing a restrained conformation of 6-MI in the duplex.

Left: The change in $\tau_{\text{mol}}$ upon HU binding versus 6-MI position in the duplex portion. Upon HU binding $\tau_{\text{mol}}$ for 6-MI at the junction increases by ~1.0 ns, indicating a reduction in phase-space sampling for 6-MI as a consequence of protein binding. In addition, $\tau_{\text{mol}}$ for 6-MI positioned adjacent to the junction in the duplex portion decreases by ~0.53 ns with HU binding. The decrease in $\tau_{\text{mol}}$ for 6-MI is an indicator that HU relaxes the restraint on 6-MI conformation in the duplex portion. All $\tau_{\text{mol}}$ values are given in Table 6-3 and are calculated from the lifetime data as indicated in Chapter 3.

(B) Right: A plot of 6-MI $\tau_{\text{mol}}$ versus its position in the single-strand portion of the 5' overhang. As the position of 6-MI is moved from the duplex and into the single-strand portion $\tau_{\text{mol}}$ decreases to ~2.0 ns. The shorter $\tau_{\text{mol}}$ is an indicator of greater phase-space sampling compared to 6-MI in the duplex portion. Left: The change in $\tau_{\text{mol}}$ upon HU binding versus 6-MI position in the single-strand portion. Upon HU binding $\tau_{\text{mol}}$ for 6-MI at the junction increases by ~1.0 ns, indicating a reduction in 6-MI conformational sampling as a consequence of protein binding. In addition, $\tau_{\text{mol}}$ for 6-MI positioned 7 to 10 bases from the junction in the single-strand portion increases by ~0.40 ns with HU binding. The increase in $\tau_{\text{mol}}$ suggests HU restraints conformational sampling of these bases. All $\tau_{\text{mol}}$ values are given in Table 6-3 and are calculated from the lifetime data as indicated in Chapter 3.
position moves from the single-strand and into the double-strand portion of the construct (Figure 6-8 and 6-9, Table 6-3-A). The excited state decay data for 6-MI indicates fraying of the DNA at the junction. For example, in the 5’ overhang $\tau_{\text{mol}}$ increases from ~4 ns for 6-MI at the junction to ~ 6.5 ns when located 3 basepairs from the junction (Figure 6-9). The faster lifetime signifies 6-MI is dynamic at these positions a consequence of DNA fraying. While in the 3’ overhang, $\tau_{\text{mol}}$ increases from ~4 ns for 6-MI positioned at the junction to ~6.5 ns when located 1 bp from the junction (Figure 6-8, Table 6-3-A). The increase in $\tau_{\text{mol}}$ is consistent with the constrained single conformation of 6-MI as reported for dsDNA containing the ATFAA sequence. The discrepancy in base-steps necessary to observe a value for $\tau_{\text{mol}}$ consistent with 6-MI in ordered duplex between overhangs suggests there is greater DNA stability at junction for the 3’overhang. As shown in Table 6-3-A, after the third basepair from the junction and into the duplex portion tau 3($\tau_3$) represents ~90% of 6-MI’s excited state population, evidence of one population or one conformation (Table 6-3-A). When we position 6-MI at the junction there is greater complexity in the fluorescence decay, three populations are required for an adequate description, indicating a greater number of conformations or populations of the 6-MI compared to the duplex portion. Moving the position of 6-MI from the junction and into the single-strand portion we observe a shift in the population distribution (Table 6-3-A). The excited state data suggests bases adjacent to the junction in the single-strand portion maintain some conformational rigidity compared to bases located further in the single-strand portion. For example, in the 5’ overhang, moving 6-MI from the junction and into the single-strand portion results in a gradual decrease in $\tau_{\text{mol}}$, plateauing 6 bases from the junction (Figure 6-8). The pattern is similar for the 5’ overhang (Figure 6-9). The gradual decrease in $\tau_{\text{mol}}$, reflects the shift in fractional populations from ~55% tau 2 for 6-MI positioned at the junction to ~60 % tau 1 for 6-MI located 9 basepairs
from the junction (Table 6-3-A). The mean lifetime for 6-MI positioned further than 5 bases from the junction in the single-strand portion is consistent with previously reported values for the pentamer ATFAA in single-strand DNA (Table 6-3-A), which suggests bases before the 5th position maintain some conformational rigidity. In summary, the composition of excited state decay for 6-MI positioned at or around the junction is in between the duplex and single-strand portion. The unique conformational heterogeneity is possibly an important factor dictating HU specificity.

The co-crystal structure and footprinting data suggests one arm of HU interacts with the junction, while the other arm binds 9 bases away, and the body of HU interacts with the portions of the overhang adjacent to the sites of proline intercalation (Figure 6-1). In the crystal structure the intercalated Pro appears to stack with the adjacent bases. Based on this data we expected that the greatest HU-induced change to 6-MI local environment would occur when the location of 6-MI is at the junction or 9 bases away from the junction.

When 6-MI is located at the junction the mean lifetime increases by 25-30 % for both overhangs with HU binding, consistent with a reduction in dynamic quenching and stabilization of the base (Figures 6-8 and 6-9). Thus, HU binding appears to limit the conformational space sampled by the base located at the junction. Interestingly, when 6-MI is located adjacent to the junction in the duplex portion \( \tau_{\text{mol}} \) decreases by 200 ± 50 to 500 ± 50 ps, indicative of 6-MI sampling more conformational space (Figure 6-8 and 6-9, Table 6-3-A) and suggestive of protein binding leading to a structural distortion propagated close to the junction site. The effect of HU binding on 6-MI excited state is attenuated as the position of 6-MI is moved further into the double-strand portion of the 3’ overhang (Figure 6-8 and 6-9, Table 6-3-A). As 6-MI is positioned further from the junction in the duplex region of the 3’ overhang, \( \tau_{\text{mol}} \) approaches those obtained in the absence of protein. For example, when
the position is 8 to 11 basepairs from the junction in the duplex portion there is no
perturbation to the excited state decay an indication that protein binding causes minimal
change to the DNA structure at these positions (Figure 6-8).

The model generated from HU footprinting data suggests the body of HU interacts
through electrostatics with the DNA backbone regions 5’ and 3’ from the points where the
Pro residues intercalate into the DNA. Assuming the single-strand portion is either 5’ or 3’ of
these points HU binding would have an effect on 6-MI microenvironment in this region. HU
binding causes an increase in $\tau_{(mol)}$ for 6-MI located 6 to 9 bases from the junction in the
single-strand portion of the 5’ overhang, reaching a maximum change of 400 ps at the 9th
position form the junction (Figure 6-9). The increase in $\tau_{(mol)}$ reflects a shift in the fractional
populations consistent with a reduction in conformational dynamics. While the change of 400
ps for the 9th base in the single-strand portion is smaller compared to the ~1 ns change for 6-
MI at the junction it is well above the error of ~50 ps for the time-resolved measurements.

Similar to the behavior observed in the duplex portion, HU binding reduces the mean lifetime
of 6-MI located adjacent to the junction in the single-strand portion of the 3’ overhang by 120
ps (Figure 6-8). When 6-MI is positioned 12 bases from the junction in the 5’ overhang, HU
binding reduces $\tau_{(mol)}$ by ~100ps (Figure 6-9). In addition, HU binding causes an increase in
$\tau_{(mol)}$ for 6-MI located past the 12th position with a maximum change of 170 ± 40 ps at the 18th
base from the junction (Figure 6-9).

In summary, HU stabilizes the base at the junction for both 5’ and 3’ overhang an indicator
of proline intercalation and stacking with the base. HU binding destabilizes the bases
adjacent to the junction propagating away into duplex portion of the 3’ overhang which is
suggestive of a kink in the DNA. There is also an increase in $\tau_{(mol)}$ for 6-MI positioned in the
single-strand portion of the 5’ overhang consistent with a reduction in base conformation.
6.2.4 *HU binding alters the dynamics of 6-MI incorporated within the overhang*

The local base rigidity in the overhang was mapped using the anisotropic decay of 6-MI at multiple positions relative to the junction. A representative anisotropic decay is shown in Figure 6-10, the decay is for 6-MI in the TFA15_15bp5 construct in the presence (red decay) and absence (blue decay) of HU. In this construct 6-MI is located at the ss-ds junction. The anisotropic decay is modeled with two rotational correlation times. The fast decay component $\theta_L$ represents the local motion of 6-MI, while the long decay component $\theta_R$ represents the DNA molecular tumbling (Table 6-3-B). HU binding increases $\theta_L$, attributed to a stabilization of 6-MI and a reduction in conformational dynamics. The mass and size of the HU-DNA complex is significantly larger than the DNA construct alone, leading to slower molecular tumbling and longer $\theta_R$ (Table 6-3-B).

The short rotational correlation time measuring local motion of 6-MI ($\theta_L$) reveals base conformational dynamics are dependent on position in the overhang DNA construct (Table 6-3-B). At the junction, $\theta_L$ of 6-MI is 0.51 ns indicative of considerable local motion compared to bases in the duplex region (Figures 6-11 and 6-12, Table 6-3-B). Further, the bases adjacent to the junction exhibit greater dynamics consistent with excited state decay, which suggested fraying of the duplex at the junction.
The above plot represents anisotropic decay incorporated into an overhang in the presence (red) and absence (blue) of HU. The anisotropic decay is modeled with 2 components, in the presence of HU, the fast decay component becomes longer indicative of reduced 6-MI motion. The long decay component, which represents DNA tumbling, also increases.

Figure 6-10 6-MI TR-anisotropy in overhang
Table 6-3.B Mapping HU-induced perturbations to base dynamics using the position dependent rotational correlation times of 6-MI within the overhang.

