ECORI ENDONUCLEASE-DNA COMPLEXES STUDIED BY THERMODYNAMICS AND ELECTRON SPIN RESONANCE SPECTROSCOPY

by

Jacqueline Elaine Townsend

B.S. Iowa State University, 2003

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

Jacqueline E. Townsend

It was defended on

July 28, 2011

and approved by

Jeffrey Brodsky, Department of Biological Sciences

Craig L. Peebles, Department of Biological Sciences

John Rosenberg, Department of Biological Sciences

Sunil Saxena, Department of Chemistry

Dissertation Advisor: Linda Jen-Jacobson, Department of Biological Sciences

ECORI ENDONUCLEASE-DNA COMPLEXES STUDIED BY THERMODYNAMICS AND ELECTRON SPIN RESONANCE SPECTROSCOPY

Jacqueline E. Townsend, PhD

University of Pittsburgh, 2011

This work focuses on adducing general principles applicable to site-specific protein-DNA interactions by linking function to structural, thermodynamic and dynamic properties. The interaction of EcoRI endonuclease with specific, miscognate, and nonspecific DNA sequences is used as a model for protein-DNA interactions. We use four pulse Double Electron-Electron Resonance (DEER) Electron Spin Resonance (ESR) experiments to map distances and distance distributions between nitroxide spin labels placed at positions within the 'arms' and the main domain of the EcoRI homodimer. These experiments show that the DNA occupies a similar binding cleft and is enfolded by the arms of the enzyme in all three classes of EcoRI-DNA complex. Additionally, changes in dynamics of main domain and arm residues within the three complexes were explored using Continuous Wave (CW) ESR spectroscopy. A position adjacent to a protein-phosphate contact shows decreased mobility relative to other arm residues that are not at the protein-DNA interface. Signal from this position shows the largest amount of an immobile component in the specific complex, progressively less immobile in the miscognate and nonspecific complexes. This fits with distribution breadths from DEER-ESR spectra and biochemical evidence that the nearby phosphate contact is made only in the specific complex. Residues at other positions show mobilities that are in agreement with our hypothesis that residues in the arms would be relatively more mobile than those in the main domain Using Electron Spin Echo Envelope Modulation (ESEEM) ESR we show that the paramagnetic Cu^{2+} ion is coordinated by an imidazole nitrogen. These experiments thus reveal a novel metal ion

binding site. DEER measurements of distances between Cu^{2+} ions and Cu^{2+}-nitroxide distances in the homodimeric EcoRI-DNA complex establish that the Cu^{2}-coordinating residue is histidine 114, which is proximal to but not at the active site. This is consistent with our biochemical studies that show that Cu^{2+} cannot replace Mg^{2+} as a catalytic cofactor but instead completely inhibits EcoRI cleavage. We also use isothermal titration calorimetry (ITC) to directly determine a stoichiometry of two Cu^{2+} ions bound per homodimeric EcoRI-DNA complex; that is, each histidine 114 coordinates one Cu^{2+} ion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I sincerely thank my dissertation adviser, Dr. Linda Jen-Jacobson, for her wisdom, patience, and guidance through the years. Her passion for science is truly inspiring. She has created a lab environment of dedicated and rigorous research which has fostered my growth as a scientist. I will forever be grateful for what I have learned during my time in graduate school.

It has been my great pleasure to work with the members of the Jen-Jacobson lab. I particularly would like to thank Mike Kurpiewski, Dr. Lisa Engler, Dr. Paul Sapienza, Dr. Steve Hancock, and Dr. Preeti Mehta for their mentoring and friendship. I have also worked with a number of gifted undergraduate researchers and would like to specifically thank Lance Mabus, Bridget Bertoni, and Max Garber for their hard work and friendship. This dissertation originated from the work of Dr. Paul Sapienza, who kindly provided my early "hands-on" training. Lance Mabus cloned the R131C mutation and worked on the first ESR experiments. This dissertation contains data from Mike Kurpiewski's investigations into the binding of Cu^{2+} to EcoRI. Dr. Preeti Mehta developed the Cu^{2+} labeling procedure and produced the wild-type Cu^{2+} sample. She is a joy to work with and the best "lunch buddy" a gal could ask for.

The work presented in this dissertation results from an extended collaboration with Dr. Sunil Saxena and the members of his lab. This invaluable collaboration has enabled us to apply cutting-edge ESR techniques to address biophysical questions that would otherwise remain unanswered. Dr. Saxena and his lab members kindly permitted me to join their lab group meetings, where I gained a lot of knowledge in the field of spectroscopy. Dr. Katherine Stone

and Jessica Sarver collected and analyzed the DEER data presented in Chapter 2, Jessica Sarver collected and analyzed the CW data presented in Chapter 3, and Dr. Zhongyu Zhang collected and analyzed the Cu^{2+} data presented in Chapter 4. I especially acknowledge Jessica for her friendship, enthusiasm, and many helpful discussions during the hours we spent sharing coffee and pouring over our data together.

I thank all of my friends and family both within and without the Department of Biological Sciences for their support during the years. I could not have done it alone! I thank the members of the Serious Science Squad for sharing late nights, commiserations, knowledge, and ideas. I thank the members of La Belle Sultana for always keeping my spirits up with music, dance, and laughter, and I thank the members of C&D, Inc. for "diplomatic consultations".

I thank my parents, sister, and the rest of my family for their unwavering love and support. I am especially grateful to my parents for putting up with all of the "experiments" I conducted in the house when I was a kid. I love you all.

Finally, I thank my best friend and partner. James has been my rock and my knight in shining armor. He has encouraged me all through grad school, and I am deeply grateful for his strength and support.

1.0 INTRODUCTION

Protein-DNA interactions are central to many biological processes. Many of these processes, including gene regulation, genetic recombination, and transposition, require proteins to recognize specific sequences or groups of sequences. But how are proteins capable of recognizing particular sequences? By thoroughly examining systems with extremely stringent specificity we hope to elucidate the greatest number of rules that can then be applied towards an understanding of less stringent interactions. Increased knowledge of the molecular mechanisms and thermodynamic rules dictating specificity enables the development of drug design, new molecular research tools, and novel therapeutic agents.

1.1 LEVELS OF SEQUENCE SPECIFICITY IN PROTEIN-DNA INTERACTIONS

Proteins bind to DNA in order to help orchestrate a vast array of different biochemical processes. Their different functions require different levels of binding specificity. For example, structural proteins such as histones and replicative proteins such as DNA polymerases form electrostatically driven nonspecific interactions with DNA. For such proteins, it is essential to their function that they interact with all DNA sequences nonspecifically. [1] On the other end of the spectrum, regulatory proteins such as transcriptional activators and repressors must exhibit extremely high specificity in order for these proteins to bind to their target sequence among a

large molar excess of nonspecific DNA sites. For example, the paradigmatic *lac* repressor must identify one operator site among ~10^7 competing nonspecific sites. It accomplishes this via a ratio of specific to nonspecific binding constants of approximately 10^8. [1] Intermediate between these extremes are repressor proteins such as *cI* and *cro* from bacteriophage λ, which bind a series of similar operator sites in a graduated fashion referred to as "permissive discrimination." Such repressor proteins still show strong discrimination between these sites and non-specific DNA. [2]

1.2 RESTRICTION ENZYMES AS A MODEL SYSTEM FOR UNDERSTANDING STRINGENT DISCRIMINATION BY DNA BINDING PROTEINS

1.2.1 Restriction Enzymes as a class of enzymes

Restriction enzymes are a large class of enzymes whose primary function in nature is to protect the organism by cleaving foreign DNA. As discussed in a number of reviews, [3-5] restriction enzymes are ubiquitously found in bacteria and archaea, and are also found in viruses of some unicellular algae.

The restriction–modification system comprises a methylase enzyme which methylates a DNA base within a specific sequence to protect the host DNA, with a companion restriction enzyme which cleaves foreign (un-methylated) DNA at the same specific sequence. Since un-methylated DNA sequences in the cell's genome remain unprotected, it is essential for survival that the restriction enzyme has the highest level of sequence discrimination in order to prevent a potentially lethal cleavage event. For example, in the *E. coli* genome there are 645 sites which

match the specific cleavage site for EcoRI- these sites are protected by methylation. However, there are ~20,800 sites which differ from the specific sequence by only one base pair (miscognate sites) that are not protected by methylation and could potentially be cleaved. Historically, these are also referred to as "star" sites, i.e. EcoRI* site. In this work, miscognate site and EcoRI* site may be used interchangeably.

In the *E. coli* genome, there are also ~4.6×10^6 nonspecific sites (sites which differ from the specific by more than one base pair) which are not cleaved. Paul Sapienza (former lab member) has shown that the competition between nonspecific and miscognate sites for binding to EcoRI is also a significant factor in the overall discrimination against cleavage of miscognate sites *in vivo*.[6] Furthermore, unlike non-catalytic DNA binding proteins, restriction enzymes are able to exhibit discrimination in both binding and catalysis. For example, this work focuses on the restriction endonuclease EcoRI; the second order rate constant for this enzyme is 60,000 to 10^9 fold higher for cleavage at a specific site than at an EcoRI* site. [7], [8] This extremely high DNA cleavage specificity has made restriction enzymes the "workhorses of molecular biology". [9]

Restriction enzymes have been generally classified into four different subtypes: I, II, III, and IV. [3], [4], [10] The largest group of restriction enzymes is type II, which comprises most of the nearly four thousand characterized restriction enzymes. [11] This group has been further divided into a number of subgroups based on structure, recognition and cleavage mechanisms. [12] For the purposes of this work, only the type IIP, also known as "orthodox type II" restriction enzymes, will be discussed. This class is the best studied and most widely used class of restriction enzymes. [9] These enzymes are generally homodimers where each of two subunits

recognizes a single half-site of the target DNA. These enzymes recognize and cleave palindromic sequences of 4 to 8bp, using Mg^{2+} as a cofactor for cleavage. [3]

1.2.2 Restriction enzymes as a model system

An ongoing postulate of the Jen-Jacobson lab is that the rules that govern protein-DNA specificity can be understood by investigating the proteins that exhibit the greatest degree of specificity. That is, the systems which exhibit the greatest degree of specificity should reveal the greatest number of rules and constraints governing specificity. Furthermore, it has been suggested that the structural and energetic rules which apply to less specific protein-DNA interactions may be understood in terms of relaxation of the rules which apply to highly specific protein-DNA interactions. As discussed above, both restriction enzymes and transcriptional activators or repressors display extremely high degrees of specificity and have thus far revealed a great amount of information on the rules which govern specificity. Additionally, for the systems that have been thoroughly examined to date, it seems to hold that less specific protein-DNA interactions can indeed be described in terms of relaxation of the rules governing the highly discriminatory interactions. [13]

Restriction enzymes have a further advantage over transcription-regulating DNA binding proteins in that as mentioned above, restriction enzymes can be examined in terms of both their binding and cleavage properties. By withholding the required Mg^{2+} catalytic cofactor, both equilibrium binding affinities and kinetic aspects of DNA binding can be investigated for different DNA sites under a vast array of conditions. By adding the catalytic cofactor, cleavage kinetics can also be investigated. This is a valuable feature of type II restriction enzymes which

permits the separation of the overall reaction into “binding” and “cleavage” components for experimental analysis.

1.3 DETERMINANTS OF SPECIFICITY

The structural and thermodynamic properties underlying the ability of highly specific protein-DNA interactions have been the subject of great scientific interest for some time. Herein I will review some of the features that have been determined for restriction enzyme binding. These features hold true for the three restriction enzymes that have been extensively studied by the Jen-Jacobson laboratory - EcoRI, EcoRV, and BamHI. However, where appropriate I will focus on the properties of EcoRI restriction endonuclease for two primary reasons. First, the EcoRI restriction endonuclease is one of the most thoroughly studied DNA binding proteins. Investigation of this system has provided many structural and thermodynamic insights into the properties of DNA binding proteins as a whole and type II restriction enzymes specifically. Second, the data presented in this dissertation focus on EcoRI, and therefore an introduction to the properties of this enzyme is relevant to provide context for the interpretation of data presented in this work.

1.3.1 Three categories of DNA binding sites for restriction endonucleases

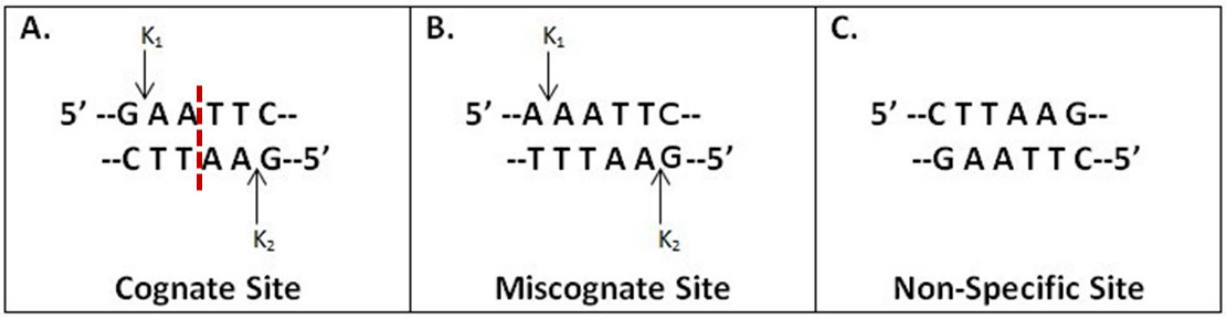


Figure 1.1 Three classes of binding site for EcoRI restriction endonuclease.

The specific (cognate) EcoRI binding site is shown in panel A. The red dashed line indicates the division of the binding site into two rotationally symmetric "half-sites". The arrows indicate the cleavage locations, k_1 and k_2 represent the cleavage rate at each half site. For the specific site these rates are equivalent. Panel B shows one of the nine possible miscognate sites, with one cognate half-site (TTC/AAG) and one miscognate half site (AAA/TTT). Panel C shows one possible non-specific site, this is the "inverted" nonspecific site which cannot make any of the normal hydrogen bonds or nonpolar interactions seen in the specific complex. [14]

The "orthodox" type IIP restriction enzymes as a class are homodimeric proteins which bind and cleave within palindromic sites of 4-8 base pairs. EcoRI is a homodimer of 31kDa subunits which hydrolyzes the bond between the G and A base pairs of the six base pair sequence GAATTC. (Figure 1.1). This site is referred to in this work as the specific or cognate binding site. The EcoRI homodimer binds to this rotationally symmetric sequence such that there are two equivalent "half-sites" (GAA base-paired with TTC) with equal cleavage rates at each binding half-site. This combined with the fact that dissociation rates for the specific complex are comparatively slow relative to the cleavage rate of the enzyme for the specific site means that a single binding event (in the presence of cofactor) generally results in a double-stranded break with four base pair 3' overhangs or "sticky ends". [15]

Miscognate sites are DNA binding sites which differ from the specific site by only one base pair. These sites contain one cognate or correct half site (GAA base-paired with TTC) and

one half site which contains the incorrect base (AAA base paired with TTT in the example in Figure 1.1). These miscognate sites disrupt the symmetry of the binding interface, resulting in what is referred to as an "adaptive" interface. In these interfaces, the cognate half site is cleaved with a faster rate than the miscognate half site. [8] Overall, miscognate sites are cleaved 60,000 to 10^9 fold more slowly than the specific site, depending on the miscognate site in question. [7], [8]

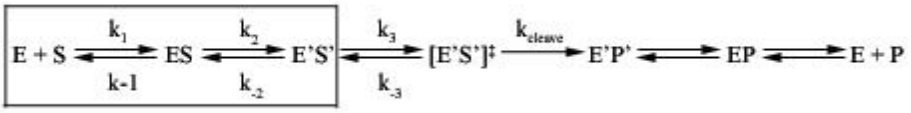
For EcoRI and BamHI, any DNA site which differs from the specific sequence by one base-pair is referred to as a "miscognate" site. For these enzymes, any DNA site that differs from the specific sequence by more than one base pair cannot be cleaved and is referred to as a "nonspecific" site as described below. For EcoRV, however, in addition to the miscognate sites that differ from the specific site by one base pair there is also one known sequence which differs from the specific site by both central base pairs but can still be cleaved. The central TA base pairs are inverted in this sequence, and crystallographic evidence indicates that despite this inversion, these bases are still oriented such that the van der Waals contacts normally made in the specific complex are preserved. However, in this case catalytic efficiency is reduced by a factor of 10^4 . [16] In general, DNA sites that differ from the specific site but can be cleaved at a reduced rate are referred to as miscognate sites, including the rare cases which differ from the specific site by two base pairs.

The term "nonspecific sites" is used in this dissertation to refer to DNA binding sites which differ from the EcoRI specific site by two or more base pairs and are completely resistant to cleavage. In the case of EcoRI, all DNA binding sites which differ from the specific by more than one base pair are resistant to cleavage. [8] In more general terms, "nonspecific binding" is used to describe DNA for which a site-specific DNA-binding protein interacts in a non-localized

manner. In other words, the protein binds to these sequences with equal probability and the interaction does not result in a detectable "footprint" when investigated with techniques such as ethylation interference footprinting. [7], [14], [17], [18] Through this work, the term "noncognate" will be used to refer to both the miscognate and nonspecific sites, while the term "cognate" will be used interchangeably with specific.

1.4 THERMODYNAMIC AND KINETIC PARAMETERS OF ECORI SPECIFICITY

The affinity of EcoRI for a given DNA sequence can be described in terms of the binding association constant (K_A) or the standard free energy change of the reaction. For a protein-DNA binding interaction:



Formation and cleavage of a DNA complex proceeds through a number of steps. The first intermediate, ES, is a nonspecific "collision complex" that is believed to progress across DNA in a linear facilitated diffusion. If a specific DNA binding site is encountered, the complex undergoes further changes to form the recognition or E'S' complex.

Several lines of evidence indicate that the specific recognition complex is a "pre-transition state" complex. Many energetic factors have been shown to make equal contributions to the E'S' complex and the transition state ($E'S'^{\ddagger}$) for EcoRI. [7] Additionally, if Mn^{2+} is diffused into crystals of the specific complex, the DNA is cleaved without shattering the crystal; indeed, the post-reactive complex shows only minor deviations from the specific complex. [19]

Furthermore, molecular dynamics simulations of EcoRI in the presence and absence of computationally added Mg^{2+} only show minor rearrangements of side-chains and solvent molecules in the interface. [20] The E'S' complex is therefore predicted to have a structure extremely close to that of the transition state complex and is referred to as the "pre-transition state" complex. [7]

In the absence of divalent metal, the reaction is confined to the boxed portion of the diagram above. Binding affinities are measured in terms of the equilibrium association constant, K_A:

$$K_A=[ES]/[E][S]$$

The association constant as measured is actually an apparent constant (K_{Obs}) that comprises contributions from both specific binding (K_S) and nonspecific binding (K_{NS}), as shown below. [21], [22] Since each base pair represents the start of a potential nonspecific binding site, K_{NS} is multiplied by the number of base pairs in the sequence (N_{NS}):

$$K_{Obs}= K_S + N_{NS}*K_{NS}$$

However, for our system $K_S >>> K_{NS}$, the ratio of K_S/K_{NS} for EcoRI is on the order of $4.7x10^7$. It should be noted, however, that the precise ratio is highly dependent on the experimental conditions. [7] Therefore, K_{Obs} is essentially equal to Ks when examining a specific substrate, given that long DNA substrates are avoided in order to minimize the contribution from K_{NS}. [14] The stability of complexes is frequently described in terms of the standard Gibbs free energy change, where R is the gas constant and T is temperature:

$$\Delta G°_{bind}= -RT\ln K_{Obs}$$

The above two equations describe the binding affinity of the enzyme for a DNA substrate. It is crucial to note that binding affinity does not equate to binding specificity,

although high affinity and high specificity are generally related. Binding specificity is described in terms of the binding specificity ratio, defined as K_S/K_{NS}. Alternatively, the binding specificity can also be described in terms of the difference in free energy changes between the two different complexes:

$$\Delta\Delta G^\circ_{bind} = (\Delta G^\circ_S - \Delta G^\circ_{NS}) = RT\ln(K_S/K_{NS})$$

Similarly, the differences in binding free energy can be calculated for the enzyme binding to different DNA sites, such as two different miscognate sites, or the same site embedded in different flanking contexts. They may also be used to calculate the difference in binding energy between the wild-type enzyme and a mutant enzyme for the same site as shown below:

$$\Delta\Delta G^\circ = (\Delta G^\circ_{Mut} - \Delta G^\circ_{WT}) = -RT\ln\frac{K_{Mut}}{K_{WT}}.$$

Or:

$$\Delta\Delta G^\circ = (\Delta G^\circ_{Site1} - \Delta G^\circ_{Site2}) = -RT\ln\frac{K_{Site1}}{K_{Site2}}.$$

As discussed previously, restriction enzymes such as EcoRI, EcoRV, and BamHI have among the highest reported specificity ratios, on the order of 10^5-10^7. [7] As binding constants are sensitive to experimental conditions including pH, temperature, ionic strength, solution ions, cosolutes, and the DNA sequence flanking the binding sequence, it is critical to choose experimental conditions for determining constants carefully. [23]

1.4.1 Energetics of sequence discrimination by EcoRI

The cleavage of DNA by EcoRI is a reaction comprising many steps, including nonspecific binding, reaction intermediates, and product release. In order to focus on the energetics of

sequence discrimination itself, Jen-Jacobson and coworkers [8] developed a way to measure the transition state interaction free energy ($\Delta G^{\circ}_{I}{}^{\ddagger}$) of DNA cleavage by restriction enzymes. This is a measure of the probability of forming the activated complex from the free enzyme and a given substrate. This measure is defined as:

$$\Delta G^{\circ}_{I}{}^{\ddagger} = \Delta G^{\circ}_{bind} + \Delta G^{\circ\ddagger}$$

Where ΔG°_{bind} is as defined earlier and $\Delta G^{\circ\ddagger}$ is the standard free energy of activation for the complex:

$$\Delta G^{\circ\ddagger} = RT\ln(kT/h) - RT\ln k_c$$

where k_c is the first order cleavage rate for the first bond-breaking step, k is the Boltzman constant, h is the Plank constant, and T the temperature in Kelvin. By determining the binding constants (K_{A1} and K_{A2}) and cleavage rates (k_{c1} and k_{c2}) for two sites, the overall discrimination between a reference and variant site can thus be measured as:

$$\Delta\Delta G^{\circ}_{I}{}^{\ddagger} = - RT\ln[(k_{c1} * K_{A1})/(k_{c2} * K_{A2})]$$

Jen-Jacobson and colleagues [8] measured the binding affinities, first order cleavage rate constants, and transition state interaction free energy for all nine possible EcoRI miscognate sites and four of the possible nonspecific sites (Table 1.1). Single base pair substitutions were shown to incur significant penalties to free energies of formation ($\Delta\Delta G^{\circ}$ = 4.1 to 5.7 kcal/mol) and transition state complex formation ($\Delta G^{0}_{I}{}^{\ddagger}$ = 6.6 to 13 kcal/mol). These penalties are large in contrast to those incurred by base analog substitutions ($\Delta G^{\circ}_{I}{}^{\ddagger}$=0 to 2.5 kcal/mol) which remove a single hydrogen bond. [7] At most, a single base pair change that disrupts two hydrogen bonds is predicted to incur 3kcal of transition state interaction free energy. Therefore, these authors proposed that the energetic penalty of miscognate binding results from a combination of

penalties from disrupting protein-base contacts ($\Delta\Delta G_{Base}$), protein-phosphate contacts ($\Delta\Delta G_{Phos}$), and conformational and entropic factors ($\Delta\Delta G_{Reorg}$). [8]

Table 1.1 Changes in free energies of formation of EcoRI-DNA complexes and transition state complexes for various DNA sites.

Site	K_A* (M^{-1})	$\Delta\Delta G°_{ED}$ § (kcal/mol)	k_1 (sec^{-1})	Relative $(k_1 \times K_A)$ ‖ $(M\text{-}sec)^{-1}$	$\Delta\Delta G°_1$‡ (kcal/mol)
GAATT C CTTAA↑G	$3.3(\pm1.6)\times10^{11}$	0	$3.4(\pm0.7)\times10^{-1}$	1	0
		First-base substitutions			
TAATT C ATTAA↑G	$3.1(\pm1.6)\times10^{8}$	4.1±0.5	$5.0(\pm0.3)\times10^{-3}$	1.5×10^{-5}	6.6±0.5
AAATT C TTTAA↑G	$1.8(\pm1.0)\times10^{8}$	4.4±0.5	$8.4(\pm1.7)\times10^{-4}$	1.4×10^{-6}	8.0±0.5
CAATT C GTTAA↑G	$5.3(\pm2.2)\times10^{7}$	5.2±0.4	$2.0(\pm0.7)\times10^{-4}$	1.0×10^{-7}	9.5±0.5
		Second-base substitutions			
GCATT C CGTAA↑G	$4.4(\pm1.7)\times10^{7}$	5.3±0.4	$6.7(\pm0.5)\times10^{-5}$	2.7×10^{-8}	10.3±0.4
GGATT C CCTAA↑G	$2.6(\pm1.7)\times10^{7}$	5.6±0.5	$6.2(\pm1.5)\times10^{-6}$	1.5×10^{-9}	12.0±0.6
GTATT C CATAA↑G	$1.6(\pm0.8)\times10^{8}$	4.5±0.5	$\sim2\times10^{-7}$	2.9×10^{-10}	~13
		Third-base substitutions			
GACTT C CTGAA↑G	$5.9(\pm0.6)\times10^{7}$	5.1±0.3	$4.5(\pm0.6)\times10^{-4}$	2.5×10^{-7}	9.0±0.3
GAGTT C CTCAA↑G	$1.1(\pm0.9)\times10^{8}$	4.7±0.6	$8.9(\pm0.3)\times10^{-5}$	8.9×10^{-8}	9.6±0.3
GATTT C CTAAA↑G	$3.4(\pm0.7)\times10^{7}$	5.4±0.4	$2.0(\pm0.5)\times10^{-6}$	6.2×10^{-10}	12.5±0.4
		Double substitutions¶			
GGATCC	$9.7(\pm4.8)\times10^{7}$	4.8±0.5	No cleavage		
GAGCTC	$4.3(\pm2.0)\times10^{7}$	5.3±0.4	No cleavage		
AAATTA	$2.0(\pm1.0)\times10^{7}$	5.7±0.5	No cleavage		
		Inverted site			
CTTAAG	$4.7(\pm2.6)\times10^{7}$	5.2±0.5	No cleavage		

***K_A values measured in binding buffer of 10mM bis-tris-propane, 0.95mM NaCl, 5µm dithiothreitol, 1mM EDTA, 50mg/ml bovine serum albumin, pH 7.5, at 25°C.**
§ Calculated from the ratio of K_A values between the variant site and the canonical GAATTC site: $\Delta\Delta G°_{ED} = -RT\ln(K_{Avariant}/K_{Acanonical})$.
‖ This ratio measures discrimination relative to GAATTC.
¶ These are examples of nonspecific sites.
Table taken directly from[8].`

1.4.2 The molecular origins of entropy and enthalpy in protein-DNA binding

In more detailed terms, $\Delta G°_{bind}$ represents the sum of the thermodynamic components contributing to the stability of a non-covalent binding complex under a given set of conditions and can be written in the following equation:

$$\Delta G°_{bind} = \Delta H° - T\Delta S°$$

In the above equation $\Delta H°$ and $\Delta S°$ represent the enthalpic and entropic contributions to the change in free energy upon complex formation. Under a specific experimental condition, these parameters represent the differences in final state between the free protein and DNA compared to the state of the protein-DNA complex. Importantly, these differences derive not only from changes within the protein and DNA molecules, but also from the solvent molecules and interactions between the solvent and macromolecules.

The origins of favorable enthalpy include intermolecular interactions such as hydrogen bonds, van der Waals, and electrostatic contacts. Upon formation of the complex, molecular strain and burial of polar surfaces make unfavorable contributions to enthalpy. [7]

A major contributor to favorable entropy upon substrate binding results from the hydrophobic effect. Water tends to form ordered "cages" around nonpolar surfaces. The release of such ordered water molecules from nonpolar surfaces has highly favorable entropy. This favorable entropy provides a major driving force for protein folding. The release of water from nonpolar surfaces is also a major driving force for protein-DNA interactions. [24] [25]

Upon EcoRI specific complex formation, ~2500Å of nonpolar surface area is buried, which would result in the release of large numbers of ordered water molecules. [7] This release of bound waters is a major source of favorable entropy contributing to the overall binding free energy for EcoRI. In an elegant treatment of the subject, Jen-Jacobson and Jacobson [23] have

estimated the contribution of the hydrophobic effect on binding, ($\Delta G°_{HE}$) as ~80 $kcalmol^{-1}$ of favorable energy to the binding of EcoRI to the specific sequence. Considering that the net $\Delta G°_{bind}$ of EcoRI to the specific site is only -15 $kcalmol^{-1,}$ it is apparent that this effect makes a large contribution to the net thermodynamics of binding.

Complex formation also results in restriction of the rotational and translational motions of the macromolecules, causing an unfavorable entropic contribution to the net $\Delta G°_{bind}$ [25] as well as unfavorable entropic contributions from the loss of configurational (conformational plus vibrational) entropy. [26] These contributions are discussed in more detail in subsequent sections.