<table>
<thead>
<tr>
<th>5' OH. SS Bases</th>
<th>DNA</th>
<th>DNA + HU</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-MI to Junct.</td>
<td>β₁ ± 0.03</td>
<td>β₁ ± 0.02</td>
</tr>
<tr>
<td></td>
<td>0.15 0.51</td>
<td>0.10 1.31</td>
</tr>
<tr>
<td></td>
<td>0.22 0.78</td>
<td>0.23 0.79</td>
</tr>
<tr>
<td></td>
<td>0.17 0.83</td>
<td>0.12 0.51</td>
</tr>
<tr>
<td></td>
<td>0.28 0.72</td>
<td>0.11 0.55</td>
</tr>
<tr>
<td></td>
<td>0.28 0.72</td>
<td>0.11 0.59</td>
</tr>
<tr>
<td></td>
<td>0.34 0.66</td>
<td>0.24 0.79</td>
</tr>
<tr>
<td></td>
<td>0.40 0.60</td>
<td>0.37 0.68</td>
</tr>
<tr>
<td></td>
<td>0.30 0.70</td>
<td>0.32 0.88</td>
</tr>
<tr>
<td></td>
<td>0.34 0.66</td>
<td>0.39 0.77</td>
</tr>
<tr>
<td></td>
<td>0.30 0.70</td>
<td>0.23 0.76</td>
</tr>
<tr>
<td></td>
<td>0.29 0.71</td>
<td>0.19 0.53</td>
</tr>
<tr>
<td></td>
<td>0.27 0.73</td>
<td>0.25 0.33</td>
</tr>
<tr>
<td></td>
<td>0.22 0.78</td>
<td>0.18 0.66</td>
</tr>
<tr>
<td></td>
<td>0.24 0.76</td>
<td>0.16 0.66</td>
</tr>
<tr>
<td>5' OH. DS bp 6-MI to Junct.</td>
<td>β₁ ± 0.01</td>
<td>β₁ ± 0.01</td>
</tr>
<tr>
<td></td>
<td>0.15 0.51</td>
<td>0.10 1.31</td>
</tr>
<tr>
<td></td>
<td>0.16 1.12</td>
<td>0.14 0.56</td>
</tr>
<tr>
<td></td>
<td>0.13 1.42</td>
<td>0.11 1.00</td>
</tr>
<tr>
<td>3' OH. SS Bases</td>
<td>β₁ ± 0.06</td>
<td>β₁ ± 0.01</td>
</tr>
<tr>
<td>6-MI to Junct.</td>
<td>0.12 0.48</td>
<td>0.12 1.62</td>
</tr>
<tr>
<td></td>
<td>0.15 0.30</td>
<td>0.11 1.44</td>
</tr>
<tr>
<td></td>
<td>0.34 0.66</td>
<td>0.16 0.48</td>
</tr>
<tr>
<td></td>
<td>0.18 0.17</td>
<td>0.13 0.74</td>
</tr>
<tr>
<td>3' OH. DS bp 6-MI to Junct.</td>
<td>β₁ ± 0.02</td>
<td>β₁ ± 0.01</td>
</tr>
<tr>
<td></td>
<td>0.12 0.48</td>
<td>0.12 1.62</td>
</tr>
<tr>
<td></td>
<td>0.10 0.98</td>
<td>0.14 0.45</td>
</tr>
<tr>
<td></td>
<td>0.11 1.46</td>
<td>0.06 1.03</td>
</tr>
<tr>
<td></td>
<td>0.18 1.50</td>
<td>0.12 1.41</td>
</tr>
<tr>
<td></td>
<td>0.16 1.37</td>
<td>0.13 1.30</td>
</tr>
<tr>
<td></td>
<td>0.12 1.38</td>
<td>0.10 1.41</td>
</tr>
<tr>
<td></td>
<td>0.12 1.40</td>
<td>0.13 1.41</td>
</tr>
</tbody>
</table>

Table 6-3.B The rotational correlation time of 6-MI in the presence and absence of HU is dependent on its position relative to the junction in the overhang. The data is separated by duplex (DS) and single-strand (SS) portions of the 5-prime overhang (5' OH.) and 3-prime overhang (3' OH.)

288
Figure 6-11 (A) The plot of 6-MI local rotational correlation time ($\theta_L$) versus its position in the duplex portion of the 3’ overhang demonstrates a progressive reduction in dynamics as its position moves from the junction and into the duplex portion. Upon HU binding, $\theta_L$ increases by ~1.0 ns when the location of 6-MI is at the junction. The increase demonstrates HU constrains the motion of the base at the junction. In contrast, HU binding reduces $\theta_L$ for 6-MI located 1 to 3 bases from the junction, evidence that HU binding destabilizes the bases adjacent to the junction in the duplex portion. Further than 5 basepairs from the junction HU binding has minimal effect on the dynamics of 6-MI. All $\theta_L$ values are given in Table 6-3 and are calculated from the anisotropic decay as indicated in Chapter 3.

Figure 6-11 6-MI rotational correlation 3’ overhang

(B) The plot of 6-MI local rotational correlation time ($\theta_L$) versus its position in the single-strand portion of the 3’ overhang indicates greater dynamics for bases in the single strand compared to the duplex portion of the overhang. Upon HU binding $\theta_L$ increases by ~1.0 ns for 6-MI positioned at the junction and adjacent to the junction in the single-strand portion of the 3’ overhang. In addition, HU binding increases $\theta_L$ by ~0.5 ns when the location of 6-MI is 5 bases from the junction in the single-strand portion. The increase in the rotational correlation time suggests HU limits the motion of the base at this location. All $\theta_L$ values are given in Table 6-3 and are calculated from the anisotropic decay as indicated in Chapter 3.
Figure 6-12(A) The plot of 6-MI local rotational correlation time ($\theta_L$) versus its position in the duplex portion of the 5' overhang demonstrates a progressive reduction in dynamics as its position moves from the junction and into the duplex portion. Upon HU binding, $\theta_L$ increases by ~0.8 ns when the location of 6-MI is at the junction. The increase demonstrates HU constrains the motion of the base at the junction. In contrast, HU binding reduces $\theta_L$ for 6-MI located 1 to 2 bases from the junction, evidence that HU binding destabilizes the bases adjacent to the junction in the duplex portion. All $\theta_L$ values are given in Table 6-3 and are calculated from the anisotropic decay as indicated in Chapter 3.

Figure 6-12 6-MI rotational correlation 5' overhang

(B) The plot of 6-MI local rotational correlation time ($\theta_L$) versus its position in the single-strand portion of the 5' overhang indicates greater dynamics for bases in the single strand compared to the duplex portion of the overhang. Upon HU binding $\theta_L$ increases by ~0.8 ns for 6-MI positioned at the junction. In addition, HU binding increases $\theta_L$ by ~0.4 ns when the location of 6-MI is 7 to 12b bases from the junction in the single-strand portion. The increase in the rotational correlation time suggests HU limits the motion of the base at this location. All $\theta_L$ values are given in Table 6-3 and are calculated from the anisotropic decay as indicated in Chapter 3.
As the position of 6-MI is moved from the junction and into the double-strand portion of the 3’-overhang, its rotational correlation time (θ_L) increases, consistent with a progressive reduction in base dynamics (Figure 6-11, Table 6-3-B). Similar behavior is also observed in the 5’ overhang, where longer θ_L times are detected, as 6-MI location progresses through the double helix away from the junction (Figure 6-12, Table 6-3-B). In both the 5’ and 3’ overhang constructs, the rotational correlation time (θ_L) increases to ~ 1.5 ns by the third basepair in the duplex away from the junction (Figure 6-11 and 6-12). This longer time likely reflects some rotation of the duplex or segmental motion in addition to local motion. These measurements also show that in the 5’ and 3’ overhangs, base dynamics in the single-strand portion are faster, θ_L ~ 0.3 ns (Figures 6-11 and 6-12, Table 6-3-B). Stabilization of local based dynamics, as measured by a longer θ_L, occurs in part because of base stacking interactions, a reduction in base stacking interactions results in faster local rotational times in ssDNA, which exhibits rotational correlation times ranging from 0.13 to 0.42 ns consistent with this hypothesis.

HU binding to the overhang produces a noticeable change in base conformational dynamics. We expected that the intercalating prolines located in HU’s β-ribbon arms would limit the motion of the stacking base and because of the distortion to the helix, adjacent bases experienced greater dynamics. As suspected, HU binding to the 3’ overhang increases the rotational correlation time of 6-MI located at the junction by 1.14 ± 0.173 ns (Figure 6-11, Table 6-3-B), consistent with local base stabilization. Binding of HU stabilizes the junction base where local dynamics are comparable to those observed at positions 5 and 7 of the helix away from the junction (Table 6-3-B). Similarly, for the 5’ overhang HU binding also reduces 6-MI conformational dynamics at the junction, Δθ_L is 0.79 ± 0.175 ns (Figure 6-12, Table 6-3-B). The trends for HU-induced conformational base dynamics in the duplex
portion are similar for the 3’ and 5’ overhang (Figure 6-11 and 6-12). HU binding reduces θ,
for 6-MI located 1 to 4 basepairs from the junction in the duplex portion, indicative of
dsDNA destabilization, consistent with TR-fluorescence data (Figure 6-11 and 6-12). As the
position of 6-MI is moved ≥ 5 basepairs from the junction and into the duplex portion, the
effect of HU binding on 6-MI rotational correlation time is minimized (Figure 6-11 and 6-
12).

In addition to altering base dynamics in the duplex region, HU binding influences the
rotational correlation time for 6-MI positioned in the single-strand portion. For example, HU
binding to the 5’ overhang increases the rotational correlation time from 330 to 560 ± 150 ps
for 6-MI located 7 to 12 bases from the junction in the single-strand portion (Figure 6-12,
Table 6-3-B). The longer rotational correlation time is attributed to reduced base dynamics
and is not as pronounced as the behavior observed for 6-MI at the junction (Figure 6-11, 6-
12). Similarly, HU binding to the 3’ overhang increases θ,
for 6-MI located 5 bases from the
junction in the single-strand portion, again consistent with stabilization of the base (Figure 6-
11). There are subtle differences in the base dynamics between the respective HU bound 5’
and 3’ overhangs. For example, when HU binds the 5’-overhang and 6-MI is adjacent to the
junction its conformational flexibility increases (Figure 6-12). In contrast, HU binding to the
3’ overhang stabilizes 6-MI adjacent to the junction in the single-strand portion while
destabilizing 6-MI adjacent to the junction in the duplex portion (Figure 6-11).