Overall, it can be seen (as summarized in Figure 1.2) that specific binding is the net result of large opposing forces. Specific site binding energy is largely driven by favorable entropy deriving from the release of ordered water molecules from nonpolar surfaces in the protein-DNA interface, as well as favorable enthalpic contributions from the formation of contacts between the protein and DNA. These favorable forces are opposed by the loss of rotational and translational freedom of the protein and DNA, as well as the energetic costs of DNA distortion, and restrictions on the configurational modes of the molecules. In contrast, nonspecific binding is the result of opposing forces of smaller magnitude and is largely driven by favorable enthalpy deriving from electrostatic interactions between the protein and phosphates of the DNA backbone opposed by restriction of translational and rotational entropies.

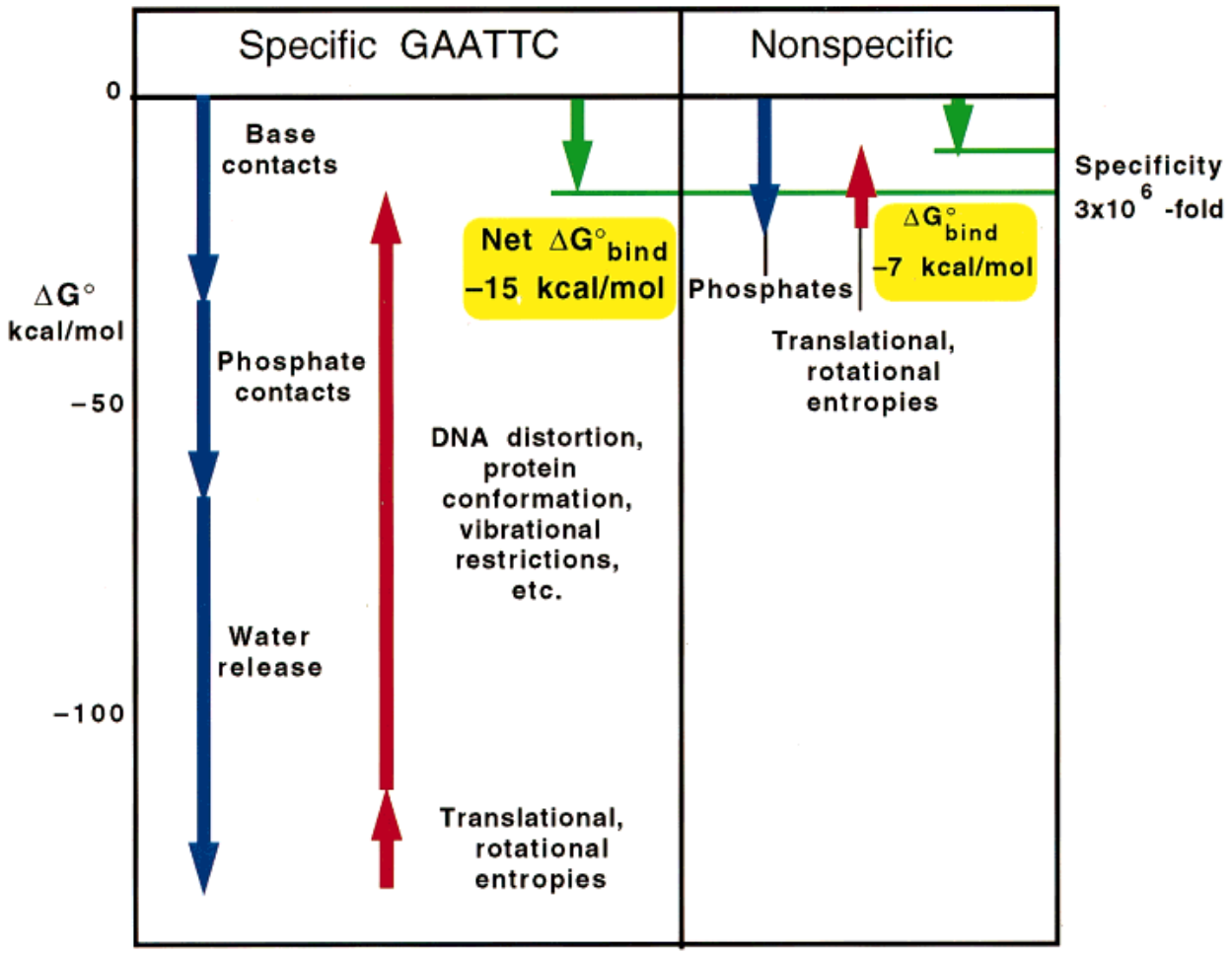


Figure 1.2 Energetic components of specific and nonspecific binding of EcoRI endonuclease.

Figure taken directly from [7]. Favorable contributions are shown with blue arrows, unfavorable contributions with red arrows, and observed net ΔG°_{bind} with green arrows. All values are estimated and may deviate from the true values by ±50% as described. [7] The "DNA distortion, protein conformation, vibrational restrictions, etc." category is estimated from the difference between the sum of the other contributions and ΔG°_{bind}. Although these contributions are divided into separate components for the purposes of this figure, they are not independent of each other.

1.4.3 Heat capacity change ΔC°p and its molecular origins

1.4.3.1 Definition of heat capacity change in relation to entropy and enthalpy

Both the enthalpic and entropic aspects of thermodynamics are related to the heat capacity (C) of a system. As reviewed in Prabhu and Sharp [27] heat capacity (C) is defined as the relationship between the amount of heat added to a system and the temperature change of the system. At constant pressure, the heat capacity (Cp) describes the overall temperature dependence of the enthalpy and entropy as shown in the following equations:

$$Cp = (\partial H/\partial T)p$$

$$Cp = T(\partial S/\partial T)p$$

Alternatively, the heat capacity can be defined in relation to the mean squared fluctuation in enthalpy or entropy as shown below where k is the Boltzmann constant and T is the absolute temperature. :

$$Cp = \langle \delta H^2 \rangle / kT^2$$

$$Cp = \langle \delta S^2 \rangle / k$$

These relationships can be used to determine the following equation which relates heat capacity to the mean fluctuation of both the entropy and enthalpy of the system:

$$Cp = \langle \delta H \delta S \rangle / kT$$

All of these equations are equivalent, such that:

$$Cp = T\frac{dS}{dT} = -T\frac{d^2G}{dT^2} = \frac{\langle \delta H^2 \rangle}{kT^2} = \frac{\langle \delta S^2 \rangle}{k}$$

The contributions to heat capacity and thus to changes in the heat capacity of the system that occur upon ligand binding ($\Delta C°_P$) as discussed in subsequent sections thus originate on the

molecular level from changes in the mean fluctuation of the entropy and enthalpy of the system that occur upon ligand binding. In our case, this the binding of DNA by EcoRI.

Put another way, the heat capacity change of the system, $\Delta C°_{P,}$ (such as the change in heat capacity upon formation of a protein-DNA complex) describes the temperature dependence of the change in enthalpy and entropy for the formation of the complex:

$$\Delta C°_P = (\partial \Delta H°/\partial T)_P$$

$$\Delta C°_P = T(\Delta \partial S°/\partial T)_P$$

If the heat capacity change of the system is nonzero and independent of temperature in the experimental range ($\Delta C°_P$ is approximately constant over small temperature ranges), then $\Delta C°_P$ can also be used to determine the dependence of ΔH and ΔS as a function of temperature. Given the values T_H and T_S for a particular system such as protein-DNA binding, where $\Delta H°=0$ at T_H, and $\Delta S° = 0$ at T_S, :

$$\Delta H° = \Delta Cp° (T-T_H)$$

$$\Delta S°= \Delta Cp° \ln (T/T_S)$$

At temperatures below T_H, binding is entropy-driven, while at temperatures above T_S binding is enthalpy-driven. Both $\Delta H°$ and $\Delta S°$ are favorable at temperatures between T_H and T_S. When discussing the enthalpic and entropic contributions to a system it is therefore always important to specify the temperature. Generally speaking, these terms tend to balance out in an effect known as temperature dependent entropy-enthalpy compensation, such that overall $\Delta G°_{bind}$ tends to be fairly resistant to temperature changes, especially at temperatures near T_S. This phenomenon is commonly observed for protein-DNA associations, because T_S generally falls in a physiologically relevant range. [13]

1.4.3.2 Molecular origins of $\Delta C°p$ in macromolecular processes

In a classic paper on the subject, Sturtevant [28] outlined six potential sources for the large heat capacity and entropy changes observed for macromolecular association processes. Of these, changes in equilibria, hydrogen bonding, and electrostatic charges were predicted to make only minimal net contributions. Key sources of both heat capacity and entropy changes for protein interactions were proposed to be the hydrophobic effect and loss of configurational/vibrational freedom, while changes in conformational entropy were proposed to provide a major source of entropy changes without contributing directly to heat capacity changes. [28]

For protein folding, the large negative heat capacity change upon folding, $\Delta Cp°_{fold}$ has been shown to correlate strongly with the burial of nonpolar surface area, using models in which the burial of nonpolar surface area makes a large negative contribution to $\Delta Cp°_{fold}$ and in which burial of polar surface area makes a small positive contribution to $\Delta Cp°_{fold}$. [29-31] However, in a large study of nearly 50 proteins, Robertson and Murphy [32] found that $\Delta Cp°_{fold}$ correlates equally well with the total buried surface area as well as it does with buried nonpolar surface area, indicating that while the large negative $\Delta Cp°_{fold}$ observed for protein folding seems to result primarily from the hydrophobic effect, other types of molecular interactions may also contribute. [32]

1.4.3.3 A large negative $\Delta Cp°$ is observed for specific protein-DNA complex formation

It has been shown that a large negative $\Delta Cp°$ is also a thermodynamic signature of specific protein-DNA complex formation. In contrast to this, nonspecific binding shows a heat capacity change of approximately zero. This observation holds true for several well-studied systems including the restriction enzymes EcoRI, [24], [26] BamHI, [26] and EcoRV [33]. Additionally,

the repressor proteins cro, [34] λ cI, [35] and trp [36] also show a large negative ΔCp° upon binding to their target sites and a ΔCp° ~ 0 for nonspecific complex formation.

Given the strong correlations between heat capacity change and the burial of nonpolar surfaces for protein folding, Record and coworkers suggested [24], [25] that the large negative ΔCp° observed for specific protein-DNA complex formation could also be explained primarily in terms of the changes in water-accessible surface areas. This model assumed that the binding of the specific DNA sequence caused induced folding of the protein regions which contact the DNA. They used this model in order to account for the observed ΔCp° values, which were too large to be accounted for by the surface area changes produced by a simple docking of the molecules. [25]

However, later work revealed that for at least some DNA binding proteins such as TATA binding protein, the observed structural changes in protein conformation upon binding are far from sufficient to explain the magnitude of the heat capacity change that is observed. [37] This heat capacity change "deficit" was also observed for other DNA-binding proteins including trp repressor, MetJ repressor, Gcn4, and the glucocorticoid receptor. [36], [38-40] Based on detailed studies of the trp and MetJ repressors, [36], [40] Ladbury and coworkers proposed that for these proteins the restricted freedom of water molecules bound in the interface may make a significant contribution to the ΔCp° which partly accounts for the "deficit". Additionally, these researchers [37], [40] suggested that the observed ΔCp° "deficit" for specific protein-DNA interactions could at least partly be the result of a stiffening of the internal degrees of freedom for the protein, as originally suggested by Sturtevant. [28] Support for this hypothesis comes from the observation that the crystallographic temperature (B) factors for MetJ and the trp repressors are reduced in the bound state, indicating a "tightening" of the structure. Although B factor changes

should be interpreted cautiously, since they are dependent on the form and packing of the crystal, this observation is consistent with the stiffening that would be predicted for the bound state. [40]

1.4.3.4 Configurational entropy in proteins

Additional support for the role of configurational entropy in protein interactions comes from more general experimental and theoretical treatments of the subject. Karplus and coworkers [41], [42] demonstrated that the residual configurational entropy of a protein is an order of magnitude greater than that of the change in configurational entropy upon protein denaturation (ΔS_{conf}). Historically, ΔS_{conf} was estimated by setting the internal vibrational entropy (S^v_p) of the folded protein equal to zero and calculating the configurational entropy change from the random-coil (denatured) state (S_{rc}) to the folded state in a way that ignored the potential contribution from the vibrational entropy. This value has generally been in agreement with experimental values obtained at room temperature for protein folding. [43] However, Karplus and coworkers found that for the protein bovine pancreatic trypsin inhibitor (BTPI), the residual configurational entropy for the protein (S^v_p) was an order of magnitude larger than ΔS_{conf}. [42] Given that S^v_p was so large, Karplus and coworkers explored how it was possible that experimentally determined ΔS_{conf} values generally agree with experimental values. [41]

The equation describing the total configurational entropy of a protein includes contributions from both the local fluctuations within a given conformation (described in the first term of the equation below) as well as the existence of multiple conformations (described in the second term of the equation below):

$$S_{\text{conf}} = \sum_{I=1}^{N} \omega_I S_I^v - k_B \sum_{I=1}^{N} \omega_I \ln \omega_I$$

Where the ω_I represents the Boltzman weighting factor of an energy well *I* with vibrational entropy S^v_I, and k_B is the Boltzman constant. The change in configurational entropy upon protein denaturation can be written as:

$$\triangle S_{\text{conf}} = S_{rc} - S_p^v = \left(\sum_{I=1}^{N} \omega_I S_I^v - k_B \sum_{I=1}^{N} \omega_I \ln \omega_I\right) - S_p^v$$

Since the experimentally determined change in configurational entropy upon protein folding generally matches that obtained from forms of the equation which ignored the first term, even though it was shown to be much larger than the second term, it was concluded that the residual vibrational entropy of a folded protein must be roughly the same as that of any given conformation of the unfolded state, such that ΔS^v_p for protein folding is minimal. This is shown in the following two equations. This would explain why the previous approximation for the *change* in configurational entropy matched experimental results even though it was missing an important contribution to the *total* entropy of the protein. [41]

$$\sum_{I=1}^{N} \omega_I S_I^v \simeq S_p^v$$

$$\triangle S_{\text{conf}} \simeq - k_B \sum_{I=1}^{N} \omega_I \ln \omega_I$$

Their work also showed that while ΔS^v_p should be minimal for protein *folding*, the magnitude of ΔS^v_p for protein *binding* to a ligand can be as significant as the hydrophobic effect to the overall ΔS of binding. [41], [43] The loss of configurational freedom is therefore shown to be a potentially important contribution to the overall entropy change of protein binding, and thus to heat capacity changes, as suggested by previous researchers. [28]

1.4.3.5 Contributions to the large negative ΔCp° observed for EcoRI binding to its target site

Similarly to the DNA binding proteins discussed in previous sections, the EcoRI and BamHI restriction enzymes exhibit a large negative ΔCp° for formation of their specific complexes and a negligible ΔCp° for formation of the nonspecific complex, shown in Table 1.2. [26] As described in a previous section, the major contributions to heat capacity for protein interactions are proposed to be the hydrophobic effect and loss of configurational (conformational-vibrational) freedom. [28] The loss of freedom upon protein-DNA complex formation can include the restriction of interfacial water molecules as well as of the protein sidechains, DNA bases, and backbone fluctuations. The formation of nonspecific protein-DNA complexes such as those for EcoRV [45] and the glucocorticoid receptor [45] do not result in the burial of much surface area. Furthermore, determination of the EcoRI and BamHI binding dependence on cosolute concentration shows that the desolvation upon binding is much greater for the specific complexes than for the nonspecific complexes. [26] Molecular dynamics simulations of the specific and nonspecific EcoRV complexes found greater motion for the protein and the DNA in the nonspecific complex. [46] In the nonspecific complexes, the negligible ΔCp° reflects a minimal change in solvation or restriction of the protein or DNA upon complex formation. [26]

In contrast to the negligible ΔCp° values observed for nonspecific DNA binding, a large negative ΔCp° is observed for specific DNA binding. The magnitude of this ΔCp° cannot be explained by burial of solvated surfaces alone. For EcoRI, the greatest predicted ΔCp° value calculated from changes in surface solvation results in an estimate of -0.3 kcal/mol*K, while the experimentally determined value for the "worst" flanking sequence is -1.2kcal/mol*K. [23], [26] The contribution to ΔCp° from changes in surface accessibility, therefore, are predicted to be

approximately a third or less of the total. Based on upper limit estimates, trapped interfacial waters are predicted to contribute up to -0.4 kcal/mol*K to ΔCp°, although this is likely an over-estimate of the contribution. However, a large deficit of unaccounted for ΔCp° remains which is likely primarily the result of restricted conformational-vibrational degrees of freedom in the complex. [26] This prediction is supported by the experiments described below.

In order to further investigate the contributions to ΔCp° for these complexes, Jen-Jacobson and coworkers [26] found that by changing the three non-contacted base pairs surrounding the specific binding site (also referred to as the "flanking" sequences), the observed ΔCp° varied significantly. By investigating the thermodynamic parameter changes for differing flanking sequences, it was therefore possible to investigate other contributions to the overall ΔCp° while holding the contributions from surface area changes constant. A striking finding from this analysis was that as the flanking context improved (better $\Delta G°_{bind}$) the ΔCp° values became more negative. This effect is quite dramatic; for EcoRI the magnitude of the negative ΔCp° for the "best" flanking sequence is nearly double that of the "worst" sequence (Table 1.2). BamHI shows a similar trend. These trends cannot be the result of the hydrophobic effect, for if this were so, the ΔS° values should become more favorable with more negative ΔCp°, however the opposite trend is seen. Furthermore, for the hydrophobic effect to contribute to the differences in ΔCp° observed for the various flanking sequences, these complexes would have to exhibit differing degrees of desolvation. However, cosolute studies with triethylene glycol (TEG) have shown that the binding dependence on TEG concentration is the same for the different flanking contexts, indicating that these complexes undergo the same amount of desolvation. [23]

It could be argued that the different ΔCp° values observed for the various flanking contexts may be explained by the formation of different protein-base or protein-phosphate

contacts at the recognition interface; however, several lines of evidence indicate that these contacts are unaltered for the different flanking contexts. [26] First, comparison of crystal structures obtained for the EcoRI-specific complex with two different flanking contexts revealed that there are no significant differences in protein-base and protein-phosphate contacts, and that there are no contacts formed directly between the protein and the base pairs outside the recognition sequence. [47] Second, the ethylation-interference footprints for the different flanking sequences are identical, indicating that the protein-phosphate contacts are the same. [8] Finally, the cleavage rate constants, which are an extremely sensitive measure of alterations in the binding interface, [48] are the same for all of the flanking contexts. [26] The only exception to this is the A-tract flank, which is known to be unusually rigid. [49] Other potential sources of heat capacity change could include linked equilibria such as proton dissociation and cation release. [50], [51] However, proton dissociation equilibria are expected to make a negligible contribution to heat capacity. [50] In the case of EcoRI binding to different flanking contexts, all of the flanking contexts have identical salt-dependence of equilibrium binding ($d\log K_{obs}/d\log[NaCl]$), indicating that these complexes release the same number of cations, so that the differences in heat capacity change for the different flanking contexts is cannot result from differential cation release. [26] Taken together, these results indicate that the specific complexes with varying flanking sequence form interfaces that are approximately the same in terms of the contact formation and desolvation of the interface.

Overall, it was determined that changes in vibrational-conformational freedom at the protein-DNA interface must be a major contributor to the large negative $\Delta Cp°$ which is seen for the formation of the specific complexes. This is consistent with the observed data, as this factor would be expected to make both $\Delta Cp°$ and $\Delta S°$ values more negative as $\Delta G°$ becomes more

favorable. This conclusion, combined with the trends seen for ΔH° and ΔS° as the flanking context change, suggest a model for these effects. For more favorable sequence contexts (more negative ΔG°) the protein and DNA are more immobilized (ΔCp° and ΔS° become more negative). Meanwhile, ΔH° is more favorable because the complexes "fit better" and therefore the unfavorable enthalpic contribution from molecular strain is reduced. [26] For the miscognate EcoRI-DNA complexes, the heat capacity changes were found to be intermediate between those of the specific and nonspecific complexes (in the range of -0.1 to -1.0 kcal/mol*K), supporting the overall observations that these complexes are intermediate in characteristics between the specific and nonspecific complexes. [47]

Table 1.2 Thermodynamic signatures of specific and nonspecific binding for EcoRI and BamHI complexes.

	EcoRI Endonuclease[a]				*BamHI Endonuclease*[b]			
	"Best"	*"Medium"*	*"Worst"*	*Nonspecific*[c]	*"Very Good"*	*"Medium"*	*"Worst"*	*Nonspecific*[c]
ΔG° bind[d] kcal/mol	−14.6 ± 0.1	−12.3 ± 0.3	−11.1 ± 0.4	− 6.8 ± 0.3	−12.7 ± 0.2	−12.4 ± 0.1	−10.9 ± 0.1	−7.5 ± 0.1
ΔC°_p [e] kcal/mol·K	− 2.5 ± 0.3	−1.5 ± 0.2	− 1.2 ± 0.3	0	−1.5 ± 0.1	− 1.3 ± 0.1	−1.2 ± 0.1	0
ΔH° [e] kcal/mol	− 9.7 ± 0.4	−0.3 ± 2.5	+ 9.3 ± 3.5	−14.2 ± 1.5	−16.6 ± 1.7	− 6.5 ± 1.0	−1.9 ± 0.9	− 7.4 ± 0.3
TΔS° [e] kcal/mol	+ 4.9 ± 1.0	+12.0 ± 4.1	+20.4 ± 5.8	− 7.4 ± 1.5	− 4.0 ± 0.7	+ 5.2 ± 0.7	+9.0 ± 3.0	+0.1 ± 0.2

Figure taken directly from [26].
a,b,c- K_A values determined for DNA sequences and under buffer conditions as described in [26]**. "Best, Medium, Worst" etc. refer to the relative binding constants for different flanking sequences.**
d- calculated from $\Delta G^\circ = -RT\ln K_A$ at 298K.
e- calculated from van't Hoff plots as described in Figure 1. [26]

1.5 STRUCTURAL FEATURES OF ECORI

As part of investigating the molecular origins of protein-DNA specificity, it is necessary to examine the structural features of protein binding enzymes and their complexes. For this reason, the crystal structures of restriction enzymes have been of some interest. For the ~4000 known restriction enzymes, over a hundred crystal structures have been solved, representing 40 distinct enzymes. [11] In the following section I describe the features of EcoRI structure for the specific complex and free enzyme.

1.5.1 General features of the EcoRI-specific complex

The structural features of EcoRI have been of interest for some time. The original specific complex structure was solved by X-ray crystallography at 3.0 Å and was the first protein-DNA complex to be characterized at atomic resolution. [52] The resolution was later improved to 2.7Å. [53] More recently, Arabela Grigorescu [47] generated a highly refined version of the EcoRI specific complex with resolution at 1.9 Å (refined version of PDB:1CKQ). Unless stated otherwise, all images of the EcoRI specific complexes presented in this work are based on those coordinates.

EcoRI forms a homodimer of 31kDa subunits. This complex displays perfect twofold symmetry: the asymmetric unit of the crystal contains one protein monomer and one strand of DNA. Two different views of the structure are shown in Figure 1.3. The GAATTC binding site is shown in orange, while the rest of the DNA is in green. The protein contains a large globular "core" domain (shown in grey). A notable feature of EcoRI is the two long chains which extend away from the rest of the protein and embrace the DNA. These protrusions are historically

referred as the “inner” (magenta) and “outer” (cyan) arms of the complex, but may also be viewed as an “arm” and “hand” (personal communication, J. M., Rosenberg).

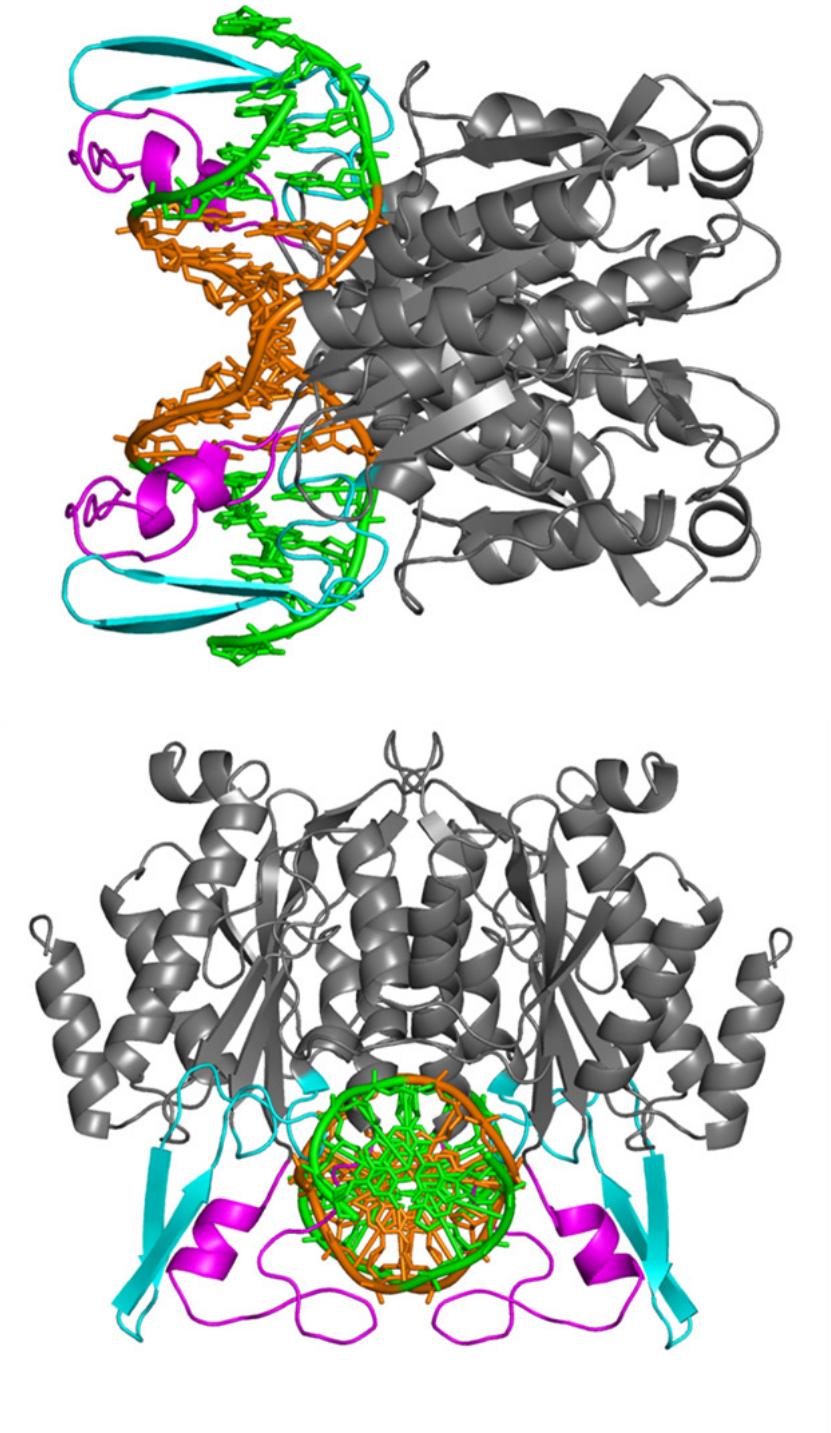


Figure 1.3 EcoRI-specific complex.

Two views of the EcoRI-specific DNA complex. GAATTC binding site is shown in orange, "inner" arm aka hand region in magenta, "outer" arm aka arm in cyan. Based on a highly refined version of PDB entry 1CKQ. [47]

1.5.2 Base or "direct" readout

A notable feature of specific protein-DNA complexes is that of intricate complementarity between the protein and the DNA bases and backbone. This complementarity is largely facilitated by extensive and highly cooperative hydrogen bonding interactions between the protein and DNA. Of the >1500 crystal structures of protein-DNA complexes in the Protein Data Bank, almost all of them exhibit evidence of a base specific pattern of hydrogen bond donors and acceptors, as well as nonpolar groups, interacting with a complementary set of amino acids. (Reviewed in [54]). Historically, this set of interactions has been referred to as the "direct readout" interactions. However in a recent review, Rohs et al. have argued that the term "base readout" is more informative. Base readout refers to the formation of specific hydrogen bonds and van der Waals contacts between amino acids and nucleotides. These bonds are often directly between the protein and base; however, they may also be mediated by highly restricted water molecules in the major groove. Additionally, hydrophobic contacts may be used by the protein to differentiate between thymine and cytosine.

In the crystal structures of restriction enzymes, it can be seen that the majority of possible hydrogen bonds are formed between the protein and the edges of the bases in the major groove, and often a majority of the possible hydrogen bonds between the protein and the minor groove are fulfilled as well. [3] In the case of the EcoRI specific complex, ([47], [52], [53]) there are 14 direct protein-base contacts and 4 water-mediated hydrogen bond contacts formed by the interface. (Figure 1.4) Additionally, Van der Waals contacts between the protein and thymine and hydrophobic interactions with cytosine contribute to the base readout of this complex. As can be seen in Figure 1.4, most of the amino acids which form hydrogen bonds to the DNA bases also form "buttressing" [2] hydrogen bonds with each other. This serves to highlight the highly

cooperative nature of this interface. Additionally, DNA base analogs have been used in structural perturbation studies to demonstrate that the free energy contributions of interactions to the inner AATT base pairs of the GAATTC site are strictly additive. [14] This suggests that if one or two interactions is deleted, the other contacts will remain in place and the complex is not disrupted.