6.2.5  

HU binding perturbs global DNA geometry as shown by Förster resonance energy transfer (FRET)
The structural specific binding mode of HU recognizes a distortion in the double-stranded DNA polymer. The junction represents one of these motifs, which is presumed to increase DNA flexibility, a characteristic shared by many constructs that HU binds to with increased specificity. The exact relationship between DNA geometry and HU binding affinity is not completely understood. This work examines energy transfer as a function of DNA conformation for the 5' and 3' overhangs in the presence and absence of HU to gain a better understanding of this relationship. We used three FRET distance measurements to define the geometry of the protein-overhang complex. The measured FRET distances include: 1) the overall distance of the DNA substrate, 2) the distance from the junction to the end of the duplex region, and 3) the distance from the junction to the end of the single-strand region as shown in Table 6-1. We have also measured the protein-induced bend angle for nicked DNA substrates. All substrates are based on a 34 bp construct of the IHF H1 consensus sequence.

To calculate the distance separation between dye pairs required measuring the excited state decay of the FAM labeled DNA and the DNA labeled with FAM and TAMRA. The measurement was then repeated in the presence of protein. The individual excited states decays were then modeled as a sum of exponentials to obtain the best fit in the program Globals. The amplitude weighted mean lifetime, based on the excited state parameters given in Tables 6-4, was then used to calculate energy transfer and distance separating dye pairs using equations 3.38 given in chapter 3. The TR-FRET was also globally modeled (donor only and donor + acceptor) in FargoFit using a probability distribution function 3.40 given in Chapter 3 to directly obtain FRET distance and the range in distance distributions. Global modeling of TR-FRET in FargoFit is an appropriate method to determine if there are distinct populations of dye-pairs separated at different distances using equation 3.41 given in chapter.
3. Calculating the distance dependent energy transfer required taking into account the percentage labeling.

In general, DNA is considered a linear polymer with predictable properties such as molecular geometry and stiffness. Therefore, it is possible to compare the measured FRET distance for labeled DNA to the theoretical linear polymer distance and calculate a bend angle between dye pairs based on the deviation between the expected and measured distances.\(^{320}\) The approximation for the end to end distance of the linear DNA construct, depends on the persistence length of DNA. Persistence length \(L_p\), refers to the length scale where a biopolymer deviates from linearity. For dsDNA, which is very stiff and can be represented as an ideal polyelectrolyte, the persistence length is \(\sim 500 \, \text{Å}.\)\(^{319,321-323}\) If the dsDNA oligomer is shorter than the persistence length, the end to end distance is approximated by its contour length \((L_{ds})\), given by \(L_{ds} = b_0N\). Where \(b_0 = 3.4 \, \text{Å}\) is the unit length of each basepair and \(N\) is the number of basepairs. In contrast, ssDNA is highly flexible with no well-defined secondary structure, which means the persistence length is considerably shorter. The persistence length for ssDNA ranges from 10 to 30 Å depending on salt concentrations; for example at 200 mM NaCl \(L_p\) is approximately equal to 22.5 Å.\(^{317}\) In this work the contour length for the single-strand portion of the overhang is \(\sim 100 \, \text{Å}\), which is much greater than its persistence length. For this reason the end to end distance of the single-strand portion is best approximated by wormlike chain (WLC), which can be used to model flexible polymers.\(^{324}\) The end to end distance for each construct was then calculated using the following equations:

\[
Model(\text{Å}) = L_{ds} + L_{ss} + L_{ linker}
\]  

\[(6.1)\]
where $L_{ds}$ is the length of duplex portion and $L_{ss}$ represents the length of the single-strand portion from the WLC model, which equals 69 Å for the 16-mer ssDNA portion. $L_{\text{linker}}$ varies depending on the dye used. It is set to 10.0 Å for the extended C6 linker between the dye FAM and the DNA.\textsuperscript{167,324,325} As rhodamine is assumed to stack with the end of the DNA the linker distance is set to zero.\textsuperscript{167,310,326} Similarly, since 6-MI is a fluorescent base analog there is no linker.

In the absence of HU, the efficiency of energy transfer for the 5’ and 3’ total overhang is $0.01 \pm 0.01$ and $0.04 \pm 0.01$ respectively, suggestive of a similar conformation for the total overhangs (Table 6-4). The efficiency for the 3’ overhang is below the detection limit of 0.02, thus energy transfer is less than 0.03 (Table 6-4). The lack of energy transfer for both overhangs means the distances between dyes must be longer than the theoretical distances of 140 Å, which corresponds to 0.03 energy transfer for a linear overhang construct (Table 6-4). The ~45 Å FWHM spread in distance distribution (Table 6-4.0, FWHM) is evidence that both 5’ and 3’ total overhang adopt a wide range of conformations. While the difference in energy transfer between the 5’ and 3’ total overhang is minor, the $0.11 \pm 0.02$ 3’ DS and $0.22 \pm 0.02$ 5’ DS energy transfer for the duplex regions is noticeable different. The FRET determined distance for the 3’ DS is $74 \pm 15.2$ Å, within error of the 71 Å theoretical distance for a linear 18bp duplex (Table 6-4). In contrast, the 5’ DS FRET distance is 10 Å shorter, which is associated with a distortion to the DNA geometry (Table 6-4). One possibility for the distorted 5’ DS region is this portion
Table 6-4 FRET determined distance and calculated bend angle for 3 and 5 prime overhang in the presence and absence of HU

<table>
<thead>
<tr>
<th>Overhang Constructs</th>
<th>Donor</th>
<th>Donor + Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α(1)</td>
<td>τ(1) (ns)</td>
</tr>
<tr>
<td>3' TOTAL</td>
<td>0.34</td>
<td>2.63</td>
</tr>
<tr>
<td>3' TOTAL + HU</td>
<td>0.35</td>
<td>2.20</td>
</tr>
<tr>
<td>3' DS</td>
<td>0.46</td>
<td>2.62</td>
</tr>
<tr>
<td>3' SS</td>
<td>0.34</td>
<td>2.63</td>
</tr>
<tr>
<td>3' SS + HU</td>
<td>0.35</td>
<td>2.20</td>
</tr>
<tr>
<td>5' TOTAL</td>
<td>0.59</td>
<td>2.83</td>
</tr>
<tr>
<td>5' TOTAL + HU</td>
<td>0.44</td>
<td>2.51</td>
</tr>
<tr>
<td>5' SS</td>
<td>0.34</td>
<td>2.72</td>
</tr>
<tr>
<td>5' SS + HU</td>
<td>0.32</td>
<td>2.38</td>
</tr>
</tbody>
</table>

The excited decay of donor only FAM labeled overhang DNA and donor FAM + acceptor TAMRA labeled overhangs. The mean lifetime \( \tau_{\text{mod}} \) was used to calculate energy transfer between dye pairs and the corresponding FRET distance. Global modeling of the excited state decay with the program \(^1\)FargoFit was also used to calculate the \(^2\)FWHM (\( \sigma \)) of the distance distribution. \(^3\)Bend angle calculated as shown in discussion \(^4\)Lab. = percent labeled The energy transfer was measured in the presence and absence of HU. N.C. refers to not calculated.
contains the IHF recognition sequence. Interestingly, the ~30 Å (Table 6-4, column σ) spread in the distance probability distribution for both 5’ and ‘3 DS is greater than previously reported distribution < 15 Å for dsDNA.\textsuperscript{228,327} One possible explanation for the discrepancy is fraying of the DNA at the junction. Energy transfer for both single stranded portion is ~0.5 (Table 6-4), consistent with the reported value for a similar dye pair separated by a 16 base single-strand portion.\textsuperscript{328} The spread in the distance distribution of FWHM ~50 Å is consistent with greater mobility of the single-strand region relative to the duplex DNA region (Table 6-4). The calculated FRET distance for the single-strand portion of 51 Å is shorter than the theoretical 79 Å distance for the WLC model (Table 6-4). There are several reasons for the discrepancy; first the model distance signifies the length with highest probability for a radius of gyration. Second, the WLC model is representative of poly dT; the diminished stacking interactions of poly dT may account for differences between experimental data and model distances.\textsuperscript{324} It is important to note the error in orientation factor was calculated using steady state anisotropy measurements for the donor and acceptor, this method is known to overestimate the error. In addition, probability distribution analysis of the overhang TR-FRET data for multiple populations with discrete average distances did not improve fits or produce unique conformations. In summary, the TR-FRET data indicates similar distorted conformational distributions for both 5’ and 3’ total overhang. In contrast to the linear 3’ DS portion, the 5’ DS portion is distorted. The spread in FRET distance distribution is greatest for the single-strand portion.

In the presence of HU energy transfer for the total 5’ and ‘3 overhang increases to ~0.18, corresponding to an end to end distance of ~65 Å(Table 6-4), consistent with a large protein induced distortion. HU binding also increases the spread in both the 3’ and ‘5 total overhang distance distribution σ by ~ 20 Å, which suggests protein binding raises the number
of possible overhang conformations. We hypothesized based on the bend angles in the HU DNA co-crystal structure\textsuperscript{165} binding of the $\beta$ arms in the duplex portion would induce a kink of $\sim 60^\circ$, corresponding to an increase in energy transfer from 0.11 to 0.22. However, when HU binds to the overhang there is no change in energy transfer for the duplex portion (5’ DS and 3’ DS) of the overhang (Table 6-4). It is possible the HU induced distortion is not symmetric between arms and results in a change in energy transfer below the detection limit of $<0.03$. It is important to reemphasize the 5’ DS portion is distorted in the absence of protein and it is possible the duplex portion is not further distorted with protein binding. Similarly, there is no change in energy transfer for the single-strand portion (3’ SS and 5’ SS) of the overhang with protein binding (Table 6-4). It possible the inherent flexibility of the single-strand portion limits the ability to detected HU-induced distortion. There was no evidence supporting multiple populations with different distance distributions when analyzing the HU-overhang TR-FRET data. In summary, HU binding distorts the global structure of both the 5’ and 3’ overhang, consistent with previous FRET measurements and the HU-DNA co-crystal structure.\textsuperscript{165} We did not observe a protein induced distortion in the duplex or the single-strand portion of the overhang. It is possible the HU induced distortion away from the junction is below the detection limit for the duplex portion or the flexibility of the single-strand portion hinders our ability to detect the protein-induced distortion.

The inherent flexibility of the single-strand portion of the overhang makes it hard to approximate the theoretical linear distance of the construct. One way to address this problem is to examine another high affinity construct where there is greater confidence in the linear model lengths. The nicked construct (Table 6-1) satisfies these requirements: 1) HU binding affinity is similar to the overhang (Table 6-2 and Figure 6-3) the complete model length of 126 Å is well defined.
Table 6-5 FRET determined distance and calculated bend angle for nicked DNA in the presence and absence of HU

<table>
<thead>
<tr>
<th>Nicked DNA</th>
<th>Donor</th>
<th>Donor + Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(a_1) (\tau_{(1)}) (ns)</td>
<td>(a_2) (\tau_{(2)}) (ns)</td>
</tr>
<tr>
<td>3’ FAM_nic_total</td>
<td>0.51 1.37 0.49 4.81 3.05</td>
<td>0.52 1.40 0.48 4.73 3.01</td>
</tr>
</tbody>
</table>

Table 6-5 The excited decay of donor only FAM labeled nicked DNA and donor FAM + acceptor TAMRA 5-prime labeled nicked DNA. The mean lifetime \(\tau_{(mol)}\) was used to calculate energy transfer between dye pairs and the corresponding FRET distance. Global modeling of the excited state decay with the program \(^1\text{FargoFit}\) was also used to calculate the \(\text{FWHM (σ)}\) of the distance distribution. \(^2\text{Bend angle calculated as shown in discussion}\) \(^3\text{Lab. = percent labeled}\) The energy transfer was measured in the presence and absence of HU. N.C. refers to not calculated.
In absence of HU, energy transfer for the complete nicked construct is 0.01 ± 0.02, if the construct is distorted it must be below the detection limit of < 0.03 (Table 6-5). The FRET determined distances for the 5’ and 3’ duplex portions (5’ FAM_nic_DS and 3’ FAM_nic_DS) are consistent with the overhang FRET data (Table 6-4 and 6-6). For example, the FRET measured distance for 3’duplex portion downstream of the nick (3’ FAM_nic_DS) and the duplex portion of the 5’ overhang (5’ DS) are 60 ± 13.8 Å and 64 ± 13.7 Å (Table 6-4 and 6-6). The distance for the 3’ duplex portion of the nick (3’ FAM_nic_DS) is 15 Å shorter than the calculated linear model distance, evidence that the DNA 3’ to the nick is distorted. In contrast, the measured FRET distance for the 5’ duplex portion of the nick (5’ FAM_nic_DS) is 74 ± 15.2 Å, consistent with the calculated linear model distance of 71 Å (Table 6-5). In the presence of HU energy transfer for the total nicked DNA construct increases to ~0.1 ± 0.025, corresponding to distance between dye pairs of ~75 ± 19 Å (Table 6-4). The FRET measured distance for HU bound total nick is similar to measured distance for the HU overhang complex, which suggest the HU-induced distortion is comparable for the two constructs (Table 6-4 and 6-5). In addition, the spread in the dye pair distance distribution FWHM for the HU-nicked complex is ~ 60 Å equivalent to the HU-overhang complex (Table 6-4 and 6-5). HU binding has no effect on energy transfer for the 5’ and 3’ portion of the nicked DNA (5’ FAM_nic_DS and 3’ FAM_nic_DS) (Table 6-5). However, the spread in distance distribution FWHM for the 5’ and 3’ duplex portion of the nick increases by ~30 Å, an indicator the portions of nicked-DNA are adopting a larger number of conformations (Table 6-5). It is important to reiterate the 3’ duplex portion of the nick is distorted in the absence of HU and it is possible HU binding in this region does not result in greater perturbation to DNA conformation. In summary, in the absence of HU the total nicked construct is slightly distorted. The HU induced conformational change to the
Table 6-6 FRET determined distance and bend angle for H1 duplex (H34aL5') in the presence and absence of IHF and HU

<table>
<thead>
<tr>
<th>Duplex</th>
<th>Donor</th>
<th>Donor + Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha_{(1)}$</td>
<td>$\tau_{(1)}$ (ns)</td>
</tr>
<tr>
<td>H34aL5' ds</td>
<td>0.42</td>
<td>2.64</td>
</tr>
<tr>
<td>H34aL5' ds + IHF</td>
<td>0.42</td>
<td>2.71</td>
</tr>
<tr>
<td>H34aL5' ds + HU 1:1</td>
<td>0.42</td>
<td>2.70</td>
</tr>
<tr>
<td>H34aL5' ds + HU 1:2</td>
<td>0.41</td>
<td>2.72</td>
</tr>
<tr>
<td>H34aL5' ds + HU 1:3</td>
<td>0.40</td>
<td>2.74</td>
</tr>
</tbody>
</table>

Table 6-6 The excited decay of donor only FAM 5-prime labeled dsDNA and donor FAM + acceptor TAMRA 5-prime labeled dsDNA. The mean lifetime $\tau_{\text{mol}}$ was used to calculate energy transfer between dye pairs and the corresponding FRET distance. Global modeling of the excited state decay with the program $^1$FargoFit was also used to calculate the $^2$FWHM ($\sigma$) of the distance distribution. $^3$Bend angle calculated as shown in discussion $^4$Lab. = percent labeled The energy transfer was measured in the presence and absence of HU. N.C. refers to not calculated.
total nicked DNA is comparable to the perturbation to the overhang construct. The HU
induced distortion to the 5’ and 3’ duplex portions of the nicked construct are below the
detection limit.

In addition to the investigating HU induced distortion, we also measured the IHF
induced distortion to the H1 duplex DNA. IHF is a structural homolog of HU, binding to a
consensus sequence in dsDNA. Measurements of the IHF-induced bend range from 120 to
160° making it a model system for comparison with HU.167,268,318,329-332 The theoretical
distance between FAM and RHO in H34aL5 is 126 Å, outside the acceptable detection range.
When IHF is bound to H34aL5 energy transfer increases to 0.47 ± 0.023 (Table 6-6), within
error of the 0.51 energy transfer reported for IHF bound to a similar 35-mer.318 The measured
energy transfer for IHF bound to H34aL5 corresponds to 53 ± 14.6 Å separating the
fluorescent dyes (Table 6-6). The IHF-DNA FRET distance is shorter than the HU-nicked
DNA distance of 75 ± 19 Å indicating the IHF induced distortion is greater. Interestingly, the
spread in the distance distribution for the IHF-DNA complex is 54 Å, slightly smaller than
the calculated distributions for HU bound to the nick and overhang. The wide distribution
suggests IHF increases the number of conformations the DNA can adopt and is consistent
with the bend distribution of IHF-DNA complexes observed with AFM.330,331

HU uses the non-specific mode when binding dsDNA. When the ratio between HU
and dsDNA is 1:1 no energy transfer is detectable, indicative of a bend angle less than 50°
(Table 6-5). Increasing the ratio of HU to DNA to 2:1 results in energy transfer of 0.02 ±
0.022, indicating minimal distortion to the 34 basepair duplex DNA (Table 6-6). At the
reported HU to DNA stoichiometric ratio60 of 3:1, the energy transfer may increase slightly to
0.03 ± 0.024 (Table 6-6). The maximum possible HU-induced change to the measured FRET
distance is a decreases of7 Å consistent with a bend angle of 82°, approximately 27° per HU
Interestingly, the bend angle per HU is similar for both the 2:1 and 3:1 ratios. The maximum possible HU-induced change to the measured FRET distance is a decrease of 7 Å indicating perturbation to the DNA structure through HU non-specific binding is minimal. To obtain a better understanding of HU induced distortion to the 34-mer requires a dye pair with an \( R_0 > 52 \) Å.

### 6.2.6 FRET mapping the HU-overhang complex

The nature of protein-DNA interactions that lead to structurally specific and non-specific HU binding modes remains elusive. In the HU-DNA co-crystal structure it was observed that HU binding kinks the DNA through intercalation of a conserved Pro residue into the DNA and stacking with an adjacent base. Electrostatic interactions between positively charged residues in the β-arm arms contribute to the stabilization of the HU-DNA complex. To better understand the differences in the two types of binding interactions this work used a combination of intermolecular FRET measurements to generate distance constraints from which a molecular model can be proposed. Two types of constructs are used to determine the distances one in which the donor 6-MI is incorporated into an overhang DNA construct at 8 positions relative to the junction and the acceptor dye FAM is used to label 5 specific locations in the protein HU through a single Cys mutation.

The five site-specific HU mutants include βQ64C and βT70C located in the β-ribbon of HUβ subunit. The location of the other three mutants are in the α-subunit. Specifically, the location of αQ43C is in β-strand 1, αA78C is in the saddle of the protein in β-strand 3, and αK90C located at the C-terminus in α-helix 3 (Figure 6-1). The mutations do not affect HU affinity for dsDNA.\(^{305}\) This work confirmed HU mutants maintain high affinity to the overhang by titration of the mutant labeled protein into a sample containing 3’ overhang.
labeled with TAMRA at constant concentration of 400 pM. Shown in Figure 6-4 and Table 6-2 are the representative binding curves and the corresponding association constants for each of the mutants. This work used 6-MI incorporated into a 34-mer at two different positions, adjusting its relative position to the junction with different lengths of 5’ or 3’ duplex portions.

To calculate the distance separation between dye pairs required measuring the excited state decay of the 6-MI labeled overhang complexed with unlabeled HU mutant. The measurement was then repeated with a HU mutant labeled with FAM. The individual excited decays were then fit to a sum of exponentials in the program Globals to obtain lifetime parameters. The τ_{mol} was used to calculate energy transfer and distance separating dye pairs. In addition, the TR-FRET was globally modeled (donor only and donor + acceptor) in FargoFit using the equation 3.40 given in Chapter 3 to directly obtain the average FRET distance and the probability distance distributions. The TR-FRET was also evaluated in the program FargoFit for multiple populations with distinct distance distributions using equation 3.41 in Chapter 3. Calculating the distance dependent energy transfer required taking into account the percentage labeling. The FRET distances provided distance constraints to evaluate four possible molecular models of the HU-DNA complex (Chapter 3 Figure 6-13). The following section highlights some of the key features of the TR-FRET data.
Table 6-7 TR-FRET between 6-MI incorporated overhangs and FAM site-specific labeled HU

<table>
<thead>
<tr>
<th></th>
<th>Donor</th>
<th>Donor + Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a_1$ (ns)</td>
<td>$\tau_1$ (ns)</td>
</tr>
<tr>
<td>TFA15_10bp5 + Q43C</td>
<td>0.71</td>
<td>2.57</td>
</tr>
<tr>
<td>TFA15_10bp5 + Q64C</td>
<td>0.54</td>
<td>1.88</td>
</tr>
<tr>
<td>TFA15_10bp5 + T70C</td>
<td>0.45</td>
<td>3.17</td>
</tr>
</tbody>
</table>
The β-ribbon of HU follows the minor groove of the DNA with P63 intercalating between bases. For this reason the two mutants, βQ64C and βT70C, are excellent indicators of the site of intercalation and location of the β-ribbon arm for the β-subunit. The mutants in HUα are representative of the amino acids making up the “body” of the protein (Figure 6-1). A core of α-helices at the dimer interface and a set of 3 β-sheets located above the helical core make up the body of HU. The position of the mutant αA78C is in β-strand 3, approximately equal distance (~28 Å) from the two intercalating prolines (Figure 6-1). The arrangement of HU’s body is such that some residues from HUα core are closer to the β-ribbon arm of HUβ than that of HUα (Figure 1). For example, the mutant αK90C, located in α-helix 3 is 25 Å from βQ64C compared to 42 Å from αQ64C (Figure 6-1). Similar to αK90C, the mutant αQ43C located in β-strand 1 is 36 Å from βQ64C compared to 40 Å from αQ64C (Figure 6-1).