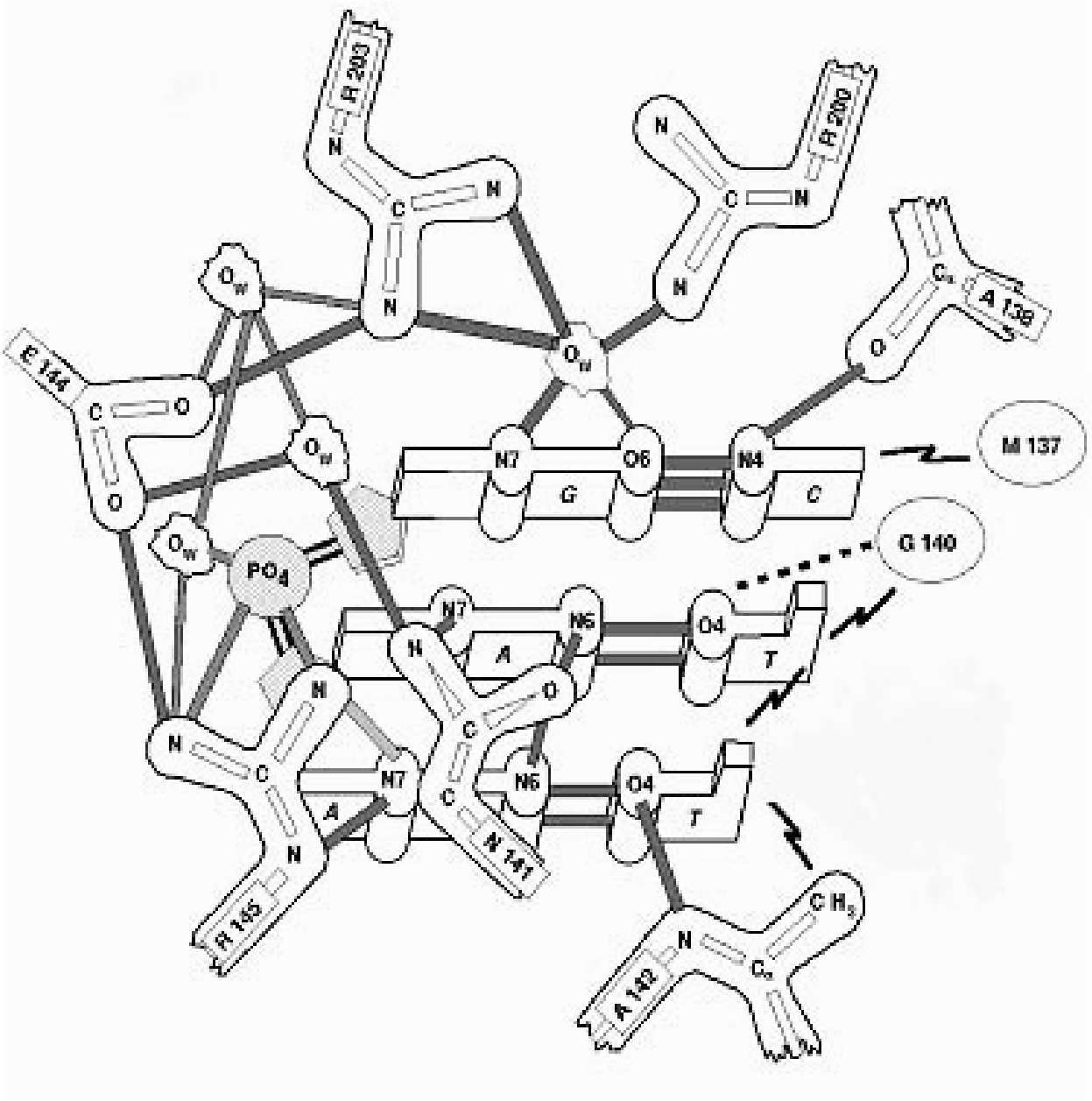


Figure 1.4 Direct readout and buttressing interactions at the EcoRI-specific DNA interface.

Hydrogen bonds are represented by rods, van der Waals interactions by "lightening," and the main-chain Cα-H-O bond by a dotted line. O_W are water molecules. Figure provided by Dr. John M. Rosenberg

1.5.3 Shape or "indirect" readout

Base readout of DNA sequences alone is not sufficient to explain the high degree of specificity exhibited by many protein-DNA interactions. Another major contribution to specificity involves sequence dependent DNA structure as well as the capacity for different sequences to be deformed from the ideal B-form as a consequence of the binding interaction. In addition to the direct base contacts described in the previous section, protein contacts to the phosphate backbone are a feature of protein-DNA binding interactions. These phosphate contacts are proposed to act as "clamps" to help in the positioning of recognition elements and to stabilize the distortion of the DNA as described in the next section. [8], [17], [55], [56] Historically, the term "indirect readout" has been used to describe the contribution of sequence-dependent DNA structure and deformability to protein-DNA interactions. [57], [58]

Crystallographic and solution studies have demonstrated that DNA sequence dictates both the intrinsic deviations from B-form and the propensity for distortion. [59], [60] Many DNA binding proteins are known to distort DNA significantly upon binding, including the classic examples of catabolic activator protein (CAP) [61], [62] and the TATA box binding protein [63]. In an extensive review of extant protein-DNA structures, Rohs, et al [54] describe forms of readout which depend on deviations from the ideal B-form DNA conformation as "shape readout." This category is then subdivided into "local shape readout" (small regions of minor groove narrowing, DNA kinks caused by unstacking of a single base pair, intercalations) and "global shape readout" (conformations where the entire binding site is distorted from the ideal B-form and B-form to A-form DNA transitions). The shape readout is important for sequence discrimination, because the conformational properties of a DNA sequence will dictate the geometric arrangement of available base and phosphate contacts. The energetic cost of this DNA

distortion has been shown to be an important factor in site discrimination. [7], [56] Overall, the base (direct) and shape (indirect) terms are used to describe aspects of the interaction between the protein and the DNA, but in reality these components of recognition are intimately interrelated.

This interrelatedness of readout mechanisms can be clearly seen in the case of EcoRI. Upon formation of the specific complex, significant distortion of the GAATTC sequence can be clearly seen. (Figure 1.5). This kinked conformation results in a wider major groove, which permits accommodation of the amino acids that form the "base readout" interactions. Additionally, it has been hypothesized that the energetic cost for EcoRI-DNA forming this kinked conformation is lowest for the GAATTC sequence. [8], [58]

The distortion of the DNA upon binding is also critical for catalysis. As can been seen in Figure 1.5, EcoRI binding moves the scissile phosphates into novel positions. These positions permit the scissile phosphate to be correctly positioned for catalysis. [47] Further investigation in the Jen-Jacobson lab has shown that introducing a DNA perturbation that causes a steric conflict in the DNA kink causes a large reduction of cleavage rate. [20]

1.5.4 Protein-phosphate contacts in the EcoRI specific complex

In many protein-DNA complexes, the formation of this complementary interface also includes precise positioning of crucial protein-phosphate interactions. Ethylation interference footprinting is an approach used to determine which phosphates of a given DNA sequence are important for binding. [64] Jen-Jacobsen and coworkers found that in the EcoRI-DNA specific complex, there are 14 contacts between the endonuclease and the phosphates (arranged as 7 per DNA strand, shown as green arrows in Figure 1.5). [8] Six of these contacts (three to each DNA strand at

NpNpGAApTTC) form what are referred to as crucial phosphate "clamps." An isosteric replacement wherein G was substituted with 7-deaza-guanosine (^{7C}GAATTC) had a pattern indistinguishable from that of the specific site (Figure 1.6 A and D), supporting the prediction that removing individual hydrogen bonds causes only minor perturbations to the structure. [8]

Miscognate sites, however, displayed dramatic differences from the specific pattern. (Figure 1.6 B and C). The ethylation interference pattern for miscognate complexes shows that two of the "supplementary" phosphate clamps seen in the specific complex are missing in the altered half-site. These complexes also display increased interference in their cognate half-site, with increased interference at the scissile phosphate, as well as a novel contact formed upstream of the binding site. The miscognate complexes are therefore forming an asymmetric "adaptive" interface, providing a structural correlation to the asymmetric cleavage rates that are seen for the two half-sites in these complexes. [8] Nonspecific complexes do not exhibit any ethylation interference pattern, suggesting that the protein is delocalized and can interact with all phosphates equally.

These data taken together indicate that the interactions between the enzyme and phosphate "clamps" play a critical role in specificity. The clamps serve to help stabilize the distortion of the DNA, which in turn enables the enzyme to form base readout contacts within the major groove. Furthermore, the failure to form all of the clamp contacts in the miscognate complexes indicates that the distorted DNA conformation is not stabilized as well. This contributes to the energetic penalty of forming miscognate complexes, because $\Delta\Delta G_{Reorg}$ is penalized in these complexes. [8] The asymmetric or adaptive nature of the ethylation interference pattern in the miscognate sites indicates that the structure of the miscognate complexes may be somewhat asymmetric. Since ethylation interference measures the average of

phosphate contacts for the entire population of complexes in the sample, the miscognate complex has an ensemble which is shifted towards more contacts with the cognate half site, and fewer contacts with the noncognate half site.

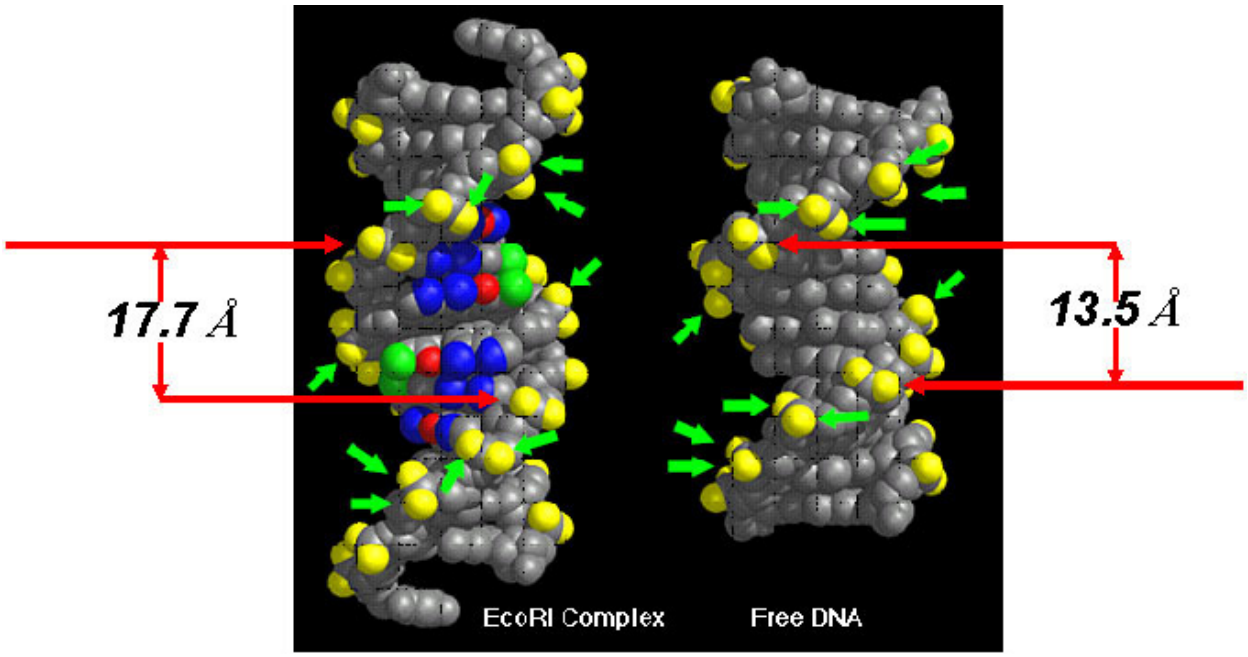


Figure 1.5 Coupling between direct and indirect readout of DNA sequence by EcoRI endonuclease.

The structures of free DNA (PDB ID 335D) and bound DNA from the EcoRI-specific complex. [47] **Scissile phosphates are highlighted with red arrows while the phosphates which contact the enzyme are highlighted with green arrows. The bound DNA exhibits a widened major grove compared to the free DNA. The contacted phosphates are repositioned in the bound state, demonstrating the link between direct and indirect readout. Figure by Dr. Linda Jen-Jacobson, analysis by Dr. Linda Jen-Jacobson and Dr. Arabela Grigorescu.**

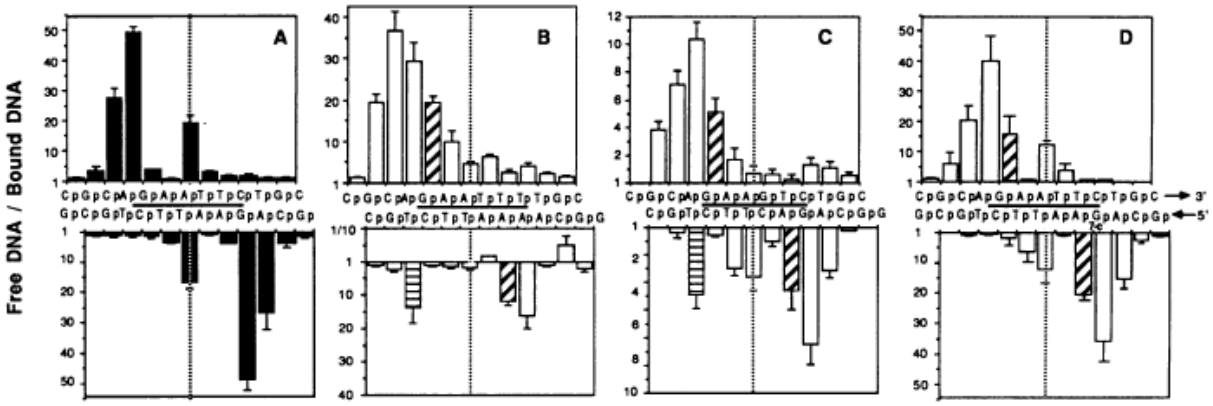


Figure 1.6 Phosphate "clamp" contacts are important for sequence discrimination.

Ethylation interference footprints of EcoRI bound to the specific site (A) GAATTC, the miscognate sites (B) AAATTC, (C) GACTTC, and the specific site with 7-deaza-guanosine isosteric site (D) ^{7C}GAATTC. The binding sites are underlined. The dotted line indicates the rotational axis of symmetry. For figures C-D, the modified half-site is shown in the bottom panel. The horizontally striped bar (C) indicates novel interference on the variant strand. The diagonally striped bars indicated increased interference at the scissile phosphates. Figure taken directly from [8]**.**

1.6 ECORI UNDERGOES A MAJOR CONFORMATIONAL CHANGE UPON BINDING THE SPECIFIC SITE

Site- specific DNA binding proteins often undergo conformational changes upon binding their recognition sequence. These rearrangements include folding and unfolding transitions, as well as large movements of subunits relative to each other. [65] DNA binding enzymes and particularly restriction enzymes often contain “arm” regions which envelop DNA in specific complexes. [3] In many cases, including for EcoRI, the specific complex thus formed is such that a large conformational change *must* take place for binding to occur, as it would be sterically impossible for the DNA to enter the binding site otherwise. [52]

Arabela Grigorescu determined the structure of the EcoRI “free” or apo enzyme. [47] Analysis of the apo structure was somewhat complicated by the fact that the unit cell of the

crystal contains three complete protein chains (monomeric units). Two of these subunits (A and B) form one dimer, while a third dimer, C, dimerizes with the symmetrical C subunit. Interestingly, the three chains exhibit very similar structure with the exception of the arm regions. For two of the three chains, (A and C) the arm regions could not be modeled; electron density was present but not interpretable (Figure 1.7). For the third chain, (B) crystal packing interactions stabilized the arm regions such that only residues 129-132 remained invisible. [47]

These data imply that the EcoRI arms are "disordered" in the free enzyme. In this work, the term disorder is used in the traditional crystallographic sense, to describe residues for which coordinates cannot be determined. There are two main classes of protein disorder. [66] Type I disorder refers to regions which are metastable under physiological conditions; in other words they contain a weak hydrophobic core and undergo frequent order-disorder transitions. Type II disorder refers to regions which are intrinsically disordered. Arabela Grigorescu [47] surmised that the arm regions of EcoRI exhibit type I disorder, based on the observations that A) one of the chains in the apo crystal does exhibit stable arm conformations in the absence of DNA, B) the arm regions have their own hydrophobic core separate from that of the core domain, and C) several hydrophobic residues in the arms are poorly packed, which is a common feature of natively metastable proteins. [67]

In addition to the disorder-order transition discovered by examining the specific and apo structures, Grigorescu also observed that several of the side chains involved in base readout of the recognition site are disordered in the apo structure. [47] This observation emphasizes the importance of the changes in conformation and flexibility of the molecules during specific site recognition.

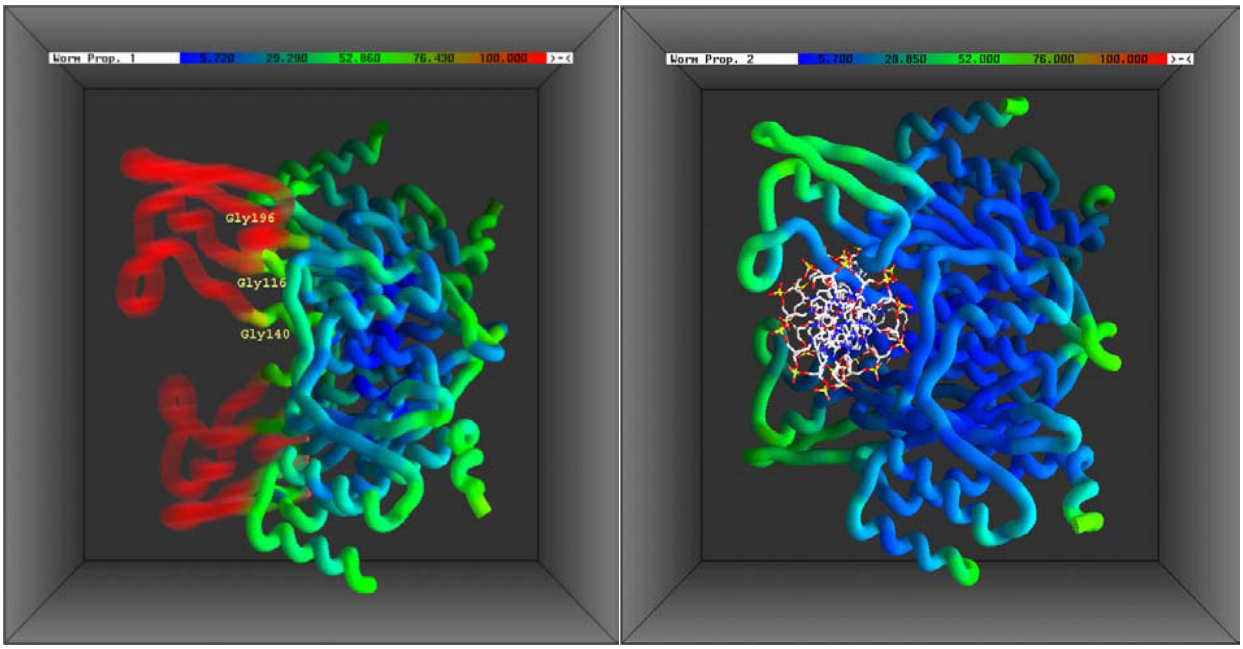


Figure 1.7 The arms of EcoRI are "disordered" in the absence of DNA.

Qualitative model depicting the distribution of structural stability in the free EcoRI dimer (left) and the specific EcoRI-DNA complex (right). The protein backbone is color-coded according to crystallographic B-factors. Dark blue regions are highly stable (B = 20Å^2) green regions are intermediate (B = 50Å^2). Red regions (B=100Å^2) shows the arm regions which are highly disordered in the absence of DNA, and thus blurred in the left panel. The B-factors were re-scaled such that the atoms that form the hydrophobic core of the main domain have identical B-factors. Structures, comparative analysis and figure by Dr. Arabela Grigorescu [47]

1.7 PROMISCUOUS ECORI MUTANTS

Thermodynamic and structural investigations into "promiscuous" EcoRI mutants described in this section highlight the importance of developing experimental methods to investigate the noncognate complex structures, as well as the dynamic properties of EcoRI complexes. These mutants are predicted to form miscognate complexes that are more "specific-like" than those of the wild-type enzyme, and they have been shown to exhibit more favorable binding to the specific site than the wild-type enzyme does. This improved binding cannot be explained by enthalpic contributions alone and is predicted to result from changes in the dynamic properties of these mutant proteins that alter the net binding entropy.

Heitman and Model conducted a screen for relaxed-specificity mutants of EcoRI endonuclease. [68] These mutant EcoRI endonucleases are referred to in the literature as "promiscuous," because they cleave *E. coli* DNA (inducing the SOS DNA-damage response) despite the presence of the EcoRI methylase, leading to effects that range from sublethal to lethal in the cell. Semi-quantitative tests determined that these enzymes are capable of cleaving miscognate sequences with much greater efficiency than the wild-type endonuclease does. [68] The amino acid changes resulting in promiscuous mutants identified by Heitman and Model were A138T, A138V, G192L, H114Y, and Y193H. Interestingly, each of these mutants are located in the arms of the complex (Figure 1.8).

As part of her HHMI summer research project, lab member Bridget Bertoni found that the I197A mutant also displays slight promiscuous behavior. The residue I197 is also positioned in the arms of the complex, and it is in close proximity to A138. The main-chain carbonyl group of this residue makes a water-mediated contact with the amino group of the cytosine flanking the

specific site. This location was chosen as one of the positions investigated by ESR experiments, as detailed in the next chapter.

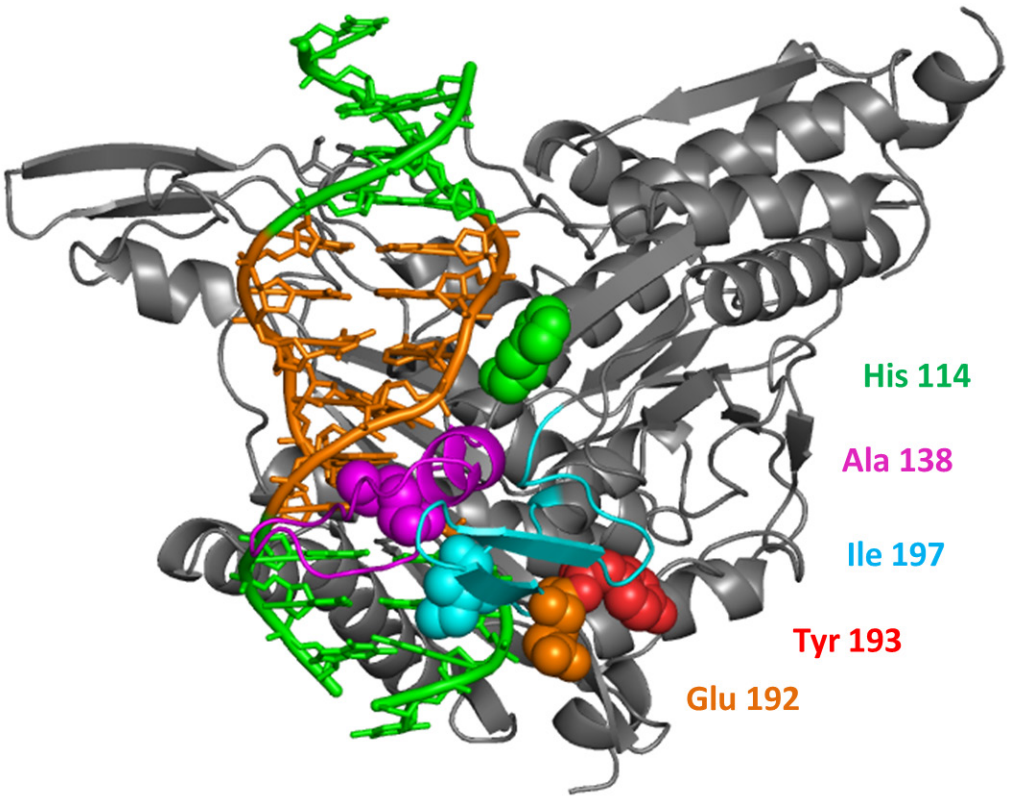


Figure 1.8 Location of the EcoRI "promiscuous" mutations.

Structure shown is the EcoRI-specific complex. The GAATTC site is shown in orange, the flanking sequences in green. Residues which have been implicated in relaxed specificity mutations are indicated with van der Waals spheres in one of the two monomers. Each of these residues is located in or near the "arm" regions. Figure based on [69]

Former lab member Paul Sapienza determined the thermodynamic and kinetic parameters for binding and cleavage for the A138T, H114Y, and E192K promiscuous EcoRI endonucleases. [69] Surprisingly, these enzymes showed a higher binding affinity than wild type enzyme for the specific site. The ΔG^0_{bind} for these mutants to the specific site is ~1.5-3 kcal/mol *more favorable* than for the wild-type protein. (Figure 1.9) Sapienza also found that these mutant enzymes exhibit enhanced discrimination between miscognate and specific sites. In other words, the energetic penalty for wild-type EcoRI binding to the miscognate sites relative to the specific site is ~6.2 to 6.7 kcal/mol, while the promiscuous mutants incur a binding penalty of ~6.7-8.4 kcal/mol (Figure 1.10). [69] For the wild-type enzyme, an important aspect of discrimination is the fact that the binding constants for the miscognate and nonspecific sites are similar, so the nonspecific DNA effectively competes against the miscognate sites in the cell. However, A138T and H114Y show unaltered binding to the nonspecific site but enhanced binding to miscognate sites, such that miscognate DNA is preferred over nonspecific DNA and nonspecific DNA is no longer an effective competitor. (Figure 1.10) The E192K mutant shows enhanced binding for all three classes of sites.

In addition, Sapienza found that the first-order cleavage rate constants of the mutant enzymes are normal for the specific site but greater than those of wild type enzyme at EcoRI miscognate sites, with the consequence that *in vivo*, cleavage of these sites is more likely than for wild type enzyme. [69] These observations taken together provided insight the characteristics of the promiscuous phenotype.

Sapienza used the temperature dependence of the equilibrium constant to dissect the enthalpic and entropic contributions to the altered $\Delta G°_{bind}$ of the A138T mutant for the specific site. [70] He found that the A138T – specific complex has $\Delta H°$ that is ~2.6 kcal/mol less

favorable than the specific complex, however the $T\Delta S^0$ is approximately 4.1 kcal/mol more favorable, resulting in the more favorable net $\Delta G°_{bind}$ for the A138T-specific complex. Therefore, the improved $\Delta G°_{bind}$ for the A138T-specific complex results from a more favorable $T\Delta S°$.

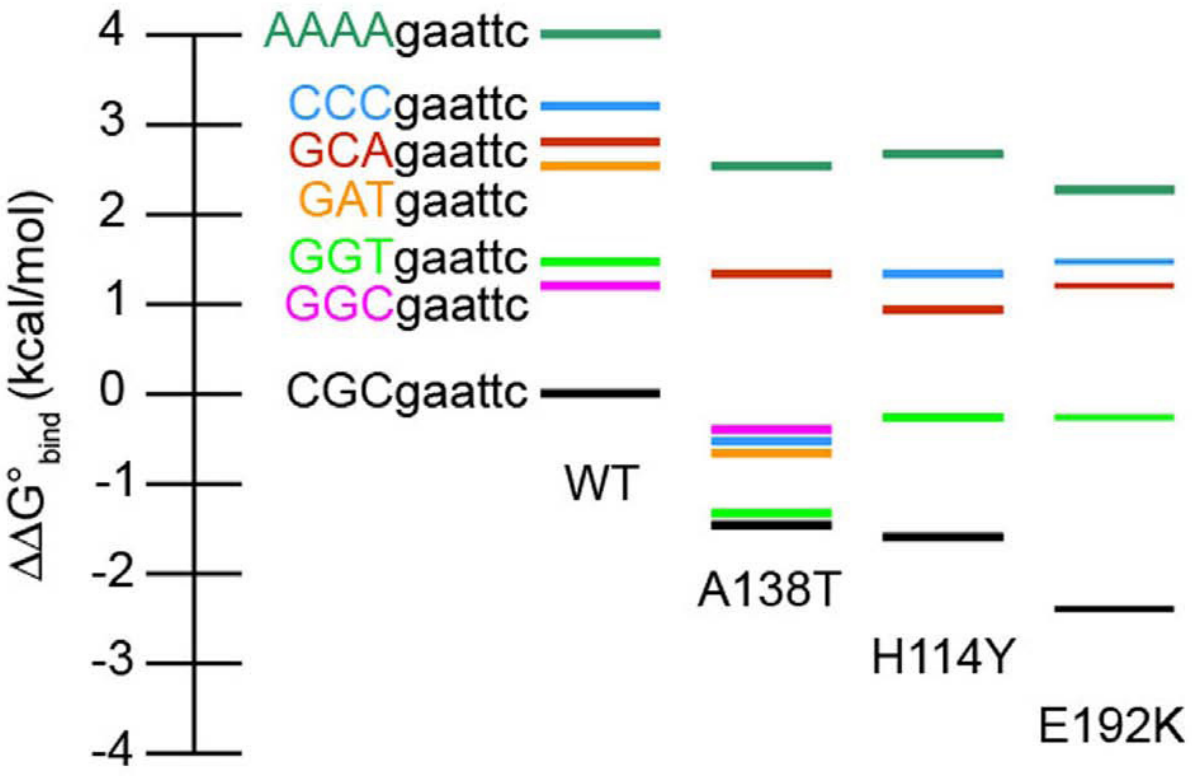


Figure 1.9 Binding of EcoRI WT enzyme and promiscuous mutants to the specific DNA site.