When the location of 6-MI is at the junction in the 3’ overhang energy transfer to FAM is greatest for mutant α78C, E.T. = 0.41 ± 0.024 compared to ~0.18 for other mutants (Table 6-7), which suggests the center of the protein is close to the junction. Energy transfer between 6-MI and the FAM labeled mutants progressively decreases as the position of 6-MI moves away from the junction and into the duplex portion of the 3’overhang to a minimum of ~0.10 when the position of 6-MI is 10 basepairs from the junction, which suggests the β-ribbon arm of the β subunit is not located in the duplex portion of the overhang (Table 6-7). The exception to this trend is αQ43C, energy transfer first increases from 0.15 ± 0.012 for 6-MI at the junction to 0.23 ± 0.006 when 6-MI is 3 basepairs from the junction followed by a decrease to 0.17 ± 0.005 when the position of 6-MI is 10 basepairs away from the junction (Table 6-7). A kink at the junction would explain this data, making 6-MI 3 basepairs from the junction closer to αQ43C compared to 6-MI at the junction and 6-MI 10 basepairs away in the duplex portion would be farther from αQ43C. Interestingly, energy transfer between
αQ43C and 6-MI positioned 3 basepairs from the junction and in the duplex portion of the 5’ overhang is 0.24 ± 0.003, consistent with the data for the 3’ overhang (Table 6-7).

When 6-MI is positioned in the single-strand portion of the overhang energy transfer is highest for the mutants located in the β-ribbon arm of the β subunit. The behavior is the same for both 5’ and 3’ overhang. For example, energy transfer from TFA07_19bp3 to βQ64C is 0.97 ± 0.011 compared to 0.17 ± 0.011 for TFA07_19bp3 to αQ43C, which suggests this is close to the site of βP63 intercalation (Table 6-7). It is important to note the position of 6-MI in TFA07_19bp3 is 9 bases from the junction in the single-strand portion of the 5’ overhang (Table 6-7). In addition, energy transfer from TFA15_10bp5 to βT70C is 0.95 ± 0.011, the position of 6-MI is 5 bases from the junction in the single-strand portion of the 3’ overhang (Table 6-7). While energy transfer between TFA15_10bp5 and βQ64C is 0.56 ± 0.068 (Table 6-7). The greater amount of energy transfer to βT70C compared to βQ64C when 6-MI is located closer to the junction is consistent with the β subunit located away from the junction. The relative location of αK90C to βQ64C is a possible reason for the similar trends in energy transfer between the mutants (Table 6-7). For example energy transfer to αK90C is greatest when the location of 6-MI is 15 bases away from the junction in the single-strand portion of the 5’ overhang (Table 6-7). When the TR-FRET was analyzed globally in FargoFit there was no evidence of multiple populations with distinct distance distributions. It is important to note the poor labeling efficiency for some of the mutants, which can lead to greater error in the approximation of labeling efficiency. We used the FRET determined distances to evaluate models of the HU-overhang complex. The evaluation of the models in terms of the FRET distance measurements is outlined in the following discussion.
6.3 Discussion

HU is one of the most abundant nucleoproteins in bacteria. HU is unique in that it does not identify a particular DNA sequence for binding, instead it recognizes DNA topology binding with different affinity dependent on DNA conformation and dynamics. The fact that HU binding to DNA is independent of sequence context has made it hard to develop a detailed mechanism describing HU specificity.

The focus of this study is to examine the interaction between HU and the high affinity overhang constructs. This work used time-resolved fluorescence with the ability to incorporate the fluorescent base analog 6-MI into the overhang construct to examine DNA conformational dynamics. This method provides information on the microenvironment and conformational dynamics with single-base resolution. In addition, FRET data from end labeled DNA constructs was used to examine DNA topology and the HU-induced distortion to the global DNA conformation. These two methods are useful for understanding both the role of DNA conformation in the mechanism of HU specificity and how HU binding alters the conformation of the DNA. Finally, an assortment of intermolecular FRET measurements was used to generate molecular distance constraints to evaluate four possible models of the HU DNA complex. The discussion section is separated into the following parts: 1) generation of the models, 2) evaluating the models in terms of FRET distances 3) validation of the model in terms of experimental data 4) new insights into the mechanism of HU specificity.

6.3.1 Generating models of the HU-overhang complex

In the literature there is a model for the specific binding of HU to cruciform and nicked DNA both of which contain a region of increased base flexibility. The overhang DNA construct
contains a related motif and HU binds with similar affinity, for these reason the interactions between HU and the overhang is thought to be comparable to the literature models. One objective was evaluate this idea by generating a model of HU bound to the overhang and comparing it with the literature models. Importantly, this is not a final model further investigation is necessary into the mechanism of HU specificity before a complete model can be brought forth. To generate the possible HU-overhang complex models, several assumptions based on experimental data from this study and the literature are made which the models must satisfy. The first step in generating a model was to approximate the HU-induced bend to the overhang constructs. Two assumptions are embedded in the mathematics: first, the location of the bend is halfway between dye pairs and second, the bend is symmetric as shown in Figure 6-13. In Figure 6-13 the bend angle is calculated using the law of cosines given by:

\[ \cos \gamma = \frac{a^2 + b^2 - c^2}{2ab} \]  

(6.2)

\[ \text{Bend Angle} \ \theta = 180 - \gamma \]  

(6.3)

where a and b are set equal to half the linear distance and c represents the FRET determined distance between dye pairs. As shown in equation 6.3 to calculate the bend angle \( \theta \), the angle \( \gamma \) is subtracted from 180 (the angle for a linear construct). The FRET distance and bend angle were calculated in the presence and absence of protein for the series of constructs and the results are shown in Tables 6-4 through 6-6.
To calculate the bend angle requires an estimation of the theoretical model distance $M$ separating the dye pair. The bend angle is calculated using the law of cosines. The length of $a$ and $b$ is assumed to be half the distance of $M$, while $c$ is the FRET measured distance between dye pairs.

\[ a^2 + b^2 - (2ab) \]

\[ \cos \gamma = \frac{a^2 + b^2 - c^2}{2ab} \]

**Bend Angle $\theta = 180 - \gamma$**

Figure 6-13 Calculate Bend angle
The inherent flexibility of the single-portion of the overhang makes it hard to approximate a distance for the linear construct and for this reason the bend angle in the absence of HU was not calculated for the total overhang or the single strand portion. We reported the duplex portion of the 3’ overhang (3’ DS) is linear (Table 6-4). In contrast, the duplex portion of the 5’ overhang exhibits an intrinsic DNA curvature of ~50° (Table 6-4). This portion of the H1 34mer contains the IHF recognition sequence, which may explain the intrinsic DNA curvature. We know from mapping 6-MI conformational dynamics in the overhang that HU binding reduces the flexibility of the single-strand portion. Assuming when bound to HU the conformation of the single-strand portion is limited we calculated the HU-induced bend angle for the overhangs. The HU induced bend for the total 5’ and 3’ overhangs is ~120° (Table 6-4). In the presence of HU there was not a detectable change in energy transfer for the duplex portions of the overhangs (Table 6-4). It is possible the induce bend in the duplex portion is below the detection limit of 0.03 energy transfer corresponding to a bend angle of ~30° and as stated above the 5’ DS portion exhibits an intrinsic curvature. We also calculated a bend angle for the single-strand portion of ~90°. This value is most likely an overestimation of the bend angle because of the poor approximation for a linear single-strand portion. A construct with a better approximation for the linear distance is the nicked 34mer.

In absence of HU the maximum possible DNA bend angle for the total nicked construct is ~70° (Table 6-5). Similar to the overhang data, the 3’ duplex portion of the nick (3’ FAM_nic_DS) exhibits an intrinsic DNA curvature of ~60° (Table 6-5). In the presence of HU the bend angle for the total nicked construct is ~110° (3’ FAM_nic_total + HU) (Table 6-5). Interestingly, we do not observe HU-induced bending in the 5’ or 3’ portion of the nicked construct (Table 6-5). It is possible HU binds to the pre-bent 3’ portion of the nicked construct without further bending of the DNA.
For comparison we measured the IHF induced bend angle to the H1 34 mer. Based on the FRET determined distance the calculated bend angle is 130°, within the established experimental range of 120-160° (Table 6-6). In addition, the HU-induced bend angle to double-stranded H1 34 mer at different three different stoichiometric ratios was calculated (Table 6-6). Previous reports indicate the HU-induced bend angle for dsDNA ranges from 25 to 150°. While the energy transfer for double-stranded 34 mer (H34aL5’ ds) is at or below the detection limit, we can approximate the maximum possible bend angle for stoichiometric ratios 1:2 and 1:3 DNA to HU. At a stoichiometric ratio of 1:2 DNA to HU the maximum protein induced bend angle is 72°, or ~35° per HU (Table 6-6). While at a stoichiometric ratio of 1:3 DNA to HU, the maximum protein induced bend angle is 82°, or ~27° per HU (Table 6-6). The FRET data demonstrates HU structurally specific binding induces a larger bend angle compared to the structurally non-specific binding mode. In summary, the HU-induced bend angle is similar for the structurally specific overhang and nicked construct. The average HU induced bend angle is ~116° for HU bound high affinity constructs, which is consistent with the literature bend angle range from 105 to 139°. The HU-DNA and IHF-DNA co-crystal structures indicate the bend is symmetric between subunits. Therefore, based on the crystal structure and FRET data one condition the models must satisfy is the HU induced distortion is a symmetric bend with each subunit equally distorting the DNA by 60°.

Other physical parameter the models must satisfy include structural data from the Anabaena HU crystal structure. These parameters include the β-ribbon arms from each monomer must follow the minor groove of the DNA, there must be nine basepairs between intercalating prolines, and finally the DNA is kinked at the site of intercalation. An additional requirement is that one of the prolines must intercalate at the junction, which
comes from footprinting data and HU-DNA co-crystal structure. Implicit to the models is the absence of HU binding polarity. Meaning HU binds the 3 and 5-prime overhangs equally. In contrast, to another study, in this work HU was shown to bind the 3 and 5-prime overhangs with similar affinity (Table 6-2). The authors of other study suggests a preference for the 3’ overhang has to with the orientation of the protein relative to the junction. The earlier work obtained the apparent dissociation constants from GMSA using the following equation:

\[ K_{D_{app}} = \left( \frac{p-d\star(c+F)}{c} \right) \cdot F \]  

(6.4)

In Eqn 6.4 p represents protein concentration, d is the DNA concentration, C is the intensity of the complex band, and F is the intensity of the free DNA band. A possible issue with this analysis is the lack of a HU-DNA complex band in the gel-image for HU binding the 5’ overhang as shown in Figure 6-14. Analysis of the gels in terms of just free DNA produces a \( K_{D_{app}} \) for the 5’ overhang of \(~14\) nM and the 3’ overhang of \(~7\) nM (Figure 6-14), which suggests HU complexed with the 5-prime overhang dissociates in the gel. In addition, when the overhang is modified to a RNA-DNA hybrid the binding polarity is absent and the association constant for both 5 and 3’ overhangs is \(~63\) \(\mu\)M\(^{-1}\).

These physical parameters were used to develop four models representing HU binding to the overhang. Shown in Figure 6-15 are the four representative models. Each model fulfills the prerequisite conditions: the β-ribbon arms follow the minor groove of the DNA, one proline kinks the overhang at the junction by 60°, and the counterpart proline
A) The above gels adapted from D. Kamashev and J. Rouviere-Yaniv\textsuperscript{103} show HU binding with 5’ overhang, dsDNA, and the 3’ overhang. The authors noted the presence of the HU-DNA complex only in the 3’ overhang gel. Analysis of the gel in terms of the band for the complex and free DNA returns a $K_{d_{app}}$ of 16 nM and 1000 nM for the 3 and 5 prime overhangs respectively. 

B) Analysis of HU binding in terms of free DNA band intensity returns a $K_{d_{app}}$ of 7 nM and 14 nM for the 3 and 5 prime overhangs respectively, consistent with the data presented in this work. As shown in the plot the intensity of the dsDNA band does not change with HU concentration.

\textbf{Figure 6-14 GMSA HU overhang}
kinks the DNA 9 bases away by a corresponding 60°. The models differ in the orientation of HU with respect to the junction. First either αHU or βHU binds at the junction while the other subunit can bind in the single-strand or duplex portion of the overhang. Shown in Figure 6-15 is the four conformations with the site-specific mutants shown in stick representation colored accordingly: 1) αQ43C = green 2) αA78C = yellow 3) αK90C = magenta 4) βQ64C = red 5) βT70C = orange. The base representing the position of 6-MI within the DNA is colored blue (Figure 2.0). The distance between each of the site-specific mutants Cα and the C1’ of the base representing the position of 6-MI was measured in Pymol. For each construct, the RMSD between FRET and model distances was calculated. In addition, the Pearson correlation coefficient (r), the coefficient of determination (R^2), and Chi square (χ^2) between FRET and predicted distances was used as method to evaluate the goodness of fit for each model.

6.3.2 Evaluation of models in terms of FRET distances

This study used an array of constructs to generate distance constraints for the position of HU on the overhang. The sites of the cysteine substitutions for labeling with FAM are spatially separated providing accurate data for distinguishing between different conformations of HU bound to overhang DNA. The distributed positions of the donor 6-MI within the overhang was chosen to provide an assortment of distance dependent energy transfer measurements. A direct correlation between FRET distances and model predicted values was not expected because of uncertainty in the experimental distances related to the inherent flexibility of the dye tethers, utilizing Cα positions of site mutations rather than the
Molecular models represent the possible binding modes of the HU heterodimer to TFA15_18bp5, a 18-basepair 3-prime overhang. The models assume HU symmetrically bends the DNA by 60°. The guanine base colored blue represents 6-MI. The HU site specific mutations are represented by the following stick colors, green for α-HU Q43C, yellow for α-HU A78C, magenta for α-HU K90C, red for β-HU Q64C, and orange for β-HU T70C: Model (A) represents α-HU binding at the junction with β-HU intercalating 9-bases away into the single-strand portion, Model (B) represents β-HU binding at the junction with α-HU intercalating 9 bases away into the single-strand portion, Model (C) represents α-HU binding at the junction with β-HU intercalating 9-bp away into the duplex portion of overhang, Model (D) represents β-HU binding at the junction with α-HU intercalating 9-bp away into the duplex portion of the overhang.
<table>
<thead>
<tr>
<th></th>
<th>FRET R (Å)</th>
<th>HU binds to single-strand</th>
<th>HU binds to double-strand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Model A HUα at junc. (Å)</td>
<td>Model B HUβ at junc. (Å)</td>
</tr>
<tr>
<td>TFA15_10bp5' + Q43C</td>
<td>34 ± 6</td>
<td>31</td>
<td>36</td>
</tr>
<tr>
<td>TFA15_10bp5' + Q64C</td>
<td>32 ± 6</td>
<td>32</td>
<td>39</td>
</tr>
<tr>
<td>TFA15_10bp5' + T70C</td>
<td>23 ± 3</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>TFA15_15bp5' + Q43C</td>
<td>46 ± 17</td>
<td>45</td>
<td>38</td>
</tr>
<tr>
<td>TFA15_15bp5' + A78C</td>
<td>40 ± 8</td>
<td>34</td>
<td>33</td>
</tr>
<tr>
<td>TFA15_15bp5' + K90C</td>
<td>48 ± 10</td>
<td>41</td>
<td>34</td>
</tr>
<tr>
<td>TFA15_15bp5' + Q64C</td>
<td>45 ± 10</td>
<td>46</td>
<td>22</td>
</tr>
<tr>
<td>TFA15_15bp5' + T70C</td>
<td>46 ± 9</td>
<td>38</td>
<td>25</td>
</tr>
<tr>
<td>TFA15_18bp5' + Q43C</td>
<td>54 ± 11</td>
<td>42</td>
<td>31</td>
</tr>
<tr>
<td>TFA15_18bp5' + Q64C</td>
<td>56 ± 13</td>
<td>53</td>
<td>22</td>
</tr>
<tr>
<td>TFA15_18bp5' + T70C</td>
<td>57 ± 13</td>
<td>44</td>
<td>27</td>
</tr>
<tr>
<td>TFA07_18bp5' + Q43C</td>
<td>49 ± 12</td>
<td>56</td>
<td>36</td>
</tr>
<tr>
<td>TFA07_18bp5' + A78C</td>
<td>65 ± 18</td>
<td>49</td>
<td>46</td>
</tr>
<tr>
<td>TFA07_18bp5' + K90C</td>
<td>68 ± 14</td>
<td>53</td>
<td>48</td>
</tr>
<tr>
<td>TFA07_18bp5' + Q64C</td>
<td>70 ± 14</td>
<td>74</td>
<td>45</td>
</tr>
<tr>
<td>TFA15_18bp5' + T70C</td>
<td>62 ± 13</td>
<td>62</td>
<td>54</td>
</tr>
<tr>
<td>TFA15_23bp3' + Q43C</td>
<td>52 ± 10</td>
<td>48</td>
<td>34</td>
</tr>
<tr>
<td>TFA15_23bp3' + Q64C</td>
<td>54 ± 11</td>
<td>59</td>
<td>22</td>
</tr>
<tr>
<td>TFA15_23bp3' + T70C</td>
<td>56 ± 12</td>
<td>50</td>
<td>31</td>
</tr>
<tr>
<td>TFA15_20bp3' + A78C</td>
<td>50 ± 10</td>
<td>28</td>
<td>30</td>
</tr>
<tr>
<td>TFA15_20bp3' + K90C</td>
<td>57 ± 10</td>
<td>41</td>
<td>26</td>
</tr>
<tr>
<td>TFA15_20bp3' + Q64C</td>
<td>42 ± 10</td>
<td>52</td>
<td>19</td>
</tr>
<tr>
<td>TFA07_19bp3' + Q43C</td>
<td>40 ± 7</td>
<td>36</td>
<td>45</td>
</tr>
<tr>
<td>TFA07_19bp3' + Q64C</td>
<td>20 ± 3</td>
<td>20</td>
<td>49</td>
</tr>
<tr>
<td>TFA07_19bp3' + T70C</td>
<td>36 ± 7</td>
<td>27</td>
<td>40</td>
</tr>
<tr>
<td>TFA07_12BP3' + A78C</td>
<td>46 ± 11</td>
<td>44</td>
<td>46</td>
</tr>
<tr>
<td>TFA07_12BP3' + K90C</td>
<td>41 ± 7</td>
<td>41</td>
<td>54</td>
</tr>
</tbody>
</table>

1)
Actual fluorophore location to predict model distances, and possible errors in calculating labeling efficiency. Thus, an acceptable value for a predicted distance is within one standard deviation of the experimental distance. In many cases more than one predicted model distance was within one standard deviation of the experimental FRET distance (Table 6-8). To be confident that the right model was selected this study measured an array of FRET distances.

A quantitative test to measure the strength and direction of a linear association between the measured FRET and predicted model distances is the Pearson linear correlation coefficient \( r \). The value of \( r \) ranges from \(-1 \leq r \leq 1\), where a value \( \geq +0.8 \) indicates a strong positive linear relationship and a value \( \leq -0.8 \) indicates a strong negative correlation between values. If the correlation is weak, the value is below \( \pm 0.5 \) closer to zero. Additional methods of evaluating the models include comparing the coefficient of determination \( R^2 \) which represents the fraction of the data that is closest to the line of best fit and the \( \chi^2 \) parameter an indicator of agreement between observed and predicted values.