Wild type and promiscuous mutant binding is shown for different flanking contexts (shown as color coded bars). The scale of $\Delta\Delta G°_{bind}$ is shown on the left, normalized to wild-type binding to the favored sequence as $\Delta\Delta G°_{bind}= 0$. Figure taken directly from [69]

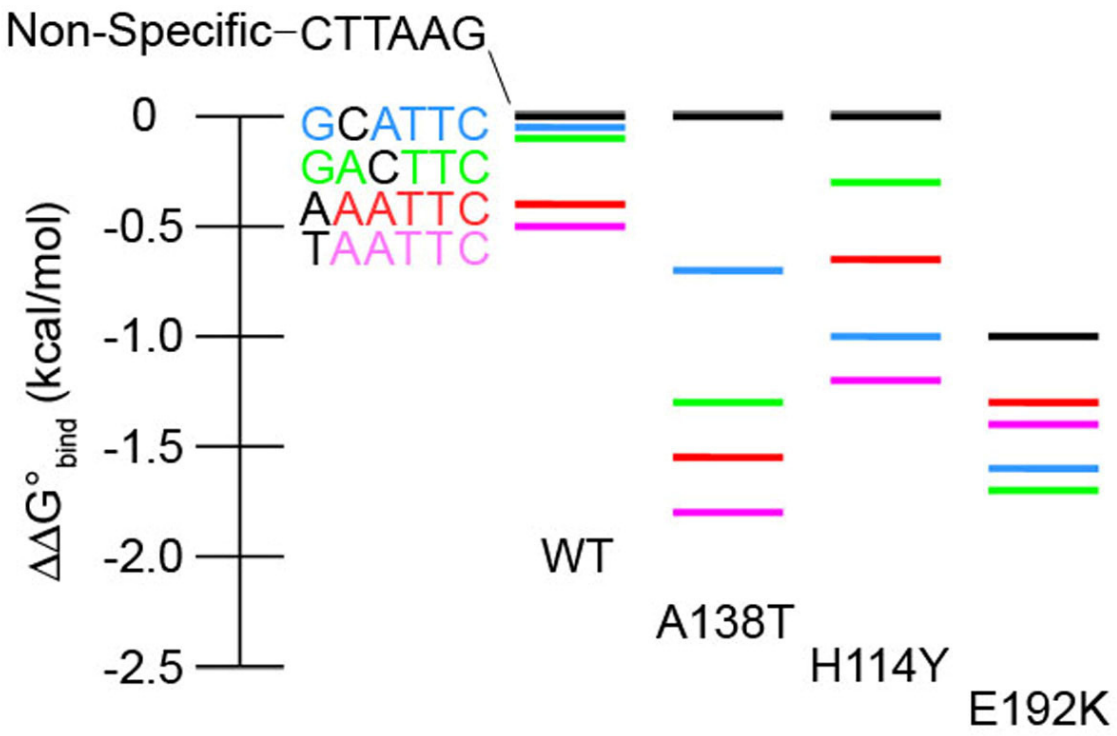


Figure 1.10 Binding of wild type and promiscuous EcoRI to noncognate sequences.

Relative binding free energies for wild-type and promiscuous mutant EcoRI binding to noncognate sequences. The scale of $\Delta\Delta G^{\circ}_{bind}$ is shown on the left, normalized to wild-type binding to nonspecific DNA as $\Delta\Delta G^{\circ}_{bind} = 0$. Figure taken directly from [69].

A large body of evidence suggests that complexes between the promiscuous mutant enzyme and miscognate sites resemble specific complexes much more than wild type-miscognate complexes do. Specifically, Sapienza showed that crucial phosphate "clamp" contacts which are lost upon wild type binding to EcoRI miscognate sites are restored in promiscuous miscognate complexes (Figure 1.11). [6] The working hypothesis is that these phosphate clamps play a key role in stabilizing the distorted DNA conformation required for assembly of the active site in both the pre-transition state (recognition) and transition state complexes and that restoration of these crucial phosphate contacts leads to a higher probability for promiscuous enzyme complexes to reach the transition state for EcoRI miscognate site cleavage. Since the EcoRI "arms" provide some of these key "clamp" contacts, an extension of this model is that the promiscuous mutations permit regions of the arm to adapt or adjust in the miscognate complexes such that these crucial phosphate contacts can be formed.

Additional lines of evidence pointing to the “specific-like” nature of the promiscuous miscognate complex include the fact that these complexes show a higher cleavage rate than the wild type miscognate complexes, indicating that the binding energy of miscognate sites is used more efficiently on the path to the transition state in promiscuous complexes than in wild-type complexes. [69] Additionally, while the wild-type miscognate cleavage rates are strongly dependent on the flanking context in which they are embedded, those of the A138T complexes are much less dependent on the flanking context. This is more similar to the characteristics of the wild-type complex with the specific binding site. [6]

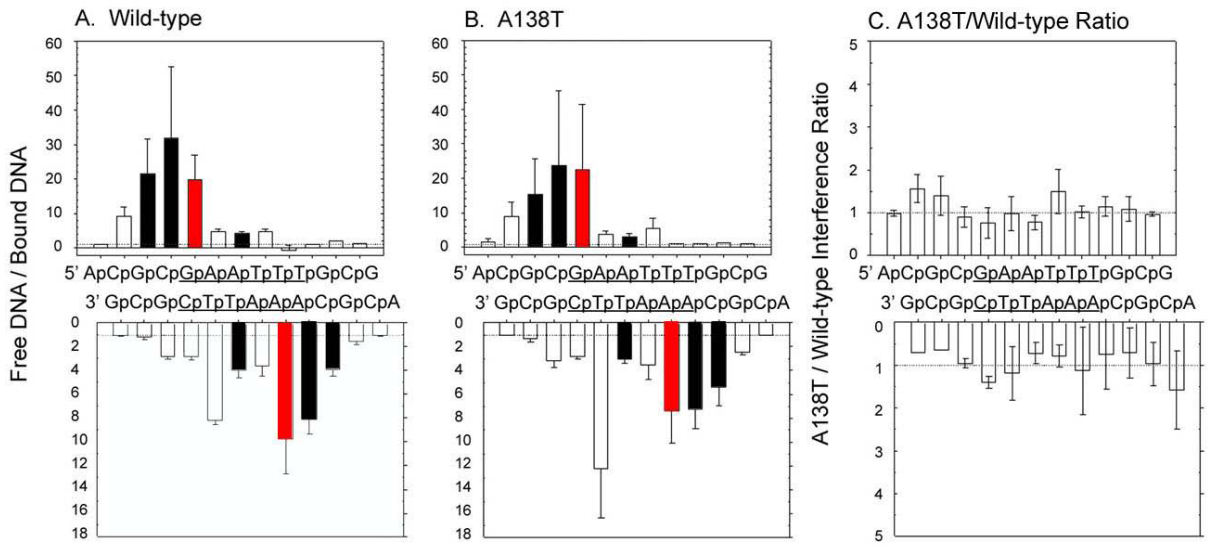


Figure 1.11 Ethylation interference patterns of wild type and A138T EcoRI bound to cgcAAATTCgcg.

Panel A shows the wild-type complex bound to the AAATTC miscognate site, Panel B shows the A138T mutant bound to the miscognate site. Panel C shows the ratio of the A138T interference values divided by the wild-type results. Interference bars for the scissile phosphates are colored red, and the bars for the "clamps" are colored black. Figure taken directly from [6]**.**

Sapienza also solved the crystal structure of the A138T mutant in complex with the specific DNA binding site and performed MD simulations to compare the wild-type and A138T specific complexes. [70] Interestingly, the crystal structure of the promiscuous A138T-DNA cognate complex showed only subtle changes between the specific and mutant complexes. These differences include the loss of a constrained water from the active site and novel packing interactions between the T138 and the DNA (Figure 1.12). These observations are in agreement with the thermodynamic changes observed between the complexes.

Sapienza predicted that another source of favorable net entropy may result if the arms of the free A138T enzyme are more constrained than those of the wild-type enzyme. If this were the case then the entropic penalty incurred by the folding and conformational constraint imposed on the arms by binding would be less for this mutant complex. His results revealed that the

differences in specific binding affinity between the wild-type and promiscuous mutants cannot be explained simply by "enthalpy-centric" differences in the contacts that are made to the DNA. This serves to highlight the importance of investigating entropic contributions to thermodynamic processes, including changes in conformational freedom upon enzyme binding. [70]

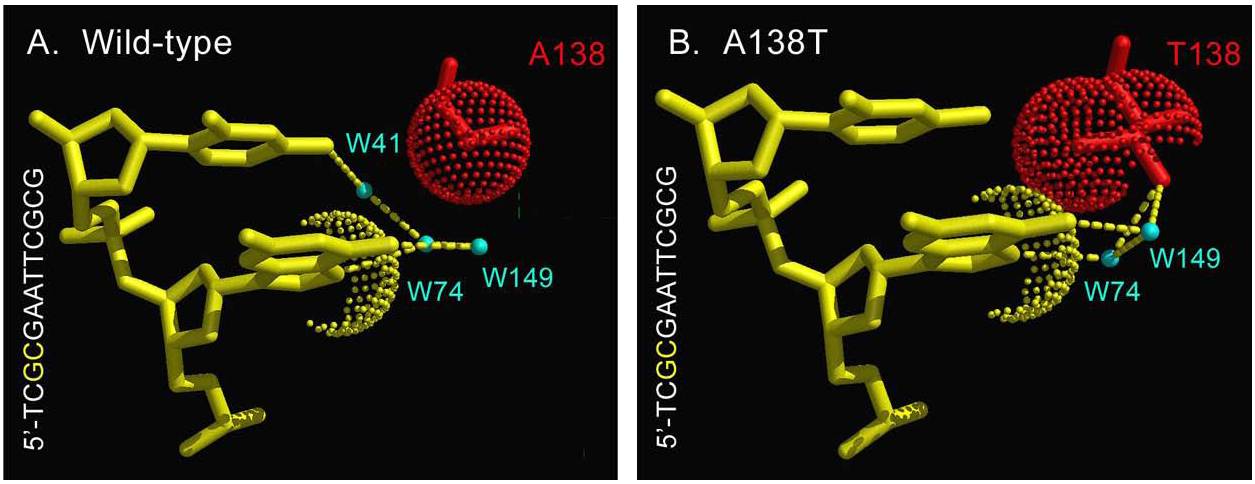


Figure 1.12 Active site differences in the A138T structure.

Panel A shows the van der Waals surfaces of the N7 guanosine (yellow dots) and the Cβ of alanine 138 (red dots) in the wild-type specific crystal structure. In the A138T complex, threonine T138 packs against the N7 guanosine, indicating a novel dipole-induced dipole interaction between these groups. Figure taken from [6]

Taken together, the prediction from the above results is that the structural and dynamical properties of the "arms" in the promiscuous complex will more closely resemble those of the arms of the specific complex. Sapienza proposed in his PhD dissertation that the characteristics exhibited by the "promiscuous" mutants are likely to be at least partly the result of perturbations of the structure and/or dynamics of the free enzyme, the miscognate complexes, and/or the dynamic behaviors of the specific complex. [6] The changes in exposed surface area and configurational entropy in the arm regions which are predicted to accompany binding could be affected by mutations in the arm regions. In his investigations of the properties of EcoRI promiscuous mutants, Sapienza made several observations which are consistent with this

prediction. First, Sapienza observed that the H114Y mutation slowed the rate-limiting step of specific DNA association, which is likely the result of a perturbation of the folding transition of the arms upon binding. Second, he observed that the crystallographic B-factors for some regions of the A138T-specific complex were lower than those for the wild type specific complex, which is consistent with the idea that the A138T mutation affects the dynamics of the arms. Third, as described in a previous section, he found that the thermodynamic behaviors and phosphate contacts for the promiscuous mutants in the miscognate complex are more "specific-like" than for the wild type complex, suggesting conformational adaptations that allow the promiscuous miscognate complexes to form a binding interface which is closer to the transition state, permitting cleavage. [6]

As suggested by Sapienza, these results highlight the importance of developing experimental methods to investigate the noncognate complex structures, as well as the dynamic properties of these complexes, in order to obtain an understanding of what types of changes in conformations and dynamics occur between the different binding modes. Such methods would permit the investigation of "invisible" entropic factors such as changes in configurational entropy between the binding modes and upon making perturbations such as promiscuous mutations. [6]

1.8 SUMMARY OF ECORI BINDING MODES

The differences observed for the wild-type EcoRI-DNA complexes are summarized in the following table.

Table 1.3 Discrimination by EcoRI: Three distinct binding modes for EcoRI-DNA complexes. [a]

	Specific Complex	**Miscognate Complex**	**Nonspecific Complex**
DNA Site	GAATTC	One incorrect base pair	2 or more incorrect base pairs
Specificity Ratio [b]	1	1.3×10^5	2.5×10^6
Phosphate contacts [c] (Footprinting)	Strong	Moderate and asymmetric	None
Salt dependence [d]	Steep	Intermediate	Very Steep
Heat capacity change [e]	Large and negative -2.6kcal/mol K	Intermediate -0.1 to 1.0 kcal/mol K	Approximately Zero
DNA Cleavage [f]	Symmetric rates	Reduced asymmetric rates	None

Table adapted from [6]**.**
[a] General trends pertain to EcoRI, EcoRV, and BamHI endonucleases. Specific numbers are for EcoRI.
[b] Specificity ratio= $K_{specific}/K_{non\text{-}cognate.}$ [7]
[c] [8]**.**
[d] (David R. Lesser unpublished data.)
[e] [26], [47]
[f] [8]

1.9 STRUCTURAL INFORMATION FOR NONCOGNATE COMPLEXES IS REQUIRED TO COMPLETE OUR UNDERSTANDING OF SPECIFICITY

In order to gain a more complete understanding of the discrimination mechanisms of restriction enzymes, it will be necessary to determine how the structures of these complexes differ at specific, miscognate, and nonspecific DNA sites. As discussed in the previous sections, extensive genetic, biochemical, and biophysical data suggest that the three classes of complexes are structurally distinct. As has been suggested by others, examination of the miscognate and nonspecific complexes of EcoRI enzyme would be particularly valuable, given the wealth of thermodynamic and structural information available for this enzyme. [6], [47], [71]

Miscognate and non-specific complexes, however, are notoriously intractable to crystallographic analysis. For the ~4000 known restriction enzymes, over a hundred crystal structures have been solved representing 40 distinct enzymes- yet only four of these structures are of restriction enzymes in complex with noncognate sites. [11] Two of these structures are of EcoRV bound to different nonspecific sites, [16], [44] one of these is of a BamHI miscognate site [72] and one is of BstYI bound to a noncognate site [71]. In addition to these, there are a handful of noncognate complexes for other site-specific DNA binding proteins. [45], [73-75]

1.9.1 Structural features of noncognate complexes

The noncognate complexes examined to date have several features that distinguish them from specific complexes. The extensive networks of hydrogen bond contacts observed in specific complexes discussed above are not present in nonspecific complexes. [44], [45], [73], [74] Additionally, the precisely constrained phosphate clamp contacts observed in the specific

complex are absent. [8], [17], [55] Instead, these interactions are primarily mediated by nonspecific Coulombic interactions which are often water-mediated. Furthermore, these "loose" complexes exhibit reduced steric complementarity and reduced burial of solvent accessible surface area. [44], [45], [74], [76] Interestingly, in the miscognate BamHI structure, base-readout contacts were lost throughout, while in the BstYI bound to a site differing by one base pair, a "hemispecific" complex was formed with the cognate half site maintaining all of the specific readout contacts, and the noncognate site missing these contacts. [72] These few structures represent a tantalizing but incomplete picture of noncognate protein-DNA complexes. A major aim of this work is to develop and establish a method for examining noncognate complexes which are currently intractable to crystallography.

1.10 THE IMPORTANCE OF DYNAMICS

"Studying a protein while ignoring its movements is rather like trying to understand a feature-length movie by gazing at a single frame of the film. How to start the projector-this is the challenge that is attracting scientists from all areas of structural biology" [77]

Crystal structures, while critical to our understanding of protein functions, have led to a static and "enthalpy centric" view of these processes. The term "dynamics" in this work refers to the internal motions of proteins in solution. In solution, proteins sample a range of conformations which play a significant role in the function of the protein. The observed solution conformational fluctuations can include major structural changes, movement of flaps or domains, local folding or unfolding transitions, helix-tilting motions, and pH-gated structural changes. [78] The functional importance of conformational switching has been observed by solution Nuclear

Magnetic Resonance (NMR) studies of proteins [74], [79-82] as well as by a number of Electron Spin Resonance (ESR) studies. [83-86]

Another important aspect of protein dynamics is that of local backbone fluctuations. Local backbone fluctuations include both smaller scale oscillations of backbone dihedral angles and larger scale rigid-body motions of secondary structural elements. [78] Both conformational switching and local backbone fluctuations are comprised of a number of interrelated motional contributions to the overall dynamics of a protein, as summarized in Table 1.5.

In general, changes in the motions of the protein and DNA are predicted to make a major contribution to the large negative ΔCp observed for specific protein-DNA binding as described earlier in this chapter. In particular, decreases in the configurational entropy, [87] especially confinement of molecular movement to below 0.5Å [88] are predicted to result in the large negative ΔCp observed for the formation of these complexes. [26]

Table 1.4 Summary of internal motions of proteins.

Motions of proteins	
I.	**Local motions (0.01-5Å, 10^{-15}-10^{-1}s)**
	Protein atomic fluctuations: small displacements for substrate binding, adaptations to accommodate other "rigid body" motions, thermal motion of bonded atoms, elastic vibration of globular regions
	Sidechain fluctuations: rotation of sidechains at surface (fast) or within interior (slow)
	Loop motions: small thermal motions, disorder-to-order transitions coupled to binding, folding, or activation, rearrangement of 'hinge' regions to facilitate rigid –body motions
II.	**Rigid-body motions (1-10 Å, 10^{-9}-1s)**
	Rigid motion of individual helices
	Relative motions between domains (hinge bending)
	Allosteric transitions between subunits
III.	**Larger-scale motions (>5Å, 10^{-7}-10^{4}s)**
	Folding and unfolding transitions
	Association/dissociation coupled structural changes

The descriptions above are compiled from the following references: [43], [78], [89]**.**

Understanding the dynamic motions of a protein and changes in these motions is critical to gaining a full understanding of how structure confers function. For example, in the case of the EcoRI A138T promiscuous mutant, a predicted contribution to the more favorable ΔS is that the mutation is hypothesized to reduce the conformational freedom of the unbound state. [70] NMR studies on proteins such as calmodulin are finding that measurements of the internal dynamics of a protein can be used to estimate conformational entropy of a protein and that these estimates of conformational entropy can be directly correlated to binding entropy in ways that assist the thermodynamic dissection of the "invisible" entropic contributions to binding energy. [90] Other NMR studies are finding that dynamic measurements can be extremely informative for structure-based drug design; for example, these measurements can detect allosteric changes upon ligand binding that do not noticeably change the structure but alter the dynamic behaviors of the protein in dramatic ways. [91]

Much of the progress in protein dynamics research comes from NMR studies, but since not all proteins are suitable for NMR, ESR is also being developed as a tool for investigating protein dynamics. [78] This will be discussed in more detail in Chapter 3. One major aim of this work is to apply emerging ESR techniques for dynamics measurements to examine the changes in conformational freedom that exist between the specific, miscognate, and nonspecific EcoRI-DNA complexes.

1.11 ELECTRON SPIN RESONANCE

In order to gain insight into the molecular mechanisms contributing to EcoRI site specificity, a major goal of this dissertation is to obtain information about the positions and dynamics of the arm regions in EcoRI complexes with specific, miscognate, and nonspecific DNA.

As discussed in the previous sections, while the EcoRI free enzyme and specific complex have been rigorously investigated by crystallography, [47] structures for the miscognate and nonspecific complex remain elusive. Furthermore, crystallography does not provide direct evidence of the dynamic behaviors of the enzyme in solution.

Other options for investigation of this information could include NMR or FRET (Förster Resonance Energy Transfer). However, as of this writing, the concentrations and conditions required for NMR studies are not suitable for the EcoRI system. Additionally, FRET has several disadvantages. FRET requires the attachment of fairly large probes, which complicates data interpretation, particularly when attempting to investigate subtle differences between structures such as those investigated in this work. FRET also requires labeling with two different probes, greatly increasing the difficulty of protocol development. We have thus chosen to use ESR spectroscopy to examine the positions and dynamics of the arms in the various complexes.

1.11.1 Introduction to ESR theory

As all of the chapters in this work include data obtained from ESR spectroscopy, I include here a very brief introduction to ESR concepts and terminology. Some of these concepts are expanded upon in relevant chapters. As these concepts are discussed in a number of textbooks and review

articles, the discussion herein is based on the following references unless otherwise noted. [92-94]

An electron has a spin quantum number of ½, with a magnetic component m_s of ± ½. In the presence of an external magnetic field of strength B_0, the magnetic moment of the electron will either align itself parallel (lower energy state) or antiparallel (higher energy state) with the magnetic field. (Figure 1.13). The energy levels for the spin are given by the spin Hamiltonian. The spin Hamiltonian describes the interactions of the electron spin with the magnetic field, adjacent nuclear spins, and other electrons. For an isolated electron in a magnetic field, the spin Hamiltonian is given as:

$$H_{spin} = g\beta_e B_0 S_z$$

Where g is the electron g factor, β_e is the Bohr magneton (a constant), B_0 is the magnetic field strength, and S_z is the electron spin state (±1/2). (Reviewed in [94].) From this equation it can be seen that the difference in energy levels between the two spin states for a given electron is proportional to the magnetic field strength, and that in the absence of a magnetic field, these two states are indistinguishable.

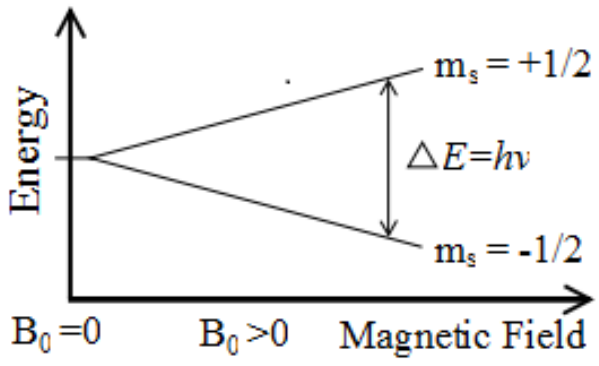


Figure 1.13 Energy diagram for an electron spin at varying magnetic field strength.

Adapted from [94]

An unpaired electron in the presence of a magnetic field can move between these energy levels by absorbing or emitting electromagnetic radiation of energy $\varepsilon = h\nu$ such that $\varepsilon = \Delta E$, where ν is the frequency and h is Plank's constant. By detecting this change of electromagnetic radiation, ESR spectra are obtained. In order to optimize the signal to noise ratio of the absorption spectrum, the magnetic field is usually modulated at a constant frequency and the absorption spectrum is detected and plotted as the first derivative of the spectrum. (Figure 1.16) The continuous wave (CW) spectra discussed in Chapter 4 are of this type.

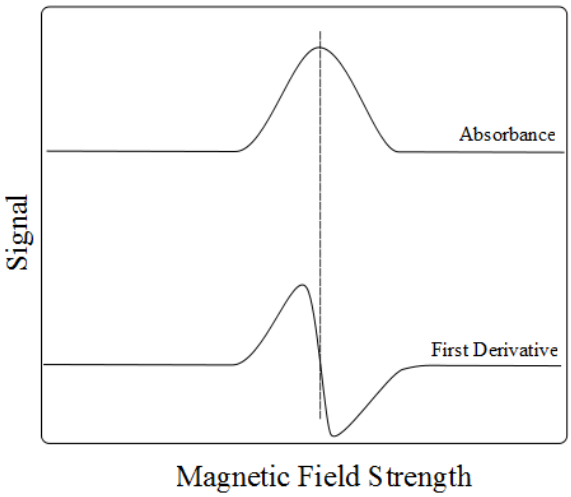


Figure 1.14 Idealized CW lineshape for a population of free electron spins.

Adapted from [95]

The discussion above is for an idealized case of a free electron spin. In reality, coupled nuclear spins, dipolar coupling between spins in a sample, and other factors can influence the shape of the absorption spectrum. Examining the shape of these spectra can provide a wealth of information on the local environment of the spin. A more detailed discussion of this is provided in Chapter 3. Pulsed techniques such as Double Electron-Electron Resonance (DEER, discussed

in Chapter 2) are able to detect dipolar coupling between spins in order to obtain distance distributions between spins.

1.11.2 Applications of ESR in biological molecules

In biological molecules, unpaired electrons are comparatively rare. Exceptions include metal ions that are bound to proteins (including EcoRI, see Chapter 4 for discussion) and transient unpaired electrons that are formed during biochemical reactions such as photosynthesis and redox reactions. However, stable electron spins can be introduced to biological molecules by covalent attachment of an organic free radical in a technique referred to as spin labeling (discussed in more detail in Chapter 3). This is the technique most commonly used for conducting ESR in biological systems. By attaching a spin label to a specific residue or residues on a protein, information about the dynamic behaviors and environment of that residue can be obtained. Distance distributions between pairs of labels can also be examined to provide information about the structure and conformation of the molecule under investigation.

ESR techniques have proven effective for investigating numerous aspects of biological molecules, including distance measurements, dynamics, conformational changes, and surface areas. These applications are discussed in a number of articles and book chapters (including [94-101]). ESR data collection methods include continuous wave (CW) measurements described briefly above and in Chapter 3 which report on the local environment of the spin label, and pulsed measurements including Double Electron-Electron Resonance (DEER, described in Chapter 2), which reports distance distributions between two probes, as well as techniques such as Electron Spin Echo Envelope Modulation (ESEEM, described in Chapter 4), which provide

information on the interactions between the electron being probed and nearby nuclei in the system.

ESR methods offer several advantages for our investigations:

1) By using different data collection methods, ESR enables the determination of positional data, dynamic information, and information on the local environment for specific residues of our protein.

2) ESR permits the collection of information about the protein in solution, under a wide range of temperature and solution conditions.

3) EcoRI is especially suited to ESR studies, having only one (buried) native cysteine residue. (The spin label reagent most commonly in use is sulfhydryl-specific.) In addition, introducing a spin label at a particular position in the homodimeric EcoRI protein has the consequence of introducing two paramagnetic side chains, thus permitting the determination of the inter-subunit distance between side chains at the two sites.

4) EcoRI can coordinate Cu^{2+} ions, enabling investigations of the copper binding site, as well as positional "triangulation" by obtaining distances between Cu^{2+} to Cu^{2+}, Cu^{2+} to spin label, and spin label to spin label positions.

2.0 ELECTRON SPIN RESONANCE SHOWS COMMON STRUCTURAL FEATURES FOR DIFFERENT CLASSES OF ECORI-DNA COMPLEXES

Portions of this chapter have been adapted from previously published work (Stone, Townsend et al., 2008) *with permission from the publisher.*

2.1 INTRODUCTION

As discussed in Chapter 1, EcoRI undergoes a conformational change upon binding to specific DNA. As part of investigating the extremely high specificity exhibited by EcoRI, it is necessary to determine how the structures of EcoRI complexes differ at specific, miscognate and nonspecific DNA sites. Footprinting results [8] suggest that these three classes of complexes are structurally distinct, and thermodynamic profiles ($\Delta G°$, $\Delta H°$, $\Delta S°$, $\Delta C°_p$) [26], [70] suggest that the specific complex has more restricted conformational–vibrational mobility of the protein and the DNA.

As has been suggested by others, an examination of the miscognate and nonspecific complexes of EcoRI enzyme would be particularly valuable, given the wealth of thermodynamic and structural information existing for this enzyme (as reviewed in the Introduction). [6], [47], [71] There are crystal structures of the free protein [47] and the metal-free specific protein–DNA complex. [47], [53] Miscognate and non-specific complexes, however, have not been readily

accessible to crystallographic analysis. As discussed in Chapter 1, crystal structures for noncognate complexes are still comparatively rare.

Herein, I describe results that demonstrate the utility of pulsed ESR distance measurements for shedding light on the structures of miscognate and non-specific complexes. I briefly introduce Double Electron-Electron Resonance-Site-Directed Spin Labeling (DEER-SDSL) methods, and demonstrate our use of these methods to examine the specific, miscognate and nonspecific complexes of EcoRI in solution.

2.2 DISTANCE MEASUREMENTS BY ESR

2.2.1 Site-directed spin labeling (SDSL)

Electron Spin Resonance is the absorption of radiation by a unpaired electron in the presence of a magnetic field. The shape of the resulting absorption spectrum provides information on the local environment of the spin (discussed in Chapters 1 and 3). For biological molecules such as proteins, an unpaired electron spin must be introduced so that ESR experiments can be performed. Two primary ways to achieve this for proteins are: 1) for metal-binding proteins, a bound paramagnetic metal can be used as a probe to investigate the metal binding site or sites (discussed in more detail in Chapter 4) or 2) by covalent attachment of a label that contains an unpaired electron spin.

Work in the lab of Wayne Hubbell introduced the powerful tool of combining site-directed mutagenesis with covalent spin labeling in a technique referred to as Site-Directed Spin Labeling (SDSL). [103], [104] This technique is now well-established for introducing spin

labels into proteins for structural and dynamic ESR measurements. [78], [100], [105] In SDSL, the introduction of a spin label to a protein is achieved by using site-directed mutagenesis to introduce a cysteine mutation at the chosen residue. After the modified protein is purified, it is reacted with a sulfhydryl-specific nitroxide spin label. This is a highly stable spin-label which is capable of forming disulfide bridges with cysteine residues. The label used in the experiments herein is the most widely used spin label, (1-Oxyl-2,2,5,5-tetramethyl-3-Pyrrolin-3-methyl-) methanethiosulfonate. (MTSSL, Toronto Research Chemicals, Figure 2.1. [106]) Nitroxide spin labels are relatively small. For example, MTSSL is approximately the size of a tyrosine, and has been shown to cause minimal perturbation to the native structure of a protein. Detailed studies of T4 lysozyme, for example, show little or no effect on the stability or structure of the protein upon incorporation of the spin label. [86]

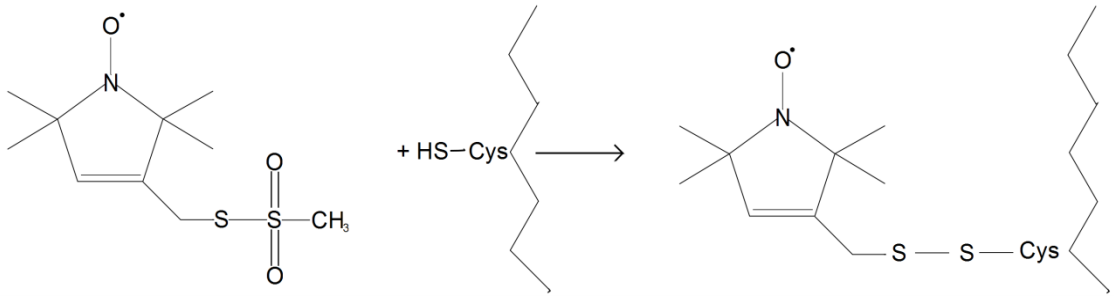


Figure 2.1 MTSSL nitroxide spin label and site-directed spin labeling.

Site-directed spin labeling is a very powerful technique, because it allows the researcher to choose precisely which locations on the protein they wish to investigate. By examining multiple specific sites for a given protein, the researcher can potentially obtain detailed information about dynamics, conformational changes, and secondary structure [78], [100], [107] as well as solvent accessibility profiles [108], [109]. Additionally, by introducing a pair of spin labels to a protein, distance measures may be obtained as discussed in the next section. These distance measures provide constraints for three-dimensional structure determination.

2.2.2 Double Electron Electron Resonance (DEER)

The distance between two unpaired electrons (or spin labels) can be obtained from ESR signals because the dipole-dipole coupling between any two spins is proportional to the inverse cube of the distance r between them. For short distances of ~8-20Å, continuous wave (CW) ESR experiments are sufficient to determine distances from broadening of the line shape. [85] For longer distances, (~20-80 Å) the relatively weak dipolar interaction must be extracted from the other contributions to the signal by using pulsed ESR techniques such as (Double Electron Electron Resonance) DEER. [110]

Briefly, the DEER experiment is a pulsed ESR technique which employs two separate microwave frequencies in order to extract the signal resulting from the dipolar coupling between two electron spins, enabling the determination of distance distributions between the two. In SDSL-DEER experiments such as those in this chapter, the distance being investigated is the distance between two spin labels attached to a single protein molecule. (These concepts are expanded in a number of excellent reviews including [100], [111], [112].)

Given a sample of randomly distributed molecules that are doubly-labeled, (as shown in Figure 2.2,) the resulting DEER signal is a time-domain signal $V_{(t)}$. This is referred to as the dipolar evolution function, which is comprised of the dipolar coupling resulting from both intra- and inter-molecular signals:

$$V_{(t)} = V_{inter} * V_{intra}$$

Given that the sample is composed of homogenously distributed molecules, V_{inter} results in a background signal that can be divided out of the total signal in order to extract the intramolecular signal. This results in the background-corrected time domain signal, V_{intra}, shown in the top panel of Figures 2.7-2.12.

This signal (V_{intra}) modulates with a frequency ν (given in MHz) which is proportional to the dipolar interaction between the spins. The frequency of the dipole-dipole interaction (ν) is described in the following equation, (Reviewed in [111], [113]) where μ_0 is the vacuum permeability constant, g_A and g_B are the g tensors of the spins, $\hbar$ is the reduced Planck (sometimes called Dirac) constant, r is the distance between spins, and θ is the angle between r and the external magnetic field :

$$\upsilon = \frac{\mu_0}{8\pi^2}\frac{g_A g_B \mu_B^2}{r_{AB}^3}\left(3\cos^2\theta_{AB} - 1\right)$$

Under the conditions used for the DEER experiment, the majority of the signal originates from excitation of the $\theta = 90°$ orientation. Assuming that the signal originates from the 90° orientation and using the g value for the unpaired electron in a nitroxide, the relationship of the modulation frequency (ν) to the distance between spins can be simplified as follows [111]:

$$r = (52.04\text{MHz nm}^{-3}/ \nu) \text{ ^ } 1/3$$

The above equations describe the relation of the frequency modulation to a single distance r between a pair of spins. Determination of the distribution of distances for a population of spins $P_{(r)}$ from the signal $V_{(t)}$, is somewhat more complex, as very similar signals can potentially result from very different distance distributions. [114], [115] Because of this, experimental noise can cause large variations in the resulting distance distributions. It is therefore crucial to optimize the balance between maximizing the resolution of the experimental signal and preventing experimental noise from causing variations in the results when obtaining distance distributions. By comparing the experimentally determined signal with simulated signals based on different potential distance distributions, the distance distribution can be optimized to identify the most accurate distance distribution. We employed one such method, called Tikhanov regularization, in our data analysis. (Described in [115].) When DEER is

conducted on a doubly-labeled protein, this permits distance mapping between the spin-labeled residues, as well as some information on conformational dynamics by comparison of distance distributions. (Reviewed in [97].)

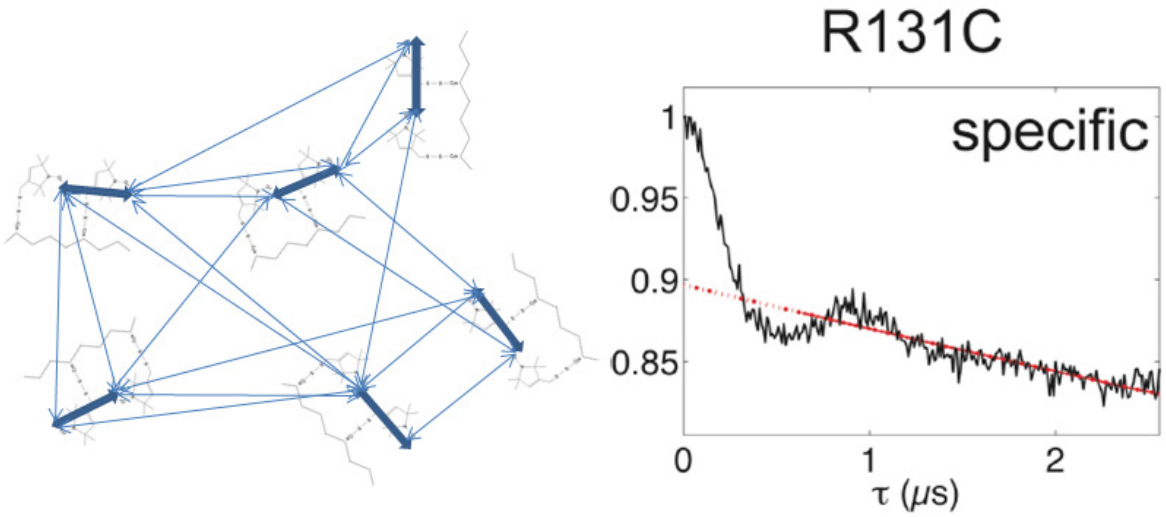


Figure 2.2 Signal contributions in the DEER experiment.

The left panel shows a population of randomly oriented "molecules" which have been labeled with a pair of nitroxides. The thick blue arrows represent the dipolar interaction resulting in the intramolecular (V_{intra}) contribution to the total signal V_t. The thin blue arrows represent a subset of the intermolecular dipolar interactions, which result in the V_{inter} contribution to the total signal. The right panel shows a representative raw time domain signal (V_t) from our DEER experiments. [116] **The red dotted line represents the baseline correction for this signal.**

Major advantages of the DEER experiment, especially in combination with SDSL, include the ability to obtain structural information for proteins that cannot be readily crystallized, such as membrane proteins [97] and noncognate complexes ([116], this work). The DEER experiment permits the determination of the solution structure of proteins in buffers that can approximate physiological conditions. DEER has been well established for measuring distance constraints in membrane proteins, [117-120] soluble proteins, ([121], [122]) oligonucleotides,

[123-125] and synthetic oligomers. ([113], [126], [127]) DEER has also been used to examine conformational changes in proteins including kinesin, ([128], [129]) myosin [129-131], and upon metal binding in anthrasis repressor (a DNA binding protein). [132] Additionally, some information on the dynamic behaviors of the protein can be obtained by examining the distance distributions. [97]

2.2.2.1 Considerations for SDSL-DEER experiments

SDSL-DEER is a very valuable and versatile tool, but there are some caveats that must be considered. The nitroxide spin label specifically reacts with cysteines, but any cysteine residue can potentially be labeled. If multiple residues are spin-labeled in a protein, the DEER signal may not be amenable to deconvolution, unless the distances are different enough in length to produce clearly resolvable signals. If a protein contains many cysteine residues, these residues may need to be removed (generally by mutation to alanine) by site-directed mutagenesis. [97]

Fortunately, while EcoRI does contain one endogenous cysteine (C218), it is buried within the hydrophobic core of the protein and therefore is less accessible to the spin label. Tests on wild-type EcoRI showed that even after exhaustive attempts at labeling, the C218 residue was less than 10% labeled. These results are shown in the Ph.D. dissertation of Katherine M. Stone, Department of Chemistry, University of Pittsburgh. [133]

Another important consideration for DEER experiments is optimizing the signal strength of the sample. As the strength of the dipolar coupling between two spins is proportional to the inverse cube of the distance between them, measuring longer distances requires careful optimization of the signal strength. The distance that can be measured between two spins is limited by the phase memory time, T_m, of the spins. For a time domain signal V(t), the signal is

dampened by an exponential factor of $-2t/T_m$. [111] Therefore, it is important to maximize the T_m for a given sample.

One way to increase the T_m is to conduct the experiments at low temperatures (80K or below). [97], [134] At these temperatures it is important to prevent ice crystals from forming in the sample. Ice crystals exclude macromolecules, which results in inhomogeneity of the sample. This prevents accurate background correction, since V_{intra} is no longer based on a homogenous population. In order to prevent this, glycerol (30%) or sucrose (40%) must be present as cryoprotecting agents. [97] Additionally, T_m is significantly longer in the presence of deuterated solvent. At low temperatures, total deuteration of solvent and sample can increase T_m up to seven-fold, but even partial deuteration of solvent can double the T_m .[111], [121], [135]

2.3 SITE-DIRECTED SPIN LABELING OF ECORI

2.3.1 Selection of residues for SDSL

The amino acid positions chosen for site-directed spin labeling were carefully selected based on a highly refined version of the crystal structure of the EcoRI specific complex. [47], [53] The residues that were chosen have high solvent-accessibility, and with the exception of I197, the residues do not participate in any contacts involved in DNA binding. (Figure 2.3).

Residue R131 (magenta, Figure 2.3) was chosen to investigate the positions and dynamics of the inner arms, while residue S180 (red, Figure 2.3) was chosen to probe the positions and dynamics of the outer arms. Residue K249 (orange, Figure 2.3) was selected as a

reference point, as it resides in a rigid helix in the main domain and has very restricted movement.

Residue I197 (blue, Figure 2.3) was selected as a more direct probe of the protein-DNA interface, as the main-chain carbonyl group of this residue makes a water-mediated contact with the amino group of the cytosine flanking the specific site. (Figure 2.4) As discussed in the introduction, we found that the I197A mutant exhibits a slightly promiscuous phenotype. The I197C mutant was also found to display characteristics of promiscuous behavior: enhanced binding to the specific sequence (discussed below) and enhanced cleavage of miscognate sequences (not shown). This mutant could therefore provide additional insight into the structural characteristics of a promiscuous mutant.

Since EcoRI is a homodimer, the introduction of one cysteine residue will necessarily generate a pair of cysteine mutations in the dimer, facilitating inter-subunit distance distribution measurements with DEER as discussed in a previous section. A double mutant, S180C/K249C was also constructed so that the distance distributions for S180-S180, K249-K249, and S180-K249 could all be compared in order to "triangulate" the position of the outer arms.

The chosen residues were mutated to cysteine via site-directed mutagenesis and confirmed by sequencing. Mutant EcoRI protein was isolated by HPLC to >99% purity as confirmed by gel electrophoresis.

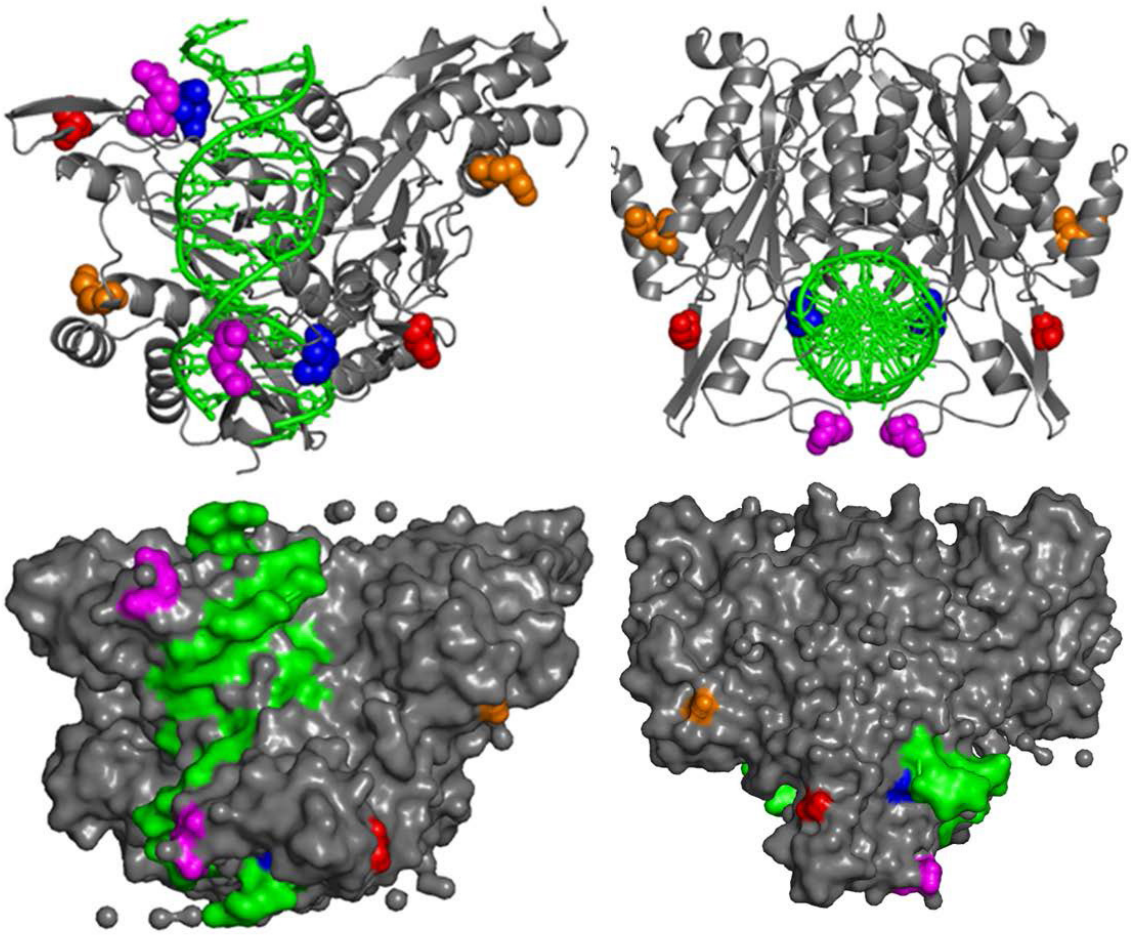


Figure 2.3 Residues selected for site-directed cysteine mutagenesis.

Cartoon representation (top two panels) and accessible surface area representation (bottom two panels) of EcoRI in complex with specific DNA. (Based on a highly refined version of PDB ID: 1CKQ, [47] **The residues chosen for site-directed mutagenesis: R131 (magenta), S180 (red), I197 (blue), and K249 (orange) are shown as colored spheres. It can be seen that all of these residues are exposed to the solvent.**

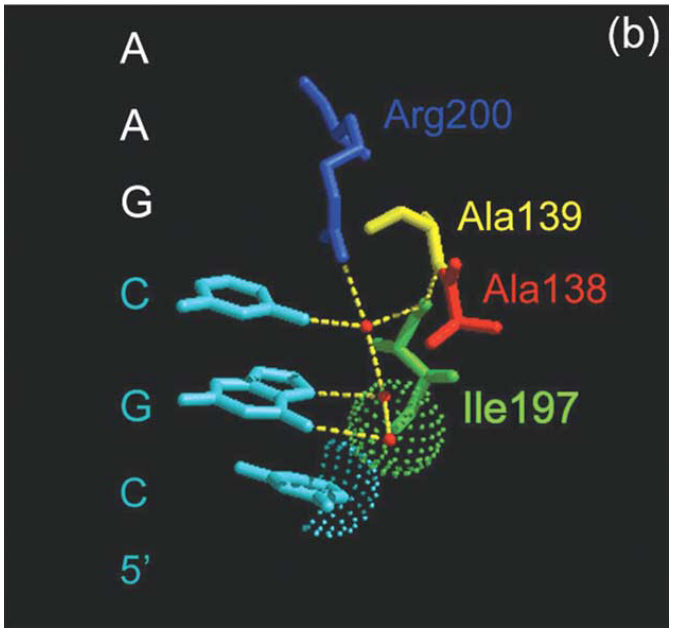


Figure 2.4 I197C forms a water-mediated contact with C-4.

This figure highlights the protein-DNA interface in the vicinity of the A138 and I197 residues in the wild-type specific complex. [47] The bases flanking the recognition sequence are shown in blue. Hydrogen bonds are shown with yellow dashed lines, van der Waals contacts are shown with dotted spheres. Residue I197 forms a water-mediated contact with the C-4 cytosine of the flanking sequence. Figure from Sapienza 2005 [6]

2.4 CHARACTERIZATION OF ECORI MUTANTS

2.4.1 The EcoRI cysteine mutant proteins do not form disulfide bridges

An initial concern in these experiments was whether the presence of a surface accessible cysteine residue would permit the formation of disulfide bridges between dimers, resulting in the formation of tetramers in solution. This was tested by concentrating wild-type and S180C mutant protein to ~100μm and incubating overnight, then checking the oligomerization state via SDS-PAGE in the presence or absence of a reducing agent (Figure 2.5). If disulfide bridges had formed, an additional band running at twice the size of the wild-type protein would have been observed in the samples not subjected to reduction by β-mercaptoethanol. As demonstrated in

Figure 2.5, this additional band is not observed. Furthermore, each of the highly concentrated spin-labeled ESR protein samples were run over SDS-PAGE for densitometry quantification, and additional bands were not observed.

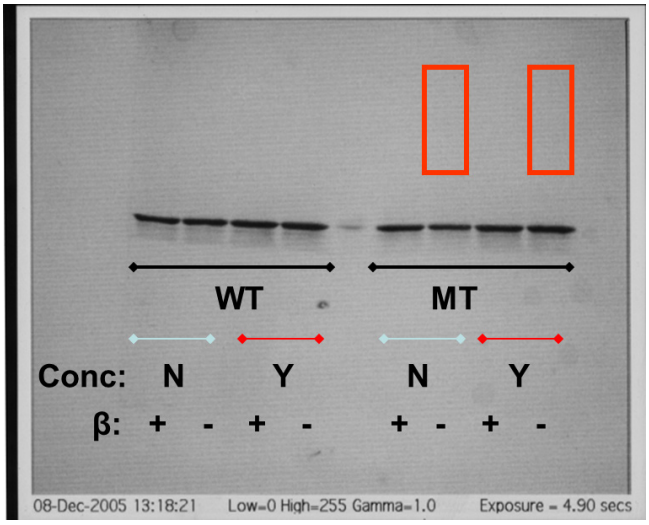


Figure 2.5 Concentrated samples show that wild-type and mutant EcoRI remain in dimer form.

Image of a native SDS-PAGE gel stained with Comassie blue. Pure wild-type and S180C EcoRI were concentrated to ~100μm and incubated for 16hr at 4°C. Half of the sample was subjected to boiling in SDS loading buffer in the presence of β-mercaptoethanol, (β+ samples) and the other half was combined with room temperature SDS loading buffer (β- samples). The samples (red underline) were run next to controls (blue underline) which had not been subjected to the concentration procedure. Red boxes indicate where tetramers would be expected.

2.4.2 EcoRI-cysteine mutants and spin-labeled derivatives exhibit DNA binding affinity that is similar to wild-type binding affinity

In order to confirm that the mutant EcoRI proteins and their spin-labeled derivatives exhibit binding affinities similar to that of the wild-type protein, equilibrium association constants (K_A) were determined using a standard filter binding assay. [136] As seen in Table 2.1, mutant proteins S180C, R131C, and K249C-S180C all show very high binding affinities for the specific site. The decrease in binding is very small when compared to the 51,000 and 96,000 fold decrease in binding of the wild type EcoRI to miscognate and non-specific DNA sites,

respectively. The I197C mutant actually exhibits binding which is *improved* by over 3-fold relative the wild-type binding affinity. For comparison, the promiscuous A138T mutant shows binding which is approximately 14-fold better than wild type for the specific site (in the CGC flanking context). [6]

Before performing the ESR experiments, a "mock" labeling experiment was performed on wild-type EcoRI in absence of the spin label. The purpose of this test was to confirm that the enzyme could survive the multiple steps of the spin-labeling protocol with little loss of activity. As shown in Table 2.1, there was little loss in activity after the procedure. The data obtained for the spin-labeled derivatives (e.g. $S180C_{spin}$) shown here were obtained from samples recovered after collecting ESR spectra. These samples were processed by cleaving the sample DNA as discussed in the methods chapter (Chapter 6) so that the filter binding assay could be performed. As seen in Table 2.1 and Figure 2.6 below, the mutations themselves, spin-labeling, and extensive handling have a combined effect on binding activity that is still relatively modest.

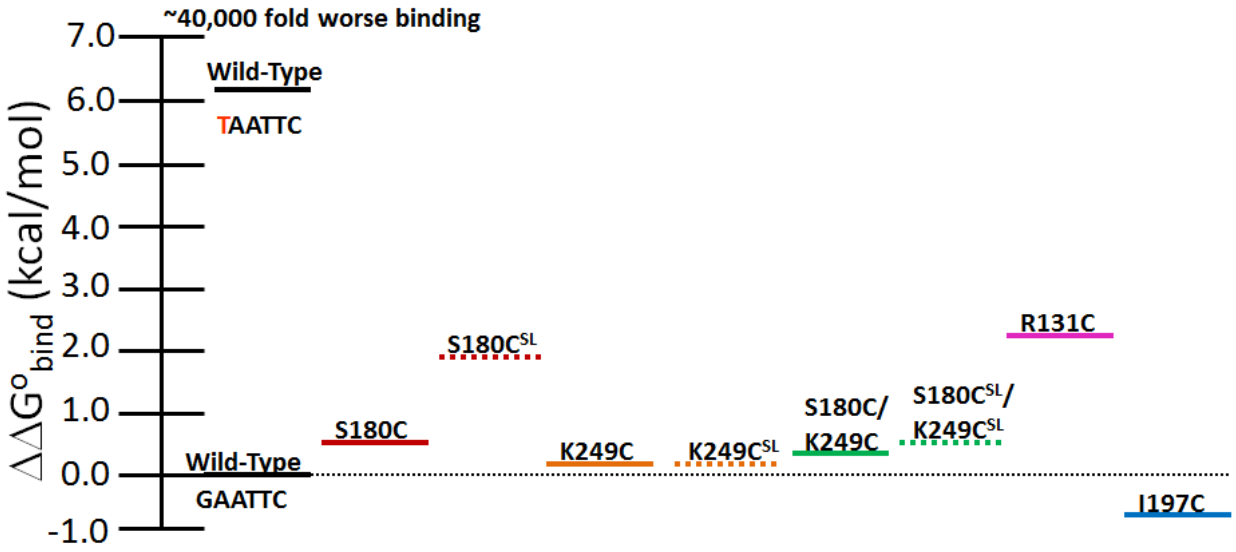


Figure 2.6 Energy level diagram for EcoRI cysteine mutants and spin-labeled mutant protein.

Data are referenced to wild-type EcoRI binding to the specific DNA sequence in the CGC flanking context. ($\Delta\Delta G°_{Bind}$=0). For comparison, the wild type binding to the TAATTC miscognate sequence is also shown. Positive $\Delta\Delta G°_{Bind}$ values represent mutant proteins for which binding is *less* favorable relative to wild type. negative values represent mutant proteins for which binding is *more* favorable than wild type. Data based on numerical values shown in the following table.

Table 2.1 EcoRI Equilibrium binding data for cysteine mutants and spin-labled samples

	$K_A(M^{-1})$	$\Delta\Delta G°_{bind}$(kcal/mol)[a]	Relative Binding (K_AWT/K_A mutant)
Wild Type	4.5 (±1.4) x 10^{10}	0	1
Wild Type Mock	3.0 x 10^{10}	0.2	1.5
S180C	2.3 (± 0.7) x 10^{10}	0.4	2
$S180C_{spin}$	1.8 (± 0.4) x 10^{9}	1.9	25
K249C	3.8 (± 0.4) x 10^{10}	0.1	1.2
$K249C_{spin}$	3.5 (± 0.5) x 10^{10}	0.2	1.3
K249C-S180C	2.3 (±0.7) x10^{10}	0.4	2
$K249C_{spin}$-$S180C_{spin}$	1.3 (±0.3) x10^{10}	0.7	3.5
R131C	1.1(±0.6) x10^{9}	2.2	43
I197C[b]	1.54 x10^{11}	-0.72	0.29[c]

Equalibrium association constants (K_A) were determined in 0.22M KCl, 20mM Cacodylate, pH 7.3, at 21°C unless otherewise noted.

[a] $\Delta\Delta G°_{bind}$=-RTln($K_A^{mutant}/K_A^{wild\ type}$) at 294K.

[b] The I197C constants were determined by rotation student Monique Brisset in 0.24M KCl, 20mM Bis-Tris-Propane, pH7.3. These reported I197C constants represent values which have been normalized to the others.

[c] This represents a 3.4 fold *increase* in binding affinity.

2.5 DEER DATA COLLECTION AND ANALYSIS

2.5.1 Sample preparation for DEER

Purified EcoRI protein was incubated with specific, miscognate, or nonspecific DNA (Appendix) to form protein-DNA complexes, then reacted with at least 100-fold excess of nitroxide spin label in order to ensure maximum saturation of cysteine residues. Excess free spin label was removed via a combination of dialysis and several buffer exchanges in an Icon concentrator. (Pierce). In all cases, excess DNA and appropriate salt concentrations (0.22M) and pH (7.3) were used to ensure that virtually all the protein exists in the bound form.

The protein concentrations of the samples ranged from 75 to 120μm, and the samples were prepared in buffer (20mM phosphate pH 7.3, 0.22M NaCl, 1mM EDTA, 100μm NaN_3) containing 30% glycerol as a cryoprotectant to prevent the formation of ice crystals in the samples. The samples were flash frozen by plunging the capillaries into liquid nitrogen or into propane cooled by liquid nitrogen. Flash-freezing the samples permits us to "capture" the full range of thermally accessible conformations which were present in the population of complexes even though the experiments are conducted at extremely low temperatures. For the S180C, K249C, R123C, and I197C samples, the samples were prepared in a deuterated version of the above buffer (prepared with 30% D-8 glycerol, 65% D_2O, 5%H_2O) to increase the T_m as discussed previously in this chapter.

2.5.2 Data collection

Four-pulse DEER experiments [137] were performed on a Bruker ElexXsys E580 CW/FT X-Band ESR spectrometer using the Bruker X-band ER 4118X-MS2, MS3, split ring resonators or MD5 dielectric ring resonator. DEER data presented in this chapter were collected by collaborators Katherine Stone and Jessica Sarver (Saxena lab). A detailed discussion of the data collection procedure can be found in the supporting information for our published results [102] and the Ph.D. dissertation of Katherine M. Stone, Department of Chemistry, University of Pittsburgh. [133]

2.5.3 DEER Data analysis

Processing and distance distribution analysis of DEER signals for the R131C complexes, S180C-K249C complexes, and the S180C specific and nonspecific complexes were initially performed as described in our published results [102] using the DeerAnalysis2006 program. [138] The time domain signals were inverted to obtain the distance distribution functions, using the Tikhonov regularization method. [134] Based on the resolution of the data, the error in the distances is estimated to be <10% for each measurement. A detailed discussion of the ESR data analysis procedures can be found in the supporting information for our published results [102] and the Ph.D. dissertation of Katherine M. Stone, Department of Chemistry, University of Pittsburgh. [133] The data presented herein for these samples were further optimized by Jessica Sarver (Saxena lab) using the DeerAnalysis2009 program. [114], [138] The data for the S180C miscognate complex, the K249C complexes, and the I197C complexes were collected and processed by Jessica Sarver using the DeerAnalysis2009 program. [114]

2.6 MEANS OF DEER DISTANCE DISTRIBUTIONS REVEAL THAT THE ARMS OF ECORI EMBRACE THE DNA IN ALL THREE CLASSES OF COMPLEX

2.6.1 DEER data

The baseline-corrected time domain signals and resulting distance distributions are shown in the following series of figures (Figure 2.7-2.12). The top panel shows the baseline corrected time domain signals, the bottom panels represent the distance distributions. Data from the specific complexes are shown in green, miscognate complex data are in purple, and nonspecific sample data are in orange.