As shown in Figure 6-16, the best linear correlation between FRET and predicted distances is achieved with Model A; this model yielded a reasonable fit to experimental data with values of 0.81 for \( r \) and a \( \chi^2 \) of 2.14. The Pearson correlation \( r \) improves further to 0.89 and a \( R^2 \) to 0.80 by removing three points with the greatest deviation. In this model the \( \alpha \text{HU} \) subunit binds at the junction and the \( \beta \text{HU} \) subunit binds in the single-strand portion of the overhang. It is worth noting, as shown in Figure 6-16 there is a negative correlation \( r = -0.77 \) between FRET and predicted distances for Model C. Model C represent \( \alpha \text{HU} \) subunit binding at the junction and the \( \beta \text{HU} \) subunit binding in the duplex portion of the overhang.
Correlation between FRET distances and predicted distances from the four models of HU bound to the overhang indicates that the best representation of HU binding the overhang is model (A) of αHU binding at the junction and βHU binding in the single-strand portion. The above plots are of the FRET distances vs those predicted from the indicated HU binding orientations. Distances and errors are listed in Table 1.0. The linear solid line in each graph depicts the perfect case of a 1:1 correspondence when the FRET distances exactly match with a particular model. Calculating the Pearson correlation coefficient ($r$), the coefficient of determination ($R^2$), and Chi square ($\chi^2$) between FRET and predicted model distances was used to determine the goodness of fit for each model.
The negative correlation indicates the orientation of the model is opposite the actual orientation of HU bound to the overhang.

All other predicted models exhibit \( r \) values \( \leq -0.01 \) and \( \chi^2 \geq 13.90 \) (Figure 6-16). For the other three models, fewer than half of the predicted distances are within error of the FRET distances consistent with the determined statistical parameters (Table 6-8). As shown in Table 6-8 nineteen of the predicted distances for Model A are within error of the experimental FRET distances. In addition to the above statistical analysis the RMSD between predicted and FRET distance was calculated (Table 6-9). The RMSD for Model A, where the αHU subunit binds at the junction and the βHU subunit binds in the single-strand portion of the overhang, is 9 Å within the average error of ± 10 Å for FRET distances, while the RMSD for all other models is \( \geq 20 \) Å.

There are several other possible explanations for the FRET distances that must be ruled out. First is the possibility that HU binds to the junction in no preferential orientation, meaning there is equal probability for each subunit to be located at the junction. If this was the case, the FRET distance measured between 6-MI located at the junction and FAM located in the β-ribbon arm (βQ64C, βT70C) would be an average of the two orientations and as such similar to the average of two orientations measured when 6-MI is located ~ 9 bp from the junction. This is not the case the FRET distance between 6-MI located 8 basepairs from the junction and βQ64C is 20 ± 3 Å (TFA07_19bp3) compared to 46 ± 10 Å when 6-MI is located at the junction (TFA15_20bp3). Importantly, the measured distance of 46 ± 10 Å between βQ64C and junction 6-MI (TFA15_20bp3) is consistent with the 40 Å between intercalating prolines. In addition, global analysis of the TR-FRET for multiple populations with distinct distance distributions suggest the presence of only a single population. Another
Table 6-9 Summary of RMSD between model and FRET distances for each construct

<table>
<thead>
<tr>
<th></th>
<th>HU Binds to single-strand</th>
<th>HU Binds to double-strand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model A</td>
<td>Model B</td>
</tr>
<tr>
<td></td>
<td>HUα at junc. (Å)</td>
<td>HUβ at junc. (Å)</td>
</tr>
<tr>
<td>TFA15_10bp5'</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>TFA15_15bp5'</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>TFA15_18bp5'</td>
<td>10</td>
<td>29</td>
</tr>
<tr>
<td>TFA07_18bp5'</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>TFA15_23bp3'</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>TFA15_20BP3'</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td>TFA07_19bp3'</td>
<td>7</td>
<td>22</td>
</tr>
<tr>
<td>TFA07_12BP3'</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>5-prime RMSD</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>3-prime RMSD</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>Total RMSD</td>
<td>9</td>
<td>20</td>
</tr>
</tbody>
</table>
possibility is the binding orientation is dependent on the type of overhang 5-prime vs 3-prime. Separating the data into the 5-prime overhang and 3-prime overhang groups produces similar results, the lowest RMSD between FRET measured and predicted distances is Model A (Table 6-9). Finally, there is a possibility that labeling the protein in the β-ribbon arm (βQ64C and βT70C) influences the orientation of HU binding to the overhang. To address this question we removed the FRET distance measurements for mutants (βQ64C and βT70C) from the dataset and plotted the FRET distances vs model predicted distances for the three other mutants (Figure 16). As shown in Figure 6-17, the best linear correlation between FRET and predicted distances for mutants in the body of HU is achieved with Model A; this model yielded a reasonable fit to experimental data with values of 0.84 for r and a \( \chi^2 \) of 0.62 (Figure 6-17). Therefore, labeling of the HU in the β-ribbon arm (βQ64C and βT70C) does not influence the binding orientation of the protein. Thus, HU appears to bind the overhang in a preferential orientation independent of overhang type. The αHU subunit binds to the junction while the βHU subunit binds in the single-strand portion of the overhang.

### 6.3.3 The model of the HU-overhang complex is consistent with other experimental data

A valid model of HU bound to the overhang must be consistent with all experimental data. It is clear from the analysis of 6-MI time-resolved fluorescence that model A accurately represents the data. For example, HU-induced perturbation to 6-MI conformational dynamics is greatest when the location of 6-MI is at the junction. Specifically, HU binding is associated with restraining 6-MI conformation dynamics and a 1.01 ± 0.056 ns increase in \( \tau_{(mol)} \) when the position of 6-MI is at the junction. The changes are an indicator of proline
Correlation between FRET distances and predicted distances from the four models of HU bound to the overhang indicates that the best representation of HU binding the overhang is model (A) of αHU binding at the junction and βHU binding in the single-strand portion. The above plots are of the FRET distances for the mutants in the body of the protein vs those predicted from the HU binding orientations. Distances and errors are listed in Table 1.0. The linear solid line in each graph depicts the perfect case of a 1:1 correspondence when the FRET distances exactly match with a particular model. Calculating the Pearson correlation coefficient \( r \), the coefficient of determination \( R^2 \), and Chi square \( \chi^2 \) between FRET and predicted model distances was used to determine the goodness of fit for each model.

**Figure 6-17** Correlation plots without Q64C and T70C FRET
intercalation and stacking with 6-MI at the junction. In addition, there is minimal change in
6-MI conformational dynamics when it is located > 7 bp from the junction and in the duplex
portion of the overhang. A sharp kink in the DNA occurs at the site of proline intercalation
altering local base microenvironment. The absence of a change in 6-MI conformational
dynamics in this region suggests HU proline does not intercalate into the duplex portion of
the junction. Another observation is HU binding alters both the mean lifetime and rotational
correlation time of 6-MI located in the single-strand portion of the overhang. In the single-
strand portion, the HU induced change to 6-MI $\tau_{(mol)}$ occurs when it is located 7 to 9 bases
from the junction, consistent with Model A: $\alpha$HU binding at the junction and $\beta$HU binding 9
bases away in the single-strand portion. A reason why the change in $\tau_{(mol)}$ is not as great
compared to the HU induced alteration at the junction is it occurs over a range of bases. It is
possible because of the inherent flexibility of the single-strand portion that HU proline can
stack with a number of bases, which means heterogeneity in the sample and not as dramatic
change in $\tau_{(mol)}$. The rotational correlation time of 6-MI located in the single-strand portion
is also altered with HU binding an indicator of a reduction in base dynamics due to
interactions between the protein’s $\beta$-ribbon arm and bases in the single-strand portion. While
the time-resolved data sets for the 5 and 3 prime overhang are different in size the trends in
the HU-induced perturbations are similar. Thus, the similar protein-induced distortion to
both 3 and 5-prime overhang suggest the HU-overhang complexes are indistinguishable.

FRET data for end labeled DNA with FAM and TAMRA is also consistent with Model
A: $\alpha$HU binding at the junction and $\beta$HU binding 9 bases away in the single-strand portion.
The data demonstrates the bend angle for the complete overhang is $\sim$120, consistent with the
crystal structure. In addition, the induced-bend is similar for both overhangs, consistent
with HU is unbiased in interacting with overhangs. It is clear from the FRET data HU does
not induce a bend greater than 30° in the duplex portion of the overhang, if HU did bend the duplex portion by 60° energy transfer would increase from 0.11 to 0.22. The absence of bending in the duplex portion rules out Models C and D, while supporting Model A. The flexibility of the single-strand portion makes it hard to confirm HU bends this portion of the overhang. It is unlikely that HU induces a single kink at the junction of 120°, indicating HU most likely introduces another kink in the single-strand portion of the overhang. It is important to note, in the duplex portion of the 5’ overhang the DNA is bent ~50° before HU binds. Analysis of this portion for theoretical curvature reveals a bend of 2°/helical turn, considerably smaller than the measured bend. This part of the sequence does contain the IHF recognition sequence and may explain the curvature of the DNA. The curvature is also observed in the nicked construct demonstrating it is not an artifact of the overhang construct. It is possible HU binds to the pre-bent duplex portion of the 5’ overhang and doesn’t induce a larger bend. Similar behavior has been observe with HU binding pre-bent A-tract sequences.175

HU affinity has similar affinity for both 3 and 5-prime overhangs. The simplest interpretation of comparable affinity is HU employs the same binding mechanism for each overhang, consistent with a single binding conformation. In addition, the structural non-specific binding affinity for single-stranded and double-stranded DNA is comparable, indicating there is little preference in terms of structural non-specific binding. Thus, there is no experimental data that argues against βHU binding in the single-strand portion of the overhang.

There is limited literature data on HU binding to the overhang construct. It was shown HU binding to the 3-prime overhang requires a single-strand length of ~8 nt to maintain high affinity.103 Assuming HU binds at the junction and in the duplex portion, the authors suggest
the 8 nt is required to interact with the body of HU. Another interpretation of the data is the 8 nt in the single-strand portion is necessary for intercalation of the HUβ subunit proline, consistent with Model A. Researchers have shown HUβ₂ binds very poorly to linear dsDNA compared to HUαβ and HUαα, consistent with the HUβ subunit binding in the single-strand portion of the overhang. In fact under stringent salt condition HUβ₂ affinity for nicked and gapped DNA is 30-fold weaker compared to HUαβ and HUαα, which supports recognition of the junction by HUα subunit. It would be informative to know the affinity of HUβ₂ for single-stranded DNA.

In the literature HU was chemically converted into a nuclease to map the nicked DNA binding site. The study used three site specific HU mutants βHU_S78C, βHU_A74C, and αHU_Q43C. Assuming a maximum cleavage distance of ~22 Å and using the IHF crystal structure to calculate the distance between the site specific mutants Cα and the C1’ of cleaved bases, the authors proposed the following model. In the model the βHU subunit binds at the junction and the location of the αHU subunit is 5-prime of the junction on the continuous strand. In this work the calculated HU induced bend to nicked DNA is ~110°. Using this information four models of HU bound to the nuclease nicked DNA construct were created (Table 6-10). The models include: Model A) the αHU subunit binds at the junction and βHU subunit binds 9bp away in the 3-prime portion, Model B) the βHU subunit binds at the junction and αHU subunit binds 9bp away in the 3-prime portion, Model C) the αHU subunit binds at the junction and βHU subunit binds 9bp away in the 5-prime portion, Model D) the βHU subunit binds at the junction and αHU subunit binds 9bp away in the 5-prime portion (Table 6-10). Shown in Table 3.0 is the calculated distance between the site specific Cα and
## Table 6-10 distance between DNA cleavage and sites site-specific mutants

<table>
<thead>
<tr>
<th>Mutants</th>
<th>Cleaved Base location</th>
<th>HU Binds to 3-prime portion of the nick</th>
<th>HU Binds to 5-prime portion of the nick</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Model A αHU at junc. (Å)</td>
<td>Model B βHU at junc. (Å)</td>
</tr>
<tr>
<td>βHU_S78C</td>
<td>-1</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td>βHU_S78C</td>
<td>-2</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>βHU_S78C</td>
<td>-3</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>βHU_A74C</td>
<td>-2</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td>βHU_A74C</td>
<td>-1</td>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>βHU_A74C</td>
<td>+1</td>
<td>28</td>
<td>23</td>
</tr>
<tr>
<td>βHU_A74C</td>
<td>+2</td>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>αHU_Q43C</td>
<td>+6</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td>αHU_Q43C</td>
<td>+11</td>
<td>23</td>
<td>32</td>
</tr>
</tbody>
</table>

5'-AGTCTAGACTGCAGTGTAGTCTTTGCTAGGAGGCATCCTCT-3'  
5'-TCAGATCTGCAGTCAACTCAGGAACGATCTGCTCTAGGGA-3'
the C1’ of the cleaved bases based on the four possible models. Based on the model
generated from FRET distance data, it was hypothesized the best fit to the nuclease data
would be with αHU at the junction. Clearly, as shown in Table 6-10 the predicted distances
are consistent with the cleavage data for model D, βHU binding at the junction and αHU
located 5-prime of the junction on the continuous strand. It is possible the discrepancy
between the two studies is related to the different constructs (overhang vs. nicked construct).
The orientation of HU with respect to the junction may depend on the type of construct or the
sequence. Another possibility for the discrepancy is the data was obtained using two
different techniques: nuclease cleavage patterns and FRET distance mapping both techniques
have been previously been used to map the binding interaction between protein and DNA.

6.3.4 Insights into the mechanism of HU specificity

The nucleoid-associated proteins HU and IHF are members of Type II DNA-binding
family employing a method of indirect readout to recognize and bind to DNA. In both IHF
and HU a conserved proline intercalates into the DNA causing kinks to form at adjacent
bases. Despite the significant sequence conservation and structural homology between the
two, the proteins identify and bind different DNA motifs. IHF binds linear DNA poorly,
instead it recognizes the DNA topology of a specific sequence interacting with nanoMolar
affinity. In contrast, HU binding is structure specific, either binding a linear nucleic acid
polymer with micromolar affinity or detecting a localized increase in nucleic acid
conformational dynamics to bind with picomolar to nanomolar affinity. In both cases binding
appears to occur at locations in the polymer with a propensity for bending. In fact,
introduction of a nick at the site in the IHF consensus sequence of proline interaction
increases protein affinity ~fourfold.333 Similarly, in this work the portion of H1 34-mer
containing the IHF recognition sequence exhibits intrinsic DNA curvature of ~60°, indicative of the topology dependent recognition (Table 6-4). Consistent with their respective crystal structures the IHF induced bend is slightly greater than HU bound to the overhang or nick (Results Table 6-4/6-5/6-6).\textsuperscript{164,165} The correlation between induced bend and affinity is worth noting. One hypothesis for the mechanism of HU specificity is increased DNA flexibility facilitates formation of salt bridges between the body of HU and the phosphate backbone stabilizing the greater protein induced distortion. In keeping with this mechanism the HU induced bend for specific binding is considerably greater ~120° compared to nonspecific binding to dsDNA with a maximum bend of ~30° bend per HU. The ~30° bend per HU is similar to the previously reported bend of 62 ± 4° for 2:1 HU binding 20 bp dsDNA.\textsuperscript{175}

HU is known to bind a variety of DNA motifs such as nicked, gapped, forked, overhang, and cruciform with high affinity. These motifs share a unique feature of greater conformational flexibility compared to dsDNA. However, conformational flexibility is not the only factor dictating HU specificity. If that was the case HU would have the highest affinity for ssDNA. Similarly, HU affinity would be reduced as the rigidity of the constructs increases. This clearly is not the case, HU has comparable affinity for the overhang and nicked DNA which display different degrees of conformational flexibility. HU specificity seems to arise from a combination of reduction in backbone strain and localized base dynamics.

Time-resolved anisotropic data indicates the dynamics of the base at the junction is between the motion of bases in single-strand and in duplex portion of the overhang (Table 6-3). In addition, the dynamics of the base linearly decreases away from the junction. This region of increased flexibility that maintains some rigidity is important for HU specificity. Previous work has shown under stringent conditions there is no binding of HU to mismatches
of 1 to 2 in a row, when the number of mismatches in a row increases to 4 HU binds with high affinity, the $K_{D_{app}} = 21$ nM. Increasing the number of mismatches is associated with greater flexibility in the construct while the surrounding dsDNA provides rigidity for intercalating proline. It is possible the intercalating proline stacks with a dynamic base which in turn is stabilized by interactions with adjacent base.

**HU activity in vivo is dependent on composition**

It is known that the expression levels of each subunit changes during the cell cycle; such that during the exponential phase HUαα is the predominate form, followed by the heterodimer during later stages of growth. Research examining the effects of ΔhupA, ΔhupB, and ΔhupAB (genes encoding the HU subunits) on genomic expression found the double mutant ΔhupAB alters the expression of 1266 genes compared to wild-type, while 512 genes change in expression in ΔhupA whereas the smallest change is observed with ΔhupB modifying only 107 genes. The α-subunit of HU appears to have a greater role in genomic regulation compared to the β-subunit. One possibility for the difference in regulation is the preferential binding of the α-subunit for sites of increased DNA distortion. For example, the overhang is known to be a critical component in DNA replication. Studies have shown only Huaβ or HUα2 play a role in DNA replication, the proteins are known to influence the unwinding of oriC (origin of replication) by DnaA. While strains deficient in hupA exhibit initiation of replication at inappropriate times in the cell cycle there is no observable defect occurs with ΔhupB mutant. In fact HUα2 stimulates the greatest amount of oriC plasmid DNA replication in vitro. Another factor distinguishing the two subunits is the inability of HUB2 to introduce negative supercoiling in the presence of topoisomerase 1.
6.4 Conclusion

HU binds with picoMolar affinity to both 5 and 3-prime overhangs in a particular orientation. The orientation is such that Hua subunit is located at the junction while HUB subunit binds in the single-strand portion of the overhang. In both overhangs HU induces a bend of \(\sim 120^\circ\). This bend angle is comparable to the bend observed in the overhang construct and in the HU-DNA co-crystal structure. The induced bend angle is considerable larger compared to the HU-induced bend in the 34-mer dsDNA. There is greater conformational dynamics for bases at junction compared to bases in the duplex. However, the flexibility is reduced compared to bases in the single-strand portion. The unique dynamics for bases in this region is possibly an important factor conferring HU specificity.
7 References


(18) Cantor, C. R.; Schimmel, P. R. *Techniques for the study of biological structure and function*; W. H. Freeman, 1980.


(24) Liang, J.; Matsika, S. *Journal of the American Chemical Society* 2011, 133, 6799.
(49) Lin, K.-Y.; Jones, R. J.; Matteucci, M. *Journal of the American Chemical Society* 1995, 117, 3873.
(81) Sherratt, D. J. *Science* 2003, 301, 780.
(89) Geiduschek, E. P.; Schneider, G. J.; Sayre, M. H. *Journal of Structural Biology* 1990, 104, 84.
(94) Craigie, R.; Arndt-Jovin, D. J.; Mizuuchi, K. *Proceedings of the National Academy of Sciences* 1985, 82, 7570.
(95) Lavoie, B. D.; Chaconas, G. *Genes & development* 1993, 7, 2510.
(112) Flashner, Y.; Gralla, J. D. *Cell* 1988, 54, 713.


(135) Prentki, P.; Chandler, M.; Galas, D. J. *EMBO J* 1987, 6, 2479.


(143) Alazard, R.; Betermier, M.; Chandler, M. *Molecular Microbiology* 2006, 6, 1707.


(159) Tanaka, I.; Appelt, K.; Dijk, J.; White, S. W.; Wilson, K. S. *Nature* 1984, 310, 376.


(254) Mathematica, W. *Wolfram Research Inc.*


Hall, K. B.; Williams, D. J. *RNA* **2004**, *10*, 34.


Bustamante, C.; Bryant, Z.; Smith, S. B. Nature 2003, 421, 423.


Hagerman, P. J. Annual review of biophysics and biophysical chemistry 1988, 17, 265.

Hagerman, P. J. Annual review of biophysics and biomolecular structure 1997, 26, 139.