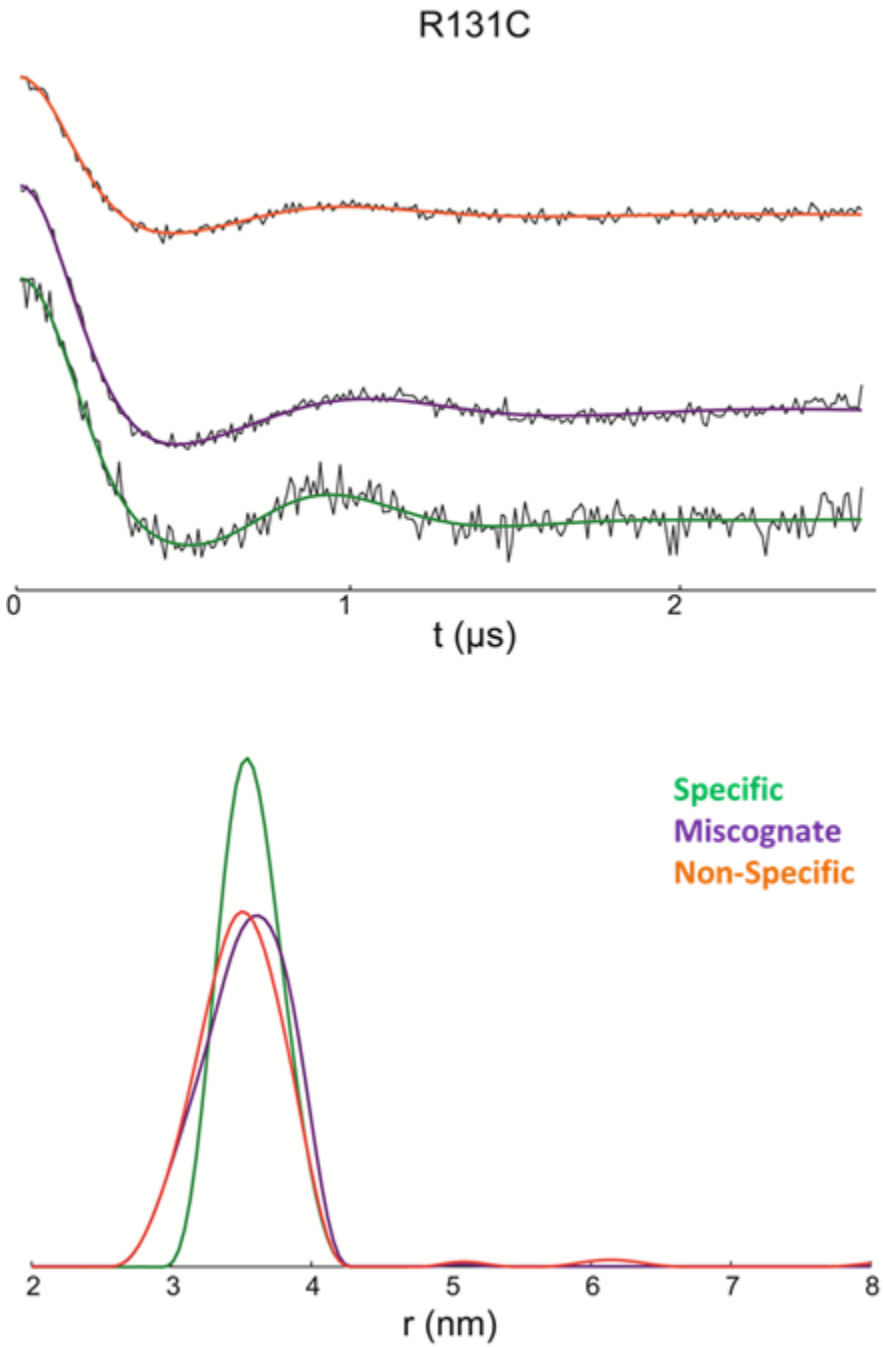


Figure 2.7 R131C DEER.

Top panel: Baseline-corrected DEER data for R131C specific (green) miscognate (purple) and nonspecific (orange) complexes. Simulated traces based on the distance distributions shown on the bottom panel are overlaid on the experimental data. Bottom panel: Normalized distance distribution functions. Data analysis and figure by Jessica Sarver.

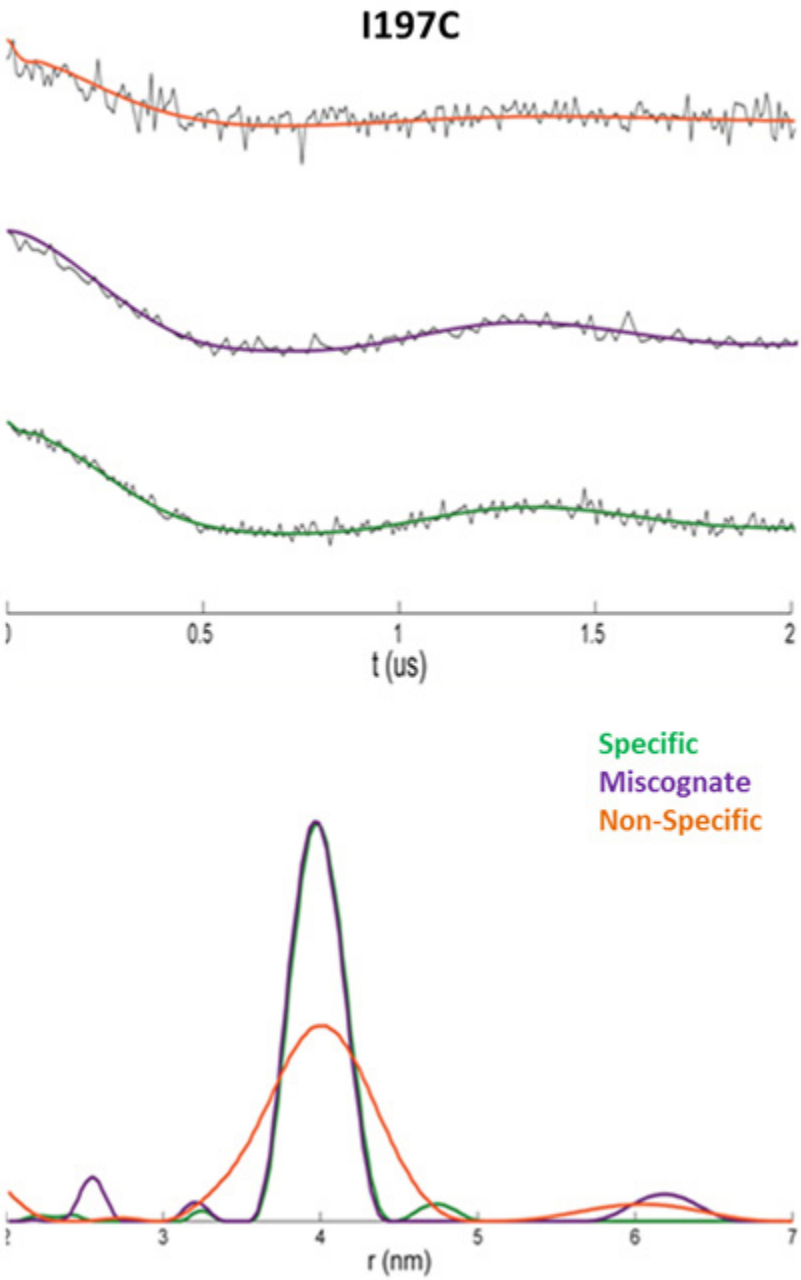


Figure 2.8 I197C DEER.

Top panel: Baseline-corrected DEER data for I197C specific (green) miscognate (purple) and nonspecific (orange) complexes. Simulated traces based on the distance distributions shown on the bottom panel are overlaid on the experimental data. Bottom panel: Normalized distance distribution functions. It should be noted that a very low signal was collected for the nonspecific sample and that this distance distribution is preliminary; the broad distribution in this figure may likely be the result of low signal. This nonspecific sample requires further signal collection and processing before the final results are determined. Data analysis and figure by Jessica Sarver.

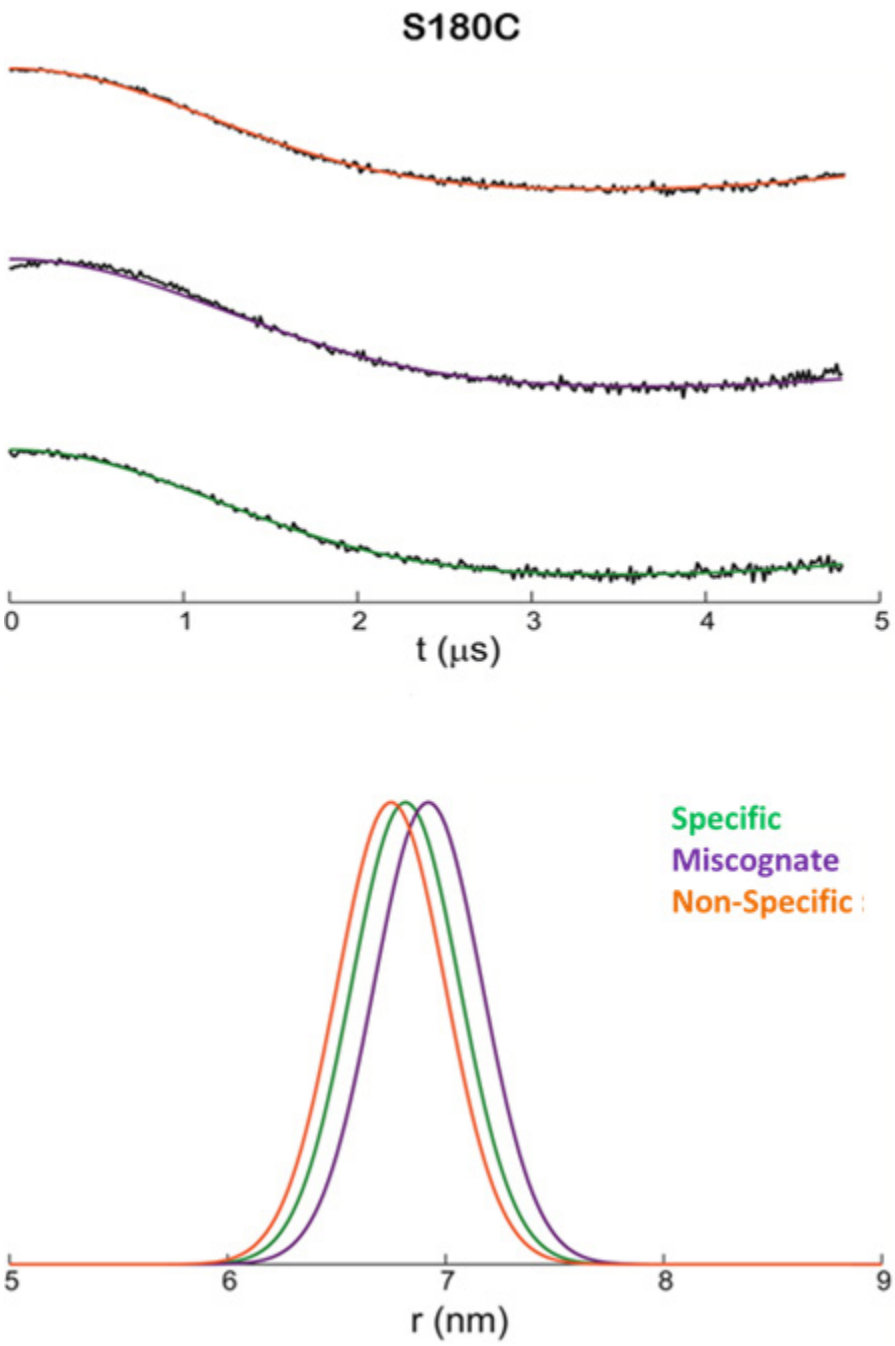


Figure 2.9 S180C DEER.

Top panel: Baseline-corrected DEER data for S180C specific (green) miscognate (purple) and nonspecific (orange) complexes. Simulated traces based on the distance distributions shown on the bottom panel are overlaid on the experimental data. Bottom panel: Normalized distance distribution functions. Data analysis and figure by Jessica Sarver.

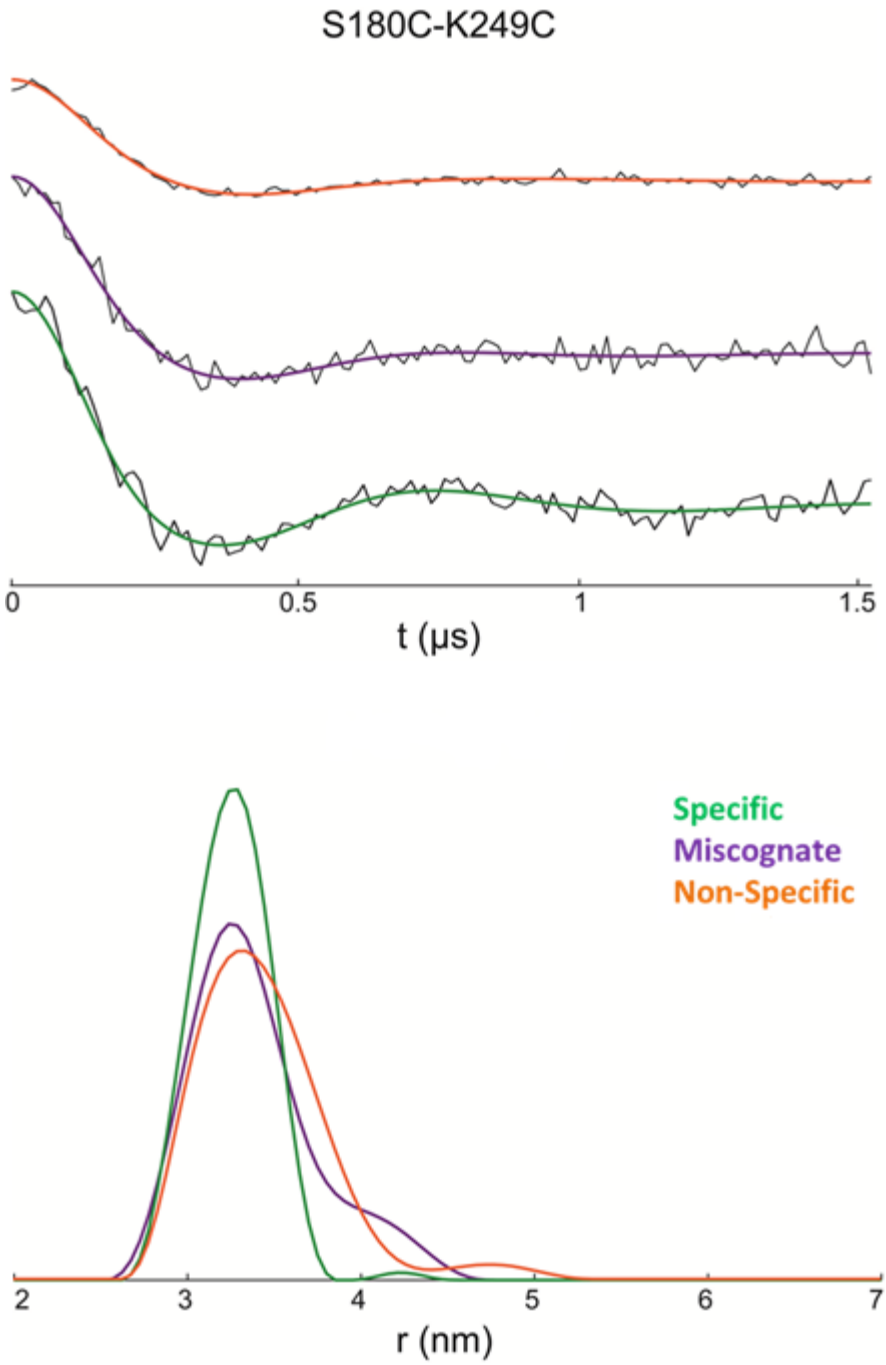


Figure 2.10 S180C-K249C DEER.

Top panel: Baseline-corrected DEER data for S180C-K249C specific (green) miscognate (purple) and nonspecific (orange) complexes. Simulated traces based on the distance distributions shown on the bottom panel are overlaid on the experimental data. Bottom panel: Normalized distance distribution functions. Data analysis and figure by Jessica Sarver.

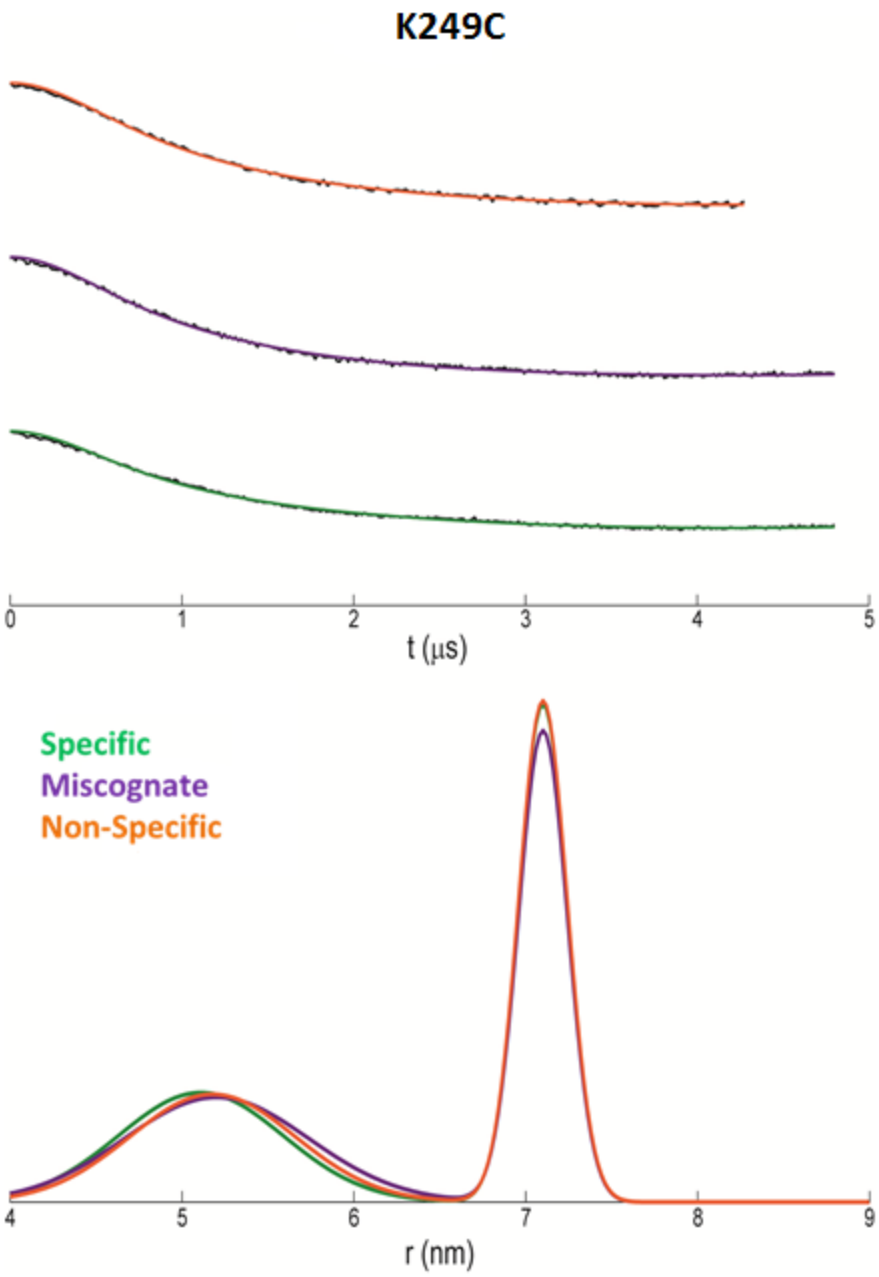


Figure 2.11 K249C DEER.

Top panel: Baseline-corrected DEER data for K249C specific (green) miscognate (purple) and nonspecific (orange) complexes. Simulated traces based on the distance distributions shown on the bottom panel are overlaid on the experimental data. Bottom panel: Normalized distance distribution functions. Data analysis and figure by Jessica Sarver.

2.6.2 Discussion of distance distribution results

2.6.2.1 Most probable distances between residues reveal that the EcoRI arms embrace the DNA the three classes of complex.

The most probable distances between the spin labels are shown in the following table (Table 2.2) as compared to Cβ-Cβ distances measured from the specific complex crystal structure. As can be seen in the table, the distances for the specific complexes are comparable to those determined from the specific complex crystal structure. [47] Due to the length of the nitroxide spin label, the distances obtained from SDSL-DEER can differ from crystal structures by up to ~10Å, depending on the orientation of the spin label. [97] The distances obtained for the specific complex DEER samples all fall within this expected range.

Table 2.2 Comparison of specific complex crystal cβ-cβ distances with distances obtained from DEER

Complex	Specific	Miscognate	Nonspecific	Specific Crystal
R131C	35 Å	36 Å	35 Å	32 Å
I197C	40 Å	40 Å	40 Å	37 Å
S180C	68 Å	69 Å	67 Å	59 Å
S180C-K249C	33 Å	33 Å	33 Å	27 Å
K249C	51 Å / 71 Å	51 Å / 71 Å	52 Å / 71 Å	60 Å

The most probable distances for the R131C EcoRI specific, miscognate, and non-specific complexes are all approximately 35-36Å. The most probable distances for the I197C complexes are all approximately 40Å. These two residues both fall within the inner arm regions of the

enzyme, and therefore we observe that the point-to-point distances across the DNA do not change much for the noncognate complexes at these positions. The most probable distances for the S180C mutant in all three complexes are all approximately 67-69Å, so the point-to-point distances across the DNA do not vary much for the outer arm position either. Although the mean distances do not change, the distance distributions show differences between the complexes as discussed in the next section.

Two distinct peaks were observed for the K249C sample, at ~ 51Å and 71Å. These could potentially represent two different conformations at this location, or an inter-dimer distance resulting from interactions between some of the protein dimers. The CW spectra for K249C samples appear to exhibit two components. (CW data are shown in Chapter 3.) This indicates that the spin label is experiencing two different "environments". This led us to tentatively assign the origin of the two peaks to two different orientations or conformations of the spin label at the K249C location. However, preliminary molecular dynamics results by collaborator Jessica Sarver indicate that the existence of such a conformation is unlikely. In order to obtain the 51Å distance, the spin label would have to adopt an orientation where it protrudes into the protein. (Jessica Sarver, personal communication). Another potential explanation is that some of the EcoRI complex in these samples has oligomerized in a way which positions the K249C sites such that the 51Å distance is the result of an intermolecular dipolar coupling, although there is currently no evidence for such an oligomer forming in solution. The true origin of this peak is currently unknown.

For the specific, miscognate, and non-specific complexes of the K249C-S180C mutant protein, the most probable experimental distance was 33Å for all cases. In principle, multiple distances corresponding to S180C-S180C, K249C-K249C, and S180C-K249C would be

anticipated for the K249-S180C double mutant. The corresponding Cβ-Cβ distances in the specific complex crystal structure are 27Å (S180C-K249C intra-monomer), 59Å (S180C-S180C), 60Å (K249-K249), and 57Å (S180-K249 inter-monomer). It is likely that the larger distances were not detectable in this series of experiments given in these experiments the phase memory times (T_m) were too short to detect the longer distances. The 33Å peak for the double mutant can thus be assigned to the S180C-K249C intra-monomer distance.

Strikingly, the experimental point-to-point distances are very similar for specific, and for all for the non-cognate EcoRI-DNA complexes for all residues examined. The data show preservation of the distances between the end of the inner arms (R131C), the inner arms within the binding site (I197C), across the outer arms (S180C), from the outer arm (S180C) to a fixed reference point (K249C) in the main domain, and across the reference point in the main domain (K249C). Taken together, the data suggest that EcoRI arms envelop the DNA and are similarly oriented in both noncognate and specific DNA complexes. This implies that the DNA in the specific and noncogate complexes occupies roughly the same binding cleft of the EcoRI dimer.

2.6.2.2 Differences in the distance distributions for the complexes may indicate that the nonspecific complex has greater access to thermally available conformations

For both the R131C and the K249C-S180C mutant proteins, the distance distribution is distinctly narrower for the specific complex than for the corresponding non-cognate complexes. This might indicate that the arms in the EcoRI complex with non-cognate DNA have greater access to thermally accessible conformations. In addition, the K249C-S180C distances show asymmetries in the non-cognate complexes. However, it is unclear if this represents an asymmetric set of accessible conformations of the arms or different orientations accessible to the spin labels. Although there are no significant differences apparent between the distance distributions for the

S180C complexes, the experimental resolution for large distances, combined with backbone and spin-label flexibilities in the arm regions of the protein, precludes detection of differences in the distributions for S180C specific and noncognate complexes. Therefore we cannot definitively say whether there are differences in the S180C distance distributions or not. The differences in the distance distributions for the DEER signals may originate from greater freedom or restriction of the nitroxide label itself, from local fluctuations of the protein carbon backbone, or from differences in the overall conformational freedom of the protein for these complexes. [78], [107], [139]

Further ESR experiments were designed to investigate potential differences in the dynamics between the complexes and are described in Chapter 3. Additionally, molecular dynamics simulations are currently being performed by collaborator Jessica Sarver (Saxena lab) in order to model the nitroxide sidechain mobility and orientation, and to assist in the deconvolution of the local backbone motions at the spin-labeled sites. The combination of molecular dynamics with DEER has been shown to increase the resolution of distance and distance distribution results. [97], [140], [141] Preliminary results from the simulation distance distributions agree with our experimental results in terms of both average distances and distance distributions. These simulations may help us increase the resolution of the distance distributions, especially for the longer distances such as the S180C-S180C data. Deconvolution of the contributions of the nitroxide side-chain mobility from the local backbone motions should provide more detailed insight into the origins of the differences seen in the distance distributions.

2.6.3 Implications for the "facilitated diffusion" model of site association

A number of kinetic studies have shown that EcoRI [142-144] like other DNA binding proteins ([145], [146]) finds its correct recognition site by "facilitated diffusion". These kinetic studies have shown that many site-specific DNA binding proteins can locate their specific site faster than the maximum diffusion-controlled "limit" that would be expected if the proteins were finding their sites by diffusing freely in three-dimensional space. [145], [147] According to this model, the protein must first bind to nonspecific DNA and then slide or hop to locate its cognate site by one-dimensional diffusion (As reviewed in [146]).

Consistent with this model is the fact that the slopes of the salt dependence for formation of specific and non-specific complexes are the same ($d\log K_A/d\log[NaCl] \approx 11$) and are consistent with the number of Coulombic interactions observed in the specific complex (using a 6 Å cutoff for cationic residues proximal to DNA phosphate) indicating that the same number of Coulombic interactions are formed in both complexes. [7] This provides additional strong evidence that the EcoRI arms encircle the DNA in both specific and nonspecific complexes. This enfolding may contribute to processivity as the protein slides along nonspecific DNA to locate its specific recognition site. [3], [7], [142-144], [146] It is striking that while sliding along the DNA to locate its GAATTC site, EcoRI embraces the nonspecific DNA.

2.7 CONCLUSIONS AND FUTURE DIRECTIONS

Our results on a protein-DNA complex by pulsed ESR have established a methodology that can measure the solution structure and range of conformational states for complexes with different

classes of DNA sites for which there is little or no prior structural information. We have determined that for EcoRI miscognate and nonspecific complexes, the DNA occupies the same binding cleft as specific DNA. Additionally, the "arm" regions for the noncognate complexes occupy approximately the same conformational position as they do in the specific complex.

Tantalizingly, it appears from comparison of the distance distributions that in the noncognate complexes the arm regions may exhibit greater conformational freedom than in the specific complex. Investigation of the potential differences in dynamic behaviors between the EcoRI-DNA complexes are discussed in the next chapter. Molecular dynamics simulations are currently underway in order to increase the resolution of these experiments. These simulations are currently being conducted by Jessica Sarver (Saxena lab). By deconvolution of the contributions from the nitroxide mobility and the backbone fluctuations to the distance distributions, we should achieve a more detailed view of the differences in conformational freedom in the different classes of complex.

3.0 CONTINUOUS WAVE EXPERIMENTS REVEAL SUBTLE DIFFERENCES IN THE DYNAMICS OF ECORI-DNA COMPLEXES

3.1 INTRODUCTION

As discussed in Chapter 1, one aim of this work is to apply emerging ESR techniques for dynamics measurements to examine differences in the dynamic behaviors in the EcoRI enzyme when bound in the specific, miscognate, and nonspecific EcoRI-DNA complexes. The DEER experiments described in Chapter 2 revealed that the arms of the protein embrace the DNA in approximately the same position for all three complexes. Tantalizingly, the R131C and the K249C-S180C complex distance distributions were distinctly narrower for the specific complex than for the corresponding non-cognate complexes, indicating that the arms may experience greater conformational freedom in the noncognate complexes. In this chapter, I review how information about the motions of a nitroxide spin label may be obtained from ESR spectra and describe our results examining the subtle differences between the specific, miscognate, and nonspecific EcoRI-DNA complexes.

3.2 MEASUREMENTS OF DYNAMICS WITH ESR

It has long been appreciated that the dynamics associated with the backbone and side-chain mobilities of a protein are an essential aspect of its function. [41], [42], [148], [149]. Investigating the dynamic motions of a protein and changes in these motions upon ligand binding is valuable towards gaining a more complete understanding of how structure confers function in a protein. A wealth of information about the dynamics of a protein can be obtained from SDSL-ESR. [78] Analysis of CW lineshapes can provide information about protein backbone dynamics, conformational changes, and local secondary structure. [78], [84-86], [100]

3.2.1 Introduction to CW

In Chapter 1, I briefly described how the ESR signal originates from an unpaired electron changing its spin state in the presence of an external magnetic field. (This splitting in energy states is also referred to as the Zeeman effect.) For this simple case, the Hamiltonian describing the energy of the system is:

$$H_{spin} = g\beta_e B_0 S_z$$

Where g is the electron g factor, β_e is the Bohr magneton (a constant), B_0 is the magnetic field strength, and S_z is the electron spin's angular momentum operator (±1/2). (Reviewed in [92], [150])

In the more complex case of nitroxide spin labels, the nitrogen nucleus itself has three possible spin states (1, 0, and -1). These states interact with those of the coupled unpaired electron in a way that results in further splitting of the potential energy levels experienced by the spin. This is referred to as the "hyperfine splitting" of the spectrum. The resulting energy level

diagram has six potential energy levels. However, in the ESR experimental spectrum (MW frequencies), only the electron transitions are excited so the nuclear spins do not change. [150] This results in three "allowed" transitions in the ESR experiment as shown in Figure 3.1 below. [92], [150] Therefore, the characteristic lineshape for a nitroxide residue consists of three peaks of equal height (Figure 3.2-A). (Reviewed in [92], [94], [150]) In some cases, the spin may interact with other nearby magnetic nuclei, resulting in a more complex lineshape.

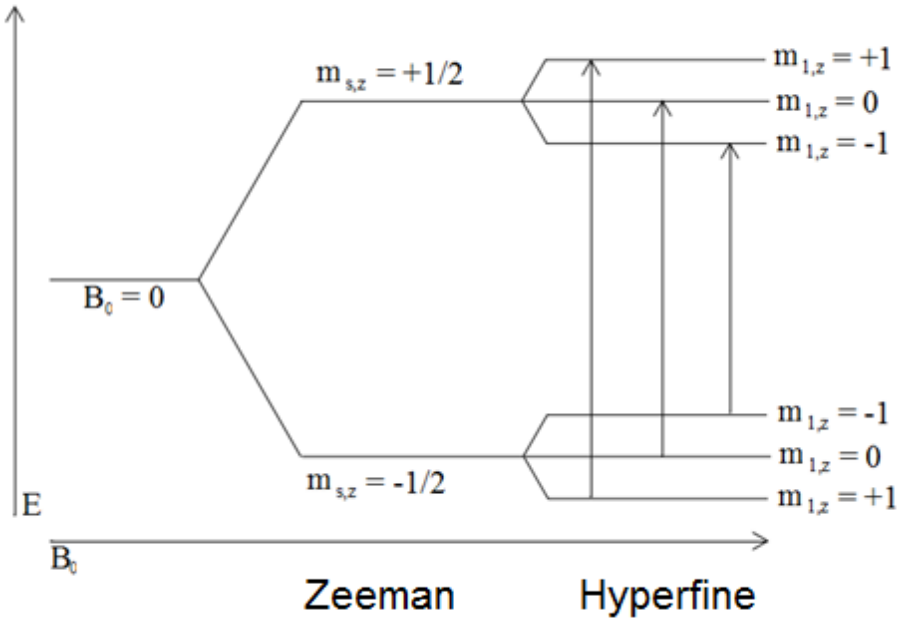


Figure 3.1 Energy level diagram for an electron spin in the context of a nitroxide residue.

The Y axis (E) represents the difference in energy, while the X axis (B_0) shows the magnetic field strength. The first split ($m_{s,z}$ = +1/2, -1/2) represents the two possible energy states for the electron spin. The interaction of the electron spin with the nitrogen nucleus of the nitroxide (which has three possible spin states, $m_{1,z}$ = +1, 0, or -1) results in further splitting of the energy states. The three allowed transitions for an electron spin in the ESR experiment are shown with arrows. Figure adapted from [150], [151]**.**

The Hamiltonian for a nitroxide spin label therefore must include the contribution from the hyperfine interaction as well as from the Zeeman interaction. (Reviewed in [92], [150], [152]). This equation can be written as:

$$H_{spin} = g\beta_e B_0 S_z + aS_z I_z$$

The first term describes the Zeeman interaction as shown previously (the interaction between the electron spin and an external magnetic field). The second term describes the hyperfine interaction (the interaction between the electron spin and the nuclear spin of the nitrogen), where I_z describes the nuclear spin state (1, 0, or -1) and "a" represents the hyperfine splitting constant which describes the strength of the interaction between the unpaired electron spin and the nuclear spin. An important point that can be seen from this equation is the fact that while the term describing the Zeeman interaction is dependent on the magnetic field strength, (B_0), the term describing the hyperfine interaction is independent of the strength of the external magnetic field. This becomes important when considering how field strength and frequency influence lineshapes as discussed in a later section.

3.2.2 The influence of orientation and motion of spin label on CW lineshapes

The previous section describes the spectrum of a nitroxide in the "isotropic" or orientation-independent case. Both the hyperfine coupling constant and the g value are orientation-dependent; the orientation of the spin relative to the external magnetic field influences the value of these terms. (The other contributions to the Hamiltonian are orientation-independent.) When the nitroxide is freely and rapidly tumbling in solution, these orientation effects are averaged out

and the result is three sharply-defined peaks of equal width as shown in Figure 3.2-A. This case is referred to as the fast motional limit. (Reviewed in [94], [152], [153])

However, when the motion of the nitroxide is restricted, such as when it is attached to a peptide, the g and A values must be treated as tensors which no longer reflect a simple averaging of the orientation effects, and the CW lineshape reflects greater anisotropy (orientation dependence). [94], [152], [154] Figure 3.2 shows a series of CW spectra which illustrate the effect of increased restriction of a nitroxide on broadening of the peaks in CW spectra. A free nitroxide (Figure 3.2 -A) shows the "isotropic" spectra of three sharp, equivalent peaks. If the label is bound to a small, denatured peptide (Figure 3.2 -B), there is a slight degree of restriction causing asymmetry to appear in the lineshape. If the label is bound to a short α-helix, even greater broadening is seen (Figure 3.2-C). If the protein is frozen in solution, the spin label motion is highly restricted – this is the opposite extreme from the fast motion limit. In this case, referred to as the "slow motion limit" or "rigid limit" spectrum, the spectra shows a great degree of broadening. For the standard X-band CW experiment, the spectrum is sensitive to motions on time scales between ~0.1-100ns. [78], [94], [153]

Overall, relative broadening of the peaks in a CW spectrum generally indicates increased immobilization of the spin label, while relative sharpening indicates increased mobility. Examination of the lineshape provides information on the mobility and local structural environment of a spin label. Changes in the protein that occur, for example, upon binding of a ligand can be investigated by analyzing the broadening or sharpening of lineshapes. [94], [98-100]

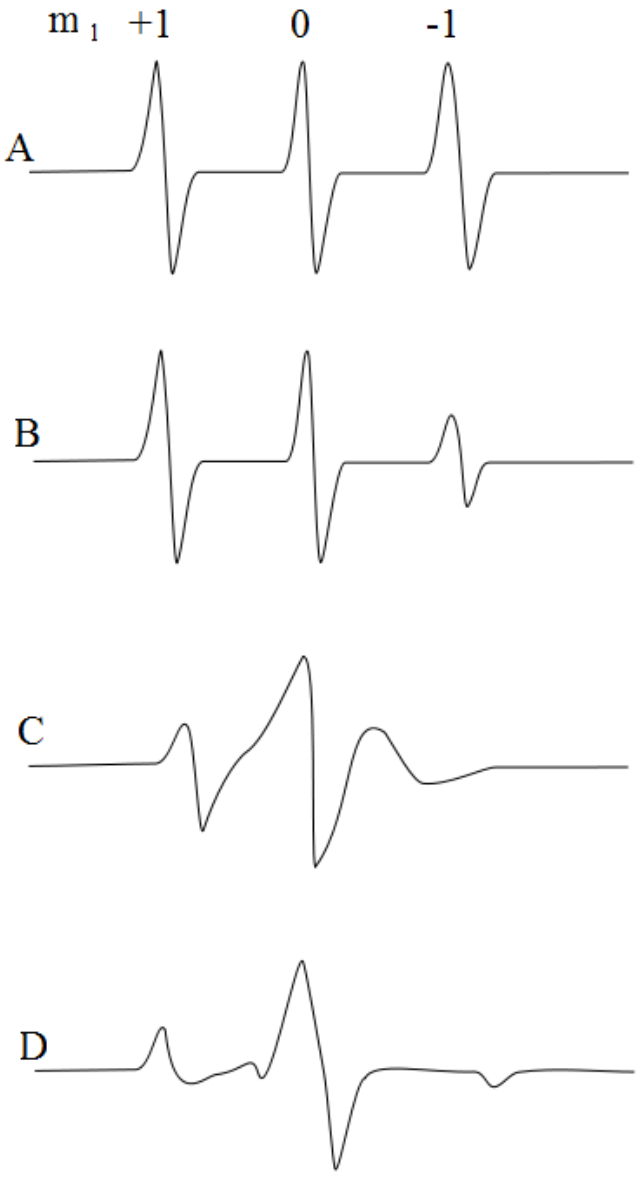


Figure 3.2 Sensitivity of CW lineshapes to the motion of the spin label.

The above panels show CW spectra for the following: A) A dilute solution of nitroxide spin label (MTSL) in solution. B) The spin label bound to small denatured peptide. C) The spin label bound to the same peptide in its folded state of a short α- helix. D) The same peptide frozen in solution. Figure adapted from [94]**.**

3.2.3 Contributions to motion of the nitroxide spin label

In order to interpret CW lineshapes for a spin-labeled protein and the changes that occur in these spectra upon events such as ligand binding, it is important to consider the various contributions to the motion (or restriction of motion) of the spin label. Firstly, the nitroxide spin label itself rotates around the bonds connecting it to the protein, producing motion of the nitroxide relative to the protein backbone. Secondly, the local fluctuations of the carbon backbone of the peptide will influence the motion of the label. Thirdly, the global tumbling of the protein itself can contribute to the spectrum. [78], [152], [155] (Figure 3.3) Generally it is the local backbone fluctuations which are of interest, as investigating these motions permits the exploration of protein dynamics. [78], [107]

These three primary modes of motion are generally described in terms of their correlation times. These are the rotational correlation time (τ_R) which represents the rotation of the entire protein, τ_B, which describes the rotation around the bonds connecting the nitroxide to the protein, and τ_S, which describes the motion of local backbone fluctuations relative to the rest of the protein structure. (Figure 3.3) [86] The rate of the global tumbling for large proteins is generally relatively slow relative to the ESR timescale (~50ns or longer), while the rate of the nitroxide motion is fast (on the order of ~1ns). [100] The rate of global tumbling for a protein is dependent on the size of the protein and the viscosity of the solution, as described in the following section.

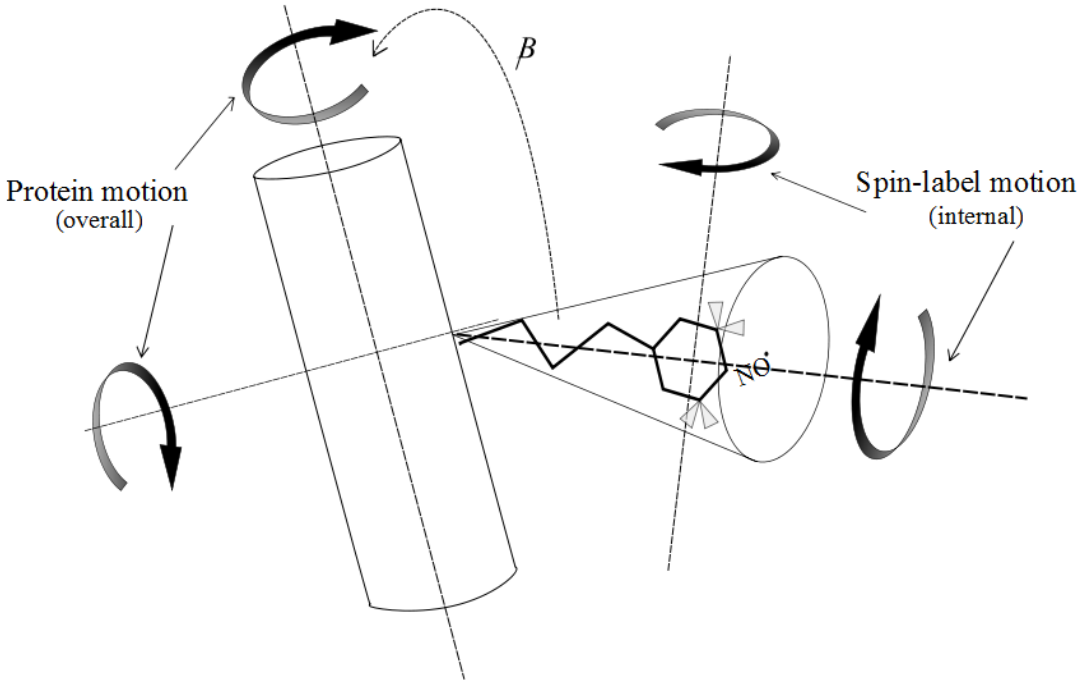


Figure 3.3 Contributions to the motion of a nitroxide spin label.

The internal motion of the nitroxide itself is largely restricted to a cone at an angle (β) to the rest of the protein. The spin labeled residue also undergoes internal motions, as does the backbone of the protein (not shown). The protein as a whole exhibits a global tumbling motion in as it undergoes rotary diffusion in solution. The amplitudes, angles, and rates of these contributions to the motion are all independent of each other. Adapted from [155]**.**

Despite the complexity of these contributions to the motion of a nitroxide bound to a protein, the inverse line width of the central resonance peak (δ^{-1}) is frequently used as a measure of the relative mobility of a spin label. For cases in which the global tumbling is too slow to influence the CW spectrum (this is discussed in next section), this parameter reports on the combined effect of the nitroxide spin label and the local fluctuations of the protein backbone. This semiquantitative parameter has found to be highly correlated with the motion of the local protein backbone fluctuations and has been established as a measure that reports on the local motions of the protein. [78], [84], [85], [107], [139], [156]

3.2.4 The effect of the global tumbling of the protein on lineshapes

The rotational motion of the protein itself (as described in the previous section) can be a confounding factor when trying to interpret CW lineshapes. The rotational correlation time for a protein tumbling in solution can be approximated by the Stokes-Einstein equation:

$$\tau_R = (\eta V/k_B T)$$

where η is the viscosity of the solution, k_B is Boltzman's constant, T is the temperature, and V is the volume of the protein. In aqueous solution conditions with low viscosity, rotational correlation times for globular proteins with mass above 50-200kDa (depending on the shape of the protein) can often be ignored, as for these proteins the τ_R is too slow for the ESR experiment to detect. [86], [157]

For smaller proteins, it is prudent to investigate whether the rotational motion of the protein is making a significant contribution to the ESR lineshapes. This can be done by increasing the viscosity of the solution used in the experiment in order to reduce τ_R. Adding

either 30% sucrose or 25% Ficoll will increase the viscosity of an aqueous solution ~3-fold, such that a comparison of the "low" and "high" viscosity conditions would show significant differences. [86], [107], [158], [159] Both sucrose and Ficoll have been shown to remove the effect of protein rotary diffusion without influencing the motion of the nitroxide. [159]

Another way to minimize the contribution of the global tumbling to the ESR spectrum is to conduct the experiments at high frequency, as discussed in the next section.

3.2.5 High field/high frequency CW experiments

The frequency which is most commonly used for collecting CW spectra is ~9.5GHz. The experiments shown in the lineshapes for previous figures are of this type, which are referred to as X-band experiments. At X-band frequencies, the ESR spectra are dominated by the anisotropy of the hyperfine interaction. [160]

As discussed previously, the Zeeman interaction is dependent on the magnetic field strength, while the hyperfine interaction is not. In the presence of high magnetic fields, the components of the Zeeman term (the g tensor) start to dominate the spectrum. [152] As shown in Figure 3.4 below, conducting ESR experiments at high field strength greatly improves the resolution of the spectra, enabling the direct determination of the components of the g tensor. [157], [161] Determination of these components can provide additional information about the orientation and mobility of the spin label, potential interactions between the nitroxide and adjacent residues, as well as the polarity of the local microenvironment of the nitroxide. [86], [96], [139], [155], [158], [161]

Recent technological advances (Reviewed in [152], [162]) have made W-band experiments (conducted at ~95 GHz) more accessible for the analysis of spin-labeled biological samples. Advantages of conducting experiments at higher frequencies include:

1) Increased signal to noise ratios, enabling the use of smaller samples. [161] For example, our CW X-Band experiments required ~5ul of sample, while the corresponding W-band experiments required <1ul of sample.
2) A dramatic gain in resolution of components to the g-factor (as described above and shown below in Figure 3.4).
3) High frequency spectra are more sensitive to fast motions (ps time scale) that are averaged out in low frequency experiments (ns time scale). This provides a picture of the molecular motion on a faster time scale than the X-band data provides. On the other hand, motions that are detected by the X-band spectra are too slow to be observed at W-band frequencies. [163] A number of studies have addressed the value of comparing CW spectra collected at different frequencies in order to examine the motions of the spin label on different time scales. [155], [158], [163-165] This also aids in the deconvolution of mobility contributions, since for W-band experiments the rotational motion of the protein is too slow to influence the spectrum. [163]

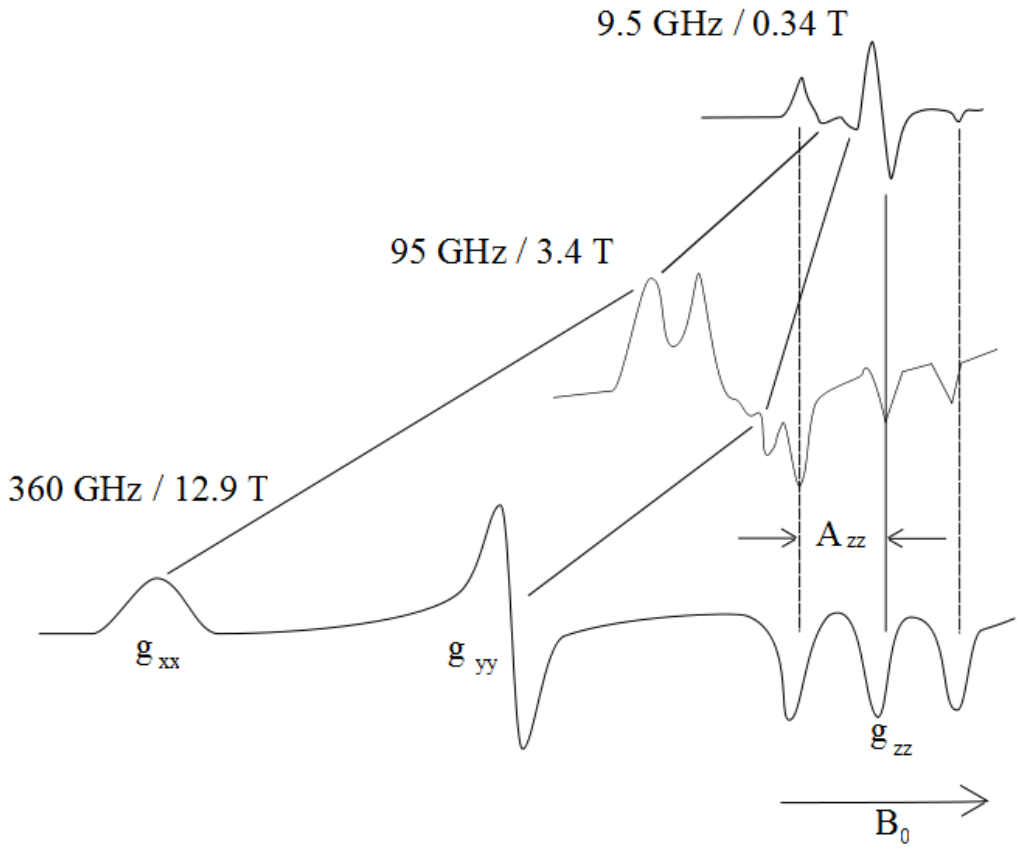


Figure 3.4 The effect of increasing frequency and field strength on the resolution of CW spectra.

Shown above are CW spectra for the same sample at three different frequency and field strengths. At the "standard" X-band frequency and field strength (9.5 GHz/0.34T), the anisotropic components of the hyperfine interaction, particularly the Azz component, dominate the spectrum. At the high W-band frequency and field strength, (360GHz/12.9T) the anisotropic components of the Zeeman interaction (the g tensors g_{xx}, g_{yy}, and g_{zz}) dominate the spectrum instead and are well-defined. The A_{zz} components can still be seen as a splitting of the g_{zz} peak. After Mobius 2005. [154]

3.3 CW METHODOLOGY

3.3.1 Choice of residues

The residues we selected for CW are the same as those described for the DEER experiments in Chapter 2. In addition, we chose to construct the R123C mutant for CW-ESR. The residue 123C lies on the "hinge point" of the arm and was selected so that we could obtain dynamic information at an additional location of the arms. An important consideration for CW experiments is that of the potential for dipolar coupling to influence lineshapes for CW spectra. These residues were all chosen such that the inter-residue distance between them is longer than the distances at which dipolar coupling makes a significant contribution to CW spectra (~8-25Å). [100]

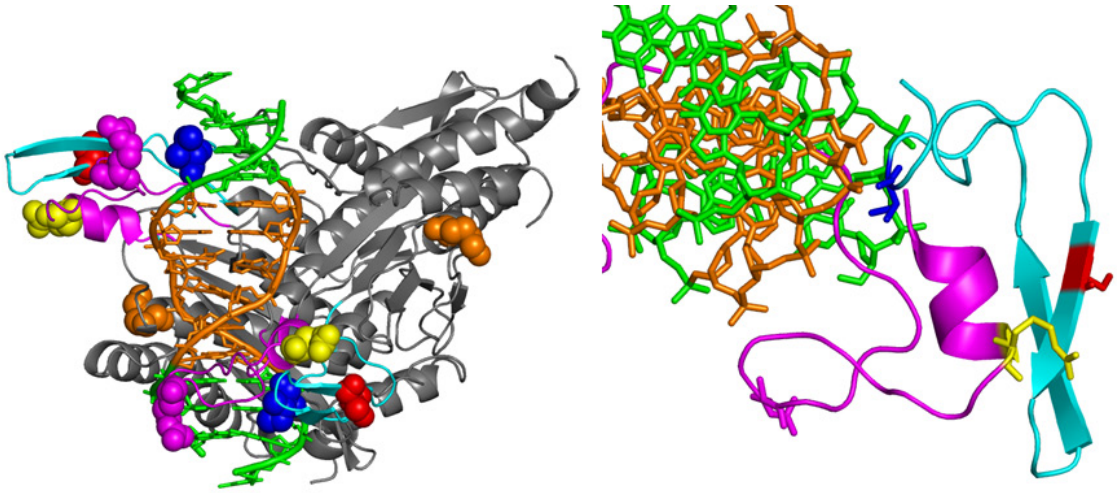


Figure 3.5 Residues chosen for site-directed-spin-labeling for CW-ESR experiments.

Cartoon representation of EcoRI in complex with specific DNA. (Based on a highly refined version of PDB ID: 1CKQ, [47] **The residues chosen for site-directed mutagenesis: R131 (magenta), S180 (red), I197 (blue), R123 (yellow) and K249 (orange) are shown as colored spheres in the left panel and colored sticks in the right panel. The left panel shows the entire protein-DNA complex, the right panel shows a zoomed in view of the DNA and the arm of one monomer of the protein. From this it can be seen that R131 resides on a loop at the end of the "hand", R123 at the end of an α-helix in the arm, S180 in a β-strand of the arm, I197 in a loop adjacent to the binding site, and K249 in the center of an α-helix in the main domain.**

3.3.2 Sample preparation for CW experiments

The protein-DNA complex samples for CW were prepared and spin-labeled as described for the non-deuterated DEER samples in Chapter 2 and the Methods chapter, with the following exceptions. The standard CW buffer contained glycerol at a concentration of 10% instead of 30%. The CW experiments were performed at higher temperatures (243-293 K) than the DEER experiments (40-80 K), so high concentrations of glycerol were not required to prevent ice crystal formation as they were for the DEER experiments. Additionally, in the case of the viscosity studies described in a subsequent section, different concentrations of glycerol or Ficoll 70 were added in order to alter the viscosity of the samples.

3.3.3 CW Data collection and analysis

The 9.5 GHz continuous wave (CW) ESR spectra were collected using 3-5 μL of sample in a 0.8 mm I.D. quartz capillary tube. For the 95 GHz spectra, less than 1 μL of sample was placed in a 0.1 mm ID quartz capillary tube. The 9.5 GHz CW experiments conducted at temperatures from 243K-283K were performed on a Bruker ElexXsys E580 CW/FT X-Band ESR spectrometer using the Bruker X-band ER 4118X-MS2, MS3, split ring resonators or MD5 dielectric ring resonator. The 95GHz CW experiments were performed using a Bruker W-band EN600-1021H resonator at the Bruker Biospin facility in Billerica, MA., under the supervision of Ralph Weber (Bruker Biospin). CW data presented in this chapter were collected and analyzed by collaborator Jessica Sarver (Saxena lab). A detailed discussion of the data collection and analysis procedures

will be forthcoming in future publications and the doctoral dissertation of Jessica Sarver, Department of Chemistry, University of Pittsburgh.

3.4 X-BAND CW EXPERIMENTS REVEAL CHANGES IN MOBILITY FOR ECORI-DNA COMPLEXES

The X-band CW spectra are shown in Figures 3.6-3.8. While the CW experiments report on the overall motion of the spin label, thereby providing information on the dynamic behaviors of the protein, an important caveat of these experiments is that the *origins* of those motions come from a number of sources. As discussed in a previous section, the spin label reports on motions including the global tumbling, local backbone dynamics of the protein, and changes in the rotational motion of the spin label itself, which can originate from spatial constraint of the nitroxide or interactions between the nitroxide and nearby side-chains. ([78], [86], [96], [107], [166] Furthermore, the detection of the motions of the spin label is highly sensitive to the frequency at which the experiments are performed, such that at any given frequency the motions which are too fast to detect are "averaged out", and motions which are too slow to detect are "frozen out." It should therefore be noted that the spectra collected at any one frequency reflect the behaviors of the spin label *which can be detected at that frequency*. [107], [154], [155], [158], [163] Ultimately these caveats reflect an advantage of the methodology, as a great wealth of information about the dynamic behaviors of a protein can be extracted from these experiments; however, it is often necessary to collect spectra under many different conditions, as well as conducting computational simulations of the spectra and molecular dynamics simulations of the spin label in order to fully extract the molecular contributions to the observed spectra.

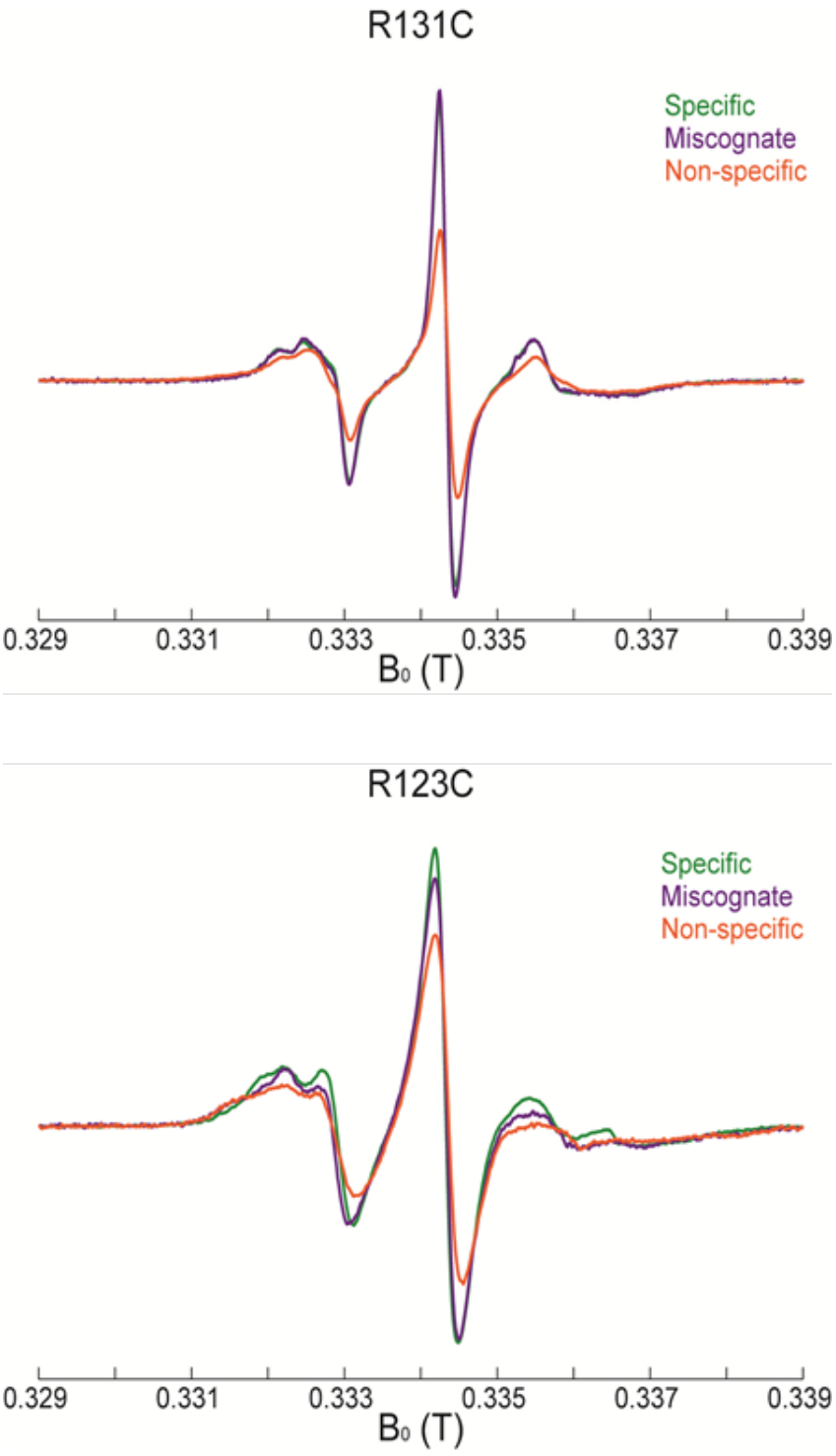


Figure 3.6 X-band continuous wave spectra for R131C and R123C complexes.

Top panel: X-band CW spectra data for R131C specific (green) miscognate (purple) and nonspecific (orange) complexes. Bottom panel: CW spectra data for R131C specific (green) miscognate (purple) and nonspecific (orange) complexes. Data analysis and figure by Jessica Sarver.

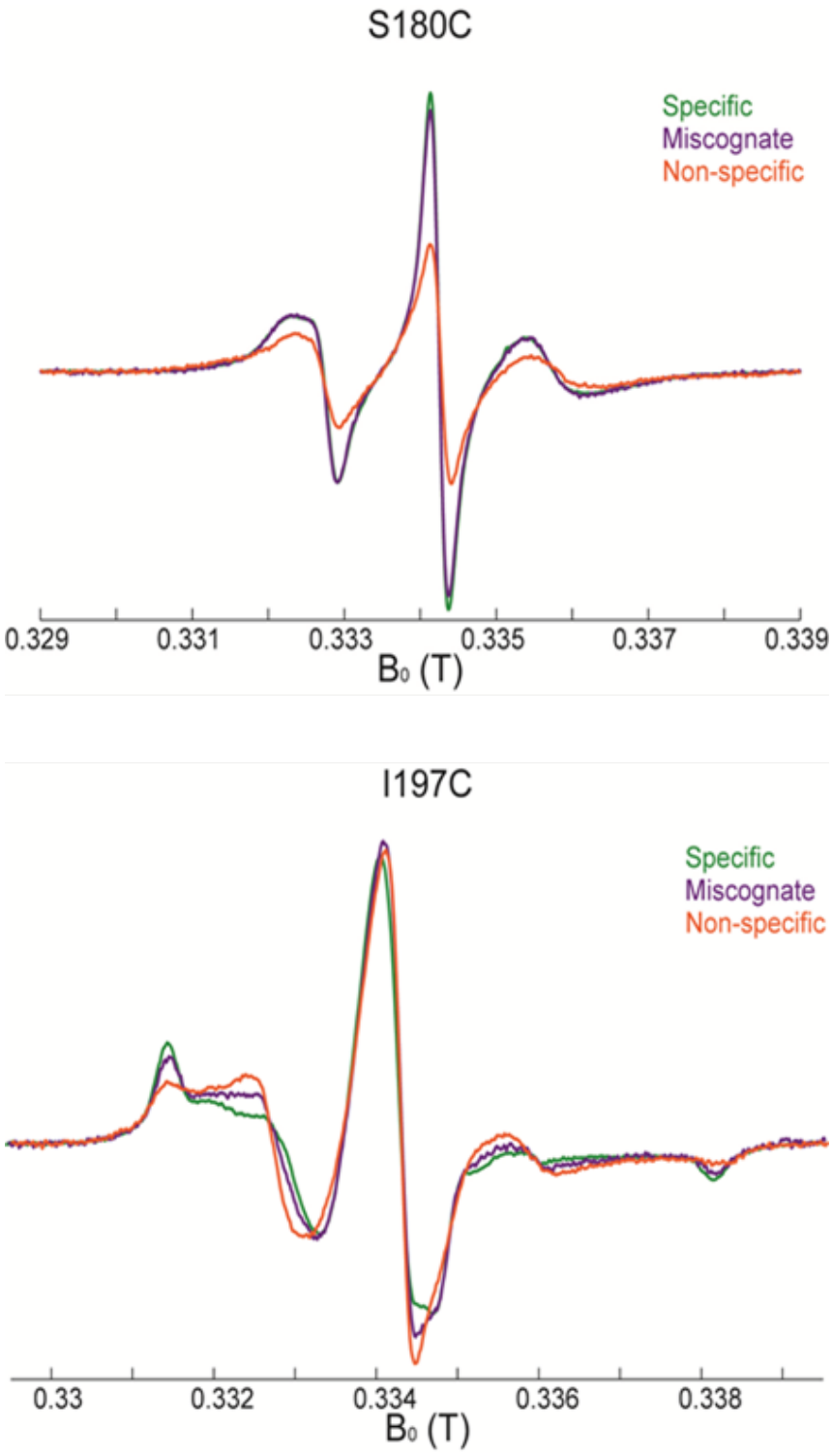


Figure 3.7 Continuous wave spectra for S180C and I197C complexes.

Top panel: X-band CW spectra data for S180C specific (green) miscognate (purple) and nonspecific (orange) complexes. Bottom panel: CW spectra data for I197C specific (green) miscognate (purple) and nonspecific (orange) complexes. Data analysis and figure by Jessica Sarver.

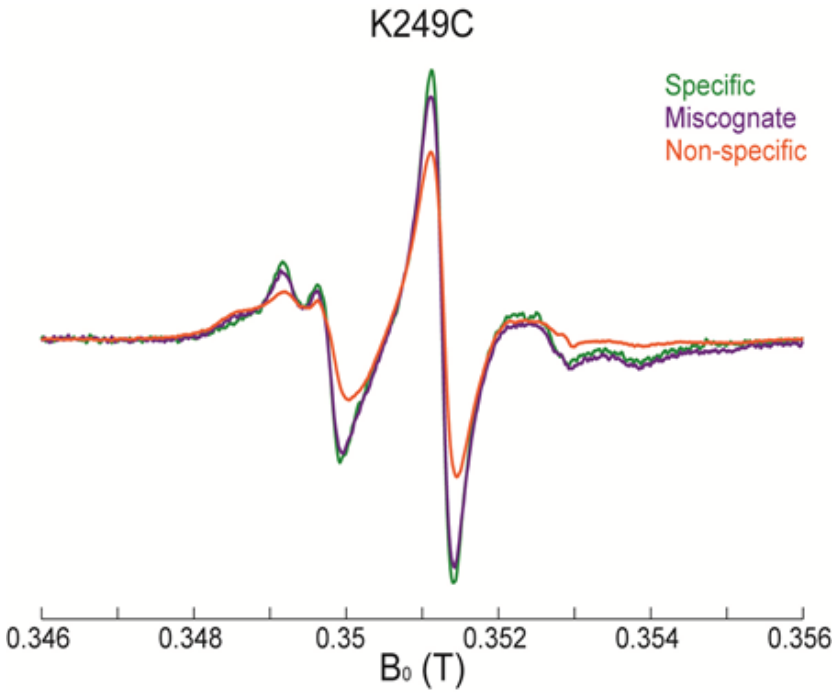


Figure 3.8 X-band continuous wave spectra for K294C complexes.

X-band CW spectra data for K249C specific (green) miscognate (purple) and nonspecific (orange) complexes. Data analysis and figure by Jessica Sarver.

3.4.1 Inverse central line width and the relative mobility of spin labels in EcoRI-DNA complexes

In this section, I describe a semi-quantitative analysis of the CW spectra. As discussed earlier in this chapter, the inverse line width of the central resonance peak (δ^{-1}) is frequently used as a measure of the relative mobility of a spin label. This parameter has been found to be highly correlated with the motion of the local protein backbone fluctuations, especially for small

proteins and within α-helices, and has been established as a measure that reports on the local backbone motions of the protein. [78] This parameter can be used to compare differences in the local backbone motions for the same residue under different conditions (such as bound to different ligands as in these experiments) or to compare differences in the local backbone motions at different locations of the protein. [139], [99] The δ^{-1} values for our X-band CW experiments are shown in the following table:

Table 3.1 Mobility of nitroxide labels in EcoRI complexes.

Residue	Specific	Miscognate	Nonspecific
R123C	0.32 G^{-1}	0.32 G^{-1}	0.27 G^{-1}
R131C	0.48 G^{-1}	0.48 G^{-1}	0.42 G^{-1}
S180C	0.42 G^{-1}	0.42 G^{-1}	0.36 G^{-1}
I197C	0.19 G^{-1}	0.20 G^{-1}	0.26 G^{-1}
K249C	0.34 G^{-1}	0.32 G^{-1}	0.29 G^{-1}

The inverse line width of the central resonance peak (δ^{-1}), measured in $gauss^{-1}$, for the X-band CW spectra. These values are obtained from measuring the width of the central resonance peak. Data analysis by Jessica Sarver.

The values in the above table provide a semiquantiative measure that confirms what can be observed by visual inspection of the lineshapes. The R131C and S180C residues (which reside on a loop and β strand in the arms, respectively) show the greatest apparent mobility, with δ^{-1} values similar to those (~0.4 G^{-1} or higher) reported for loop residues in T4 lysozyme. [86], [99] The K249C reference position (in the main domain) and the R123C location (on an α helix at the "hinge" of the arm) show less mobility, with δ^{-1} values similar to those (~0.3 G^{-1}) reported for T4 lysozyme residues which are located on an α helix (such as K249C and R123C) or

participate in a tertiary contact. The I197C residue shows the least mobility, with δ^{-1} values similar to those (~0.2 G^{-1} or lower) reported for residues which form a highly stable tertiary contact or are deeply buried within the protein. [86] The values we obtained for the different positions are therefore in agreement with the mobility parameters we would expect to see at these locations based on the crystal structure of the specific complex.

The I197C position provides evidence of reduced mobility for the local backbone in the specific complex relative to the nonspecific complex. For the other residues, the spin label experiences comparatively restricted mobility in the nonspecific complex. For R123C, R131C, and S180C, the specific and miscognate complexes show approximately the same degree of mobility, while in the I197C and the K249C positions, the label in the miscognate complex shows intermediate mobility.

3.4.2 The R131C and S180C positions show similarities for the specific and miscognate complexes at all temperatures

In order to further refine our analyses of these complexes, spectra for the R131C and S180C positions were collected at several additional temperatures (243K, 253K, 263K, 273K, 283K), beyond the initial 293K experiment. Additional differences between the spectra can potentially be resolved in this way, because the different contributions to the dynamics measured by the spectra (backbone motions, local interactions, etc.) may show differential temperature dependence. [139]

For all of the complexes examined, the lineshapes reflect the standard behavior of increasing anisotropy as the temperature is decreased. For both the R131C and S180C series of spectra, the specific and miscognate complexes continued to show indistinguishable spectra for

all temperatures, indicating a high degree of similarity in the dynamics at these positions (at the level of sensitivity of the X-band experiment) for the miscognate and specific complexes (Figure 3.10 and 3.11). The R123C and I197C complexes were examined over the same temperature series in solutions at two different viscosities as discussed in the next section (Figure 3.12). These .lineshapes showed the same pattern of increasing anisotropy at decreasing temperature. There are a few additional components that appear in the 243K spectra for all complexes, but this temperature reflects the "rigid limit" spectra which freezes out the hyperfine interactions, as discussed earlier in this chapter.

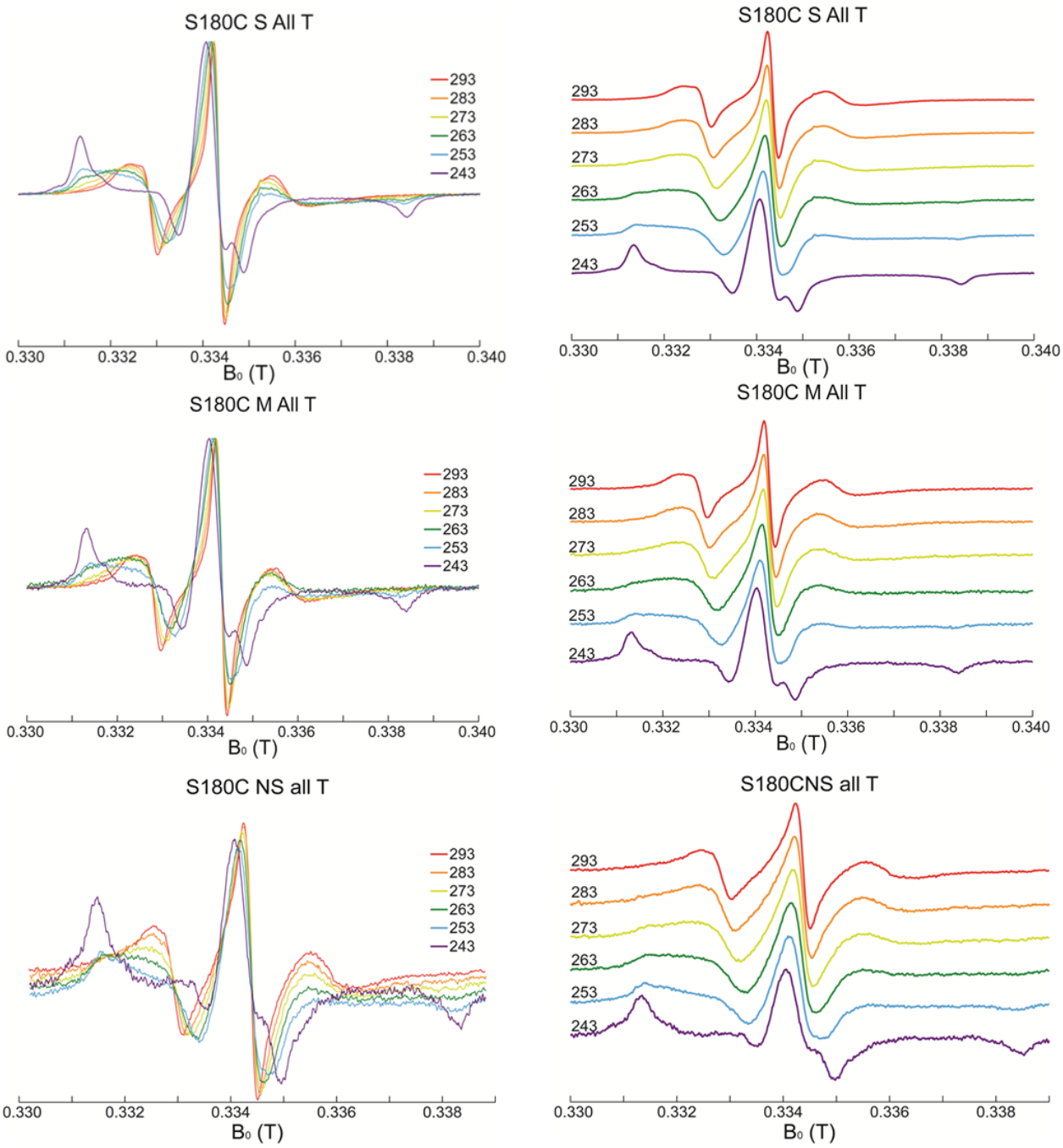


Figure 3.9 S180C CW temperature series.

Continuous wave spectra for S180C complexes collected at 10° temperature increments from 243-293K. Top row: CW spectra data for S180C specific complexes. Middle row: CW spectra for S180C miscognate complexes. Bottom row: CW spectra for S180C nonspecific complexes. Data analysis and figure by Jessica Sarver.

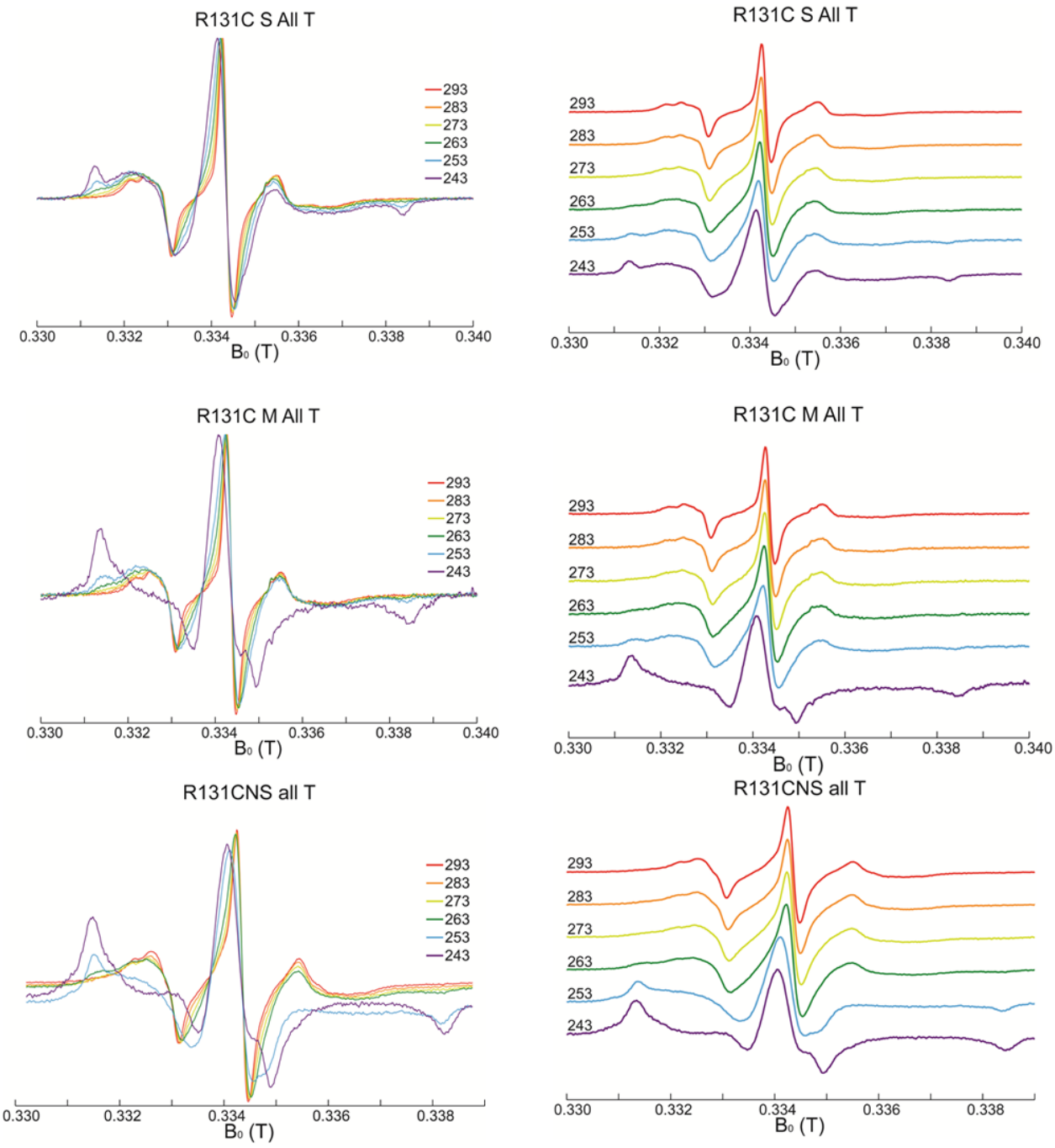


Figure 3.10 R131C CW temperature series.

Continuous wave spectra for R131C complexes collected at 10° temperature increments from 243-293K. Top row: CW spectra data for R131C specific complexes. Middle row: CW spectra for R131C miscognate complexes. Bottom row: CW spectra for R131C nonspecific complexes. Data analysis and figure by Jessica Sarver.

3.4.3 Viscosity studies show that the rotational motion of EcoRI is not a contributing factor to our CW results

As discussed earlier in this chapter, the rotational motion of the protein itself can potentially be a confounding factor to X-band CW lineshapes. By comparing spectra obtained for samples in solution at high viscosity (25% Ficoll 70 or higher), the potential contribution of the rotational motion to the CW spectra can be examined. [107], [159] We obtained the spectra for the R123C and I197C complexes for temperatures (243K, 253K, 263K, 273K, 283K and 293K) under conditions of "low" (10% glycerol) and "high" (30% w/v Ficoll 70) viscosity. For these residues, the low and high viscosity spectra were nearly identical for all temperatures, except at the lowest temperature (243K). This indicates that the rotational motion of the protein is not a significant confounding variable in our studies. The rigid limit spectra represent motion that is too slow to be observable at the given frequency. Since the samples differ in viscosity, the rigid limit will occur at different temperatures, which is likely why the spectra differ at 243K. (Jessica Sarver, personal communication).

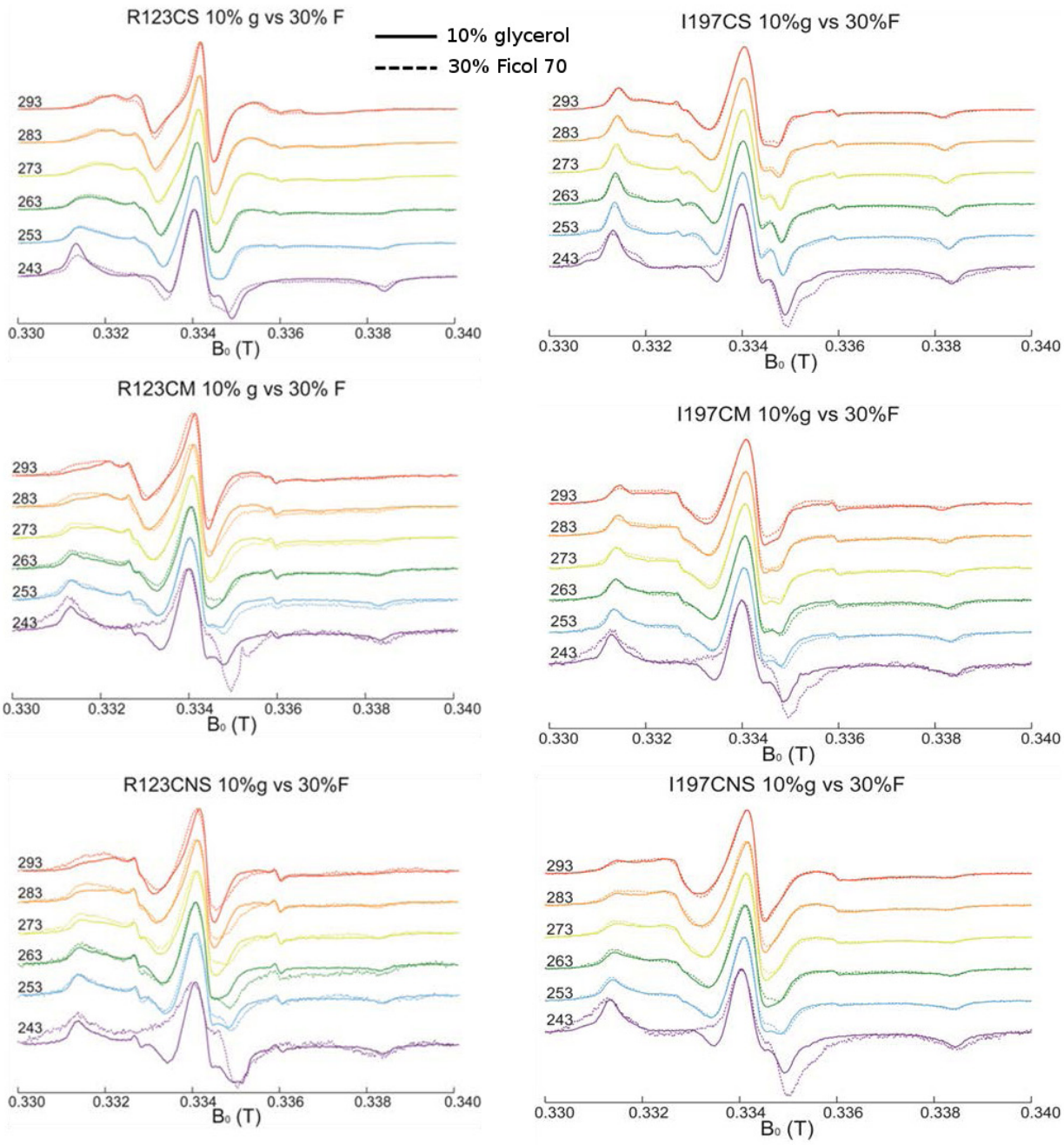


Figure 3.11 Comparison of CW spectra for R123C and I197C complexes collected under different temperature and viscosity conditions.

Continuous wave spectra for R123C and I197C complexes collected at 10° temperature increments from 243-293K for samples prepared in 10% glycerol (dotted lines) and 30% Ficoll 70 (solid lines). Top row: CW spectra data for specific complexes. Middle row: CW spectra for miscognate complexes. Bottom row: CW spectra for nonspecific complexes. Data analysis and figure by Jessica Sarver.

3.4.4 Discussion of X-band spectra

The I197C spectra all show a very high degree of anisotropy, as would be expected for a label on a residue which is in a buried position in the protein-DNA binding site. Anisotropy can originate from orientation restriction or slowing of the spin label motions; however, since as discussed previously this restriction may be at the level of the local backbone motions, or the nitroxide label itself, it is impossible to definitively determine the origins of this restriction without further deconvolution of the spectrum. The I197C complexes exhibit the trend of increased anisotropy for the miscognate and specific complexes. For this position, the specific complex shows a great degree of anisotropy, with reduced anisotropy for the miscognate complex and further reduced anisotropy for the nonspecific complex. As the main-chain carbonyl group of this residue makes a water-mediated contact with the amino group of the cytosine flanking the specific site (Figure 2.4), the fact that the label at this position shows a high degree of anisotropy is in agreement with our expectations. (This interaction would be expected to restrict the motion of the spin label attached to I197C.) In addition to the indirect contact described above, the adjacent residue (G196) forms a contact between the backbone amine group and the flanking phosphate three bases upstream of the recognition sequence. [7] This contact is absent in the miscognate complex. [8] This location would therefore be expected to be a good reporter of changes in motion of residues at the protein-DNA interface upon binding.

Comparison of the X-band CW spectra for the nonspecific, miscognate, and specific complexes for the K249C complexes showed little difference among the complexes. These spectra show the characteristics of a fairly anisotropic spectrum (as per the mobility parameter, also compare to Figure 3.2-A in the discussion of lineshapes). Since the K249 residue was chosen as a “reference” position located on a fairly stable α-helix in the main domain (as

determined from the specific complex and apo crystal structures), the observation that that these spectra are fairly anisotropic is in agreement with our expectations. There also seems to be a slight but noticeable trend of increasing anisotropy for the miscognate and nonspecific complexes.

Compared to the K249C "reference" position, the residues at S180C (on a β strand of the arm) and R131C (on the loop at the outer edge of the arm) show spectra which exhibit less anisotropy than the K249C residue. This is agreement with our predictions for residues within the "arms" of the protein. The crystallographic temperature (B) factors observed for the "arms" indicated that these residues *may* experience greater thermal motion than those in the core domain. While B factors depend on a number of contributions including thermal motion of the atoms, the ensemble of conformations present, and packing arrangements, this observation from the crystal structure is consistent with our CW results. Comparing the different DNA complexes, the CW spectra for each these two positions display similar lineshapes between the specific and miscognate complexes, but greater anisotropy for the nonspecific complex. The spectra for the R123C position (end of α-helix in the arm) exhibit anisotropy which is intermediate between that of the K249C and those of the S180C. Like the K249C complexes, the spectra for R123C show a trend of increasing anisotropy for the miscognate and nonspecific complexes.

Despite our predictions from thermodynamics and our DEER results that the miscognate and nonspecific complexes experience more thermally accessible conformations, the spectra for K249C, R123C, S180C, and R131 indicate an apparent decrease in the mobility of the spin label for the noncognate complexes. However, since all of these residues are in positions far outside of the DNA binding site, they may not report on changes which occur at the DNA-binding interface upon complex formation with the different sequences. Although at first glance it may seem that

these results contradict the DEER results indicating that the arms experience greater conformational freedom, this is not necessarily the case. For example, greater overall conformational freedom of the protein backbone in the arms of the nonspecific complexes may permit spin labels at these residues to experience conformations in which they are more restricted in space. This could lead to an overall restriction of the spin label's mobility. Another possibility is that in "looser" nonspecific complexes there are conformations where the spin label is free to form local interactions which it cannot achieve as frequently in the specific complex, again potentially restricting the mobility of the spin label. These examples illustrate that the many potential contributions to the mobility or restriction of a spin label complicate the interpretation of CW spectra.

3.5 W-BAND EXPERIMENTS PROVIDE A HIGHER RESOLUTION VIEW OF THE DIFFERENCES BETWEEN ECORI-DNA COMPLEXES

As discussed earlier in this chapter, CW experiments conducted at W-band frequencies can provide a "snapshot" of motions occurring at a faster time scale than those of the X-band frequencies. Additionally, CW experiments conducted at W-band frequencies can enable much greater resolution of the components contributing to the signal. This enables the potential deconvolution of contributions to nitroxide mobility, including the orientation of the nitroxide and local interactions. Conducting experiments on the same samples at both X and W band frequencies (and more if possible) can provide a much more detailed picture of the behaviors of the spin label, since contributions to the spectra which are too fast to detect at the X-band frequencies may be observed at W band frequencies, while contributions which are too slow to

be observed at W band frequencies may be observed at X band frequencies. [152] Ideally, spectra may be collected at multiple frequencies to provide a more complete deconvolution of the molecular contributions to the observed spectra. [155], [158], [163] However, such experiments are extremely intensive in terms of data collection and processing and are unfortunately beyond the scope of the current work.

The W-band spectra are shown in Figures 3.13-3.15. Overall, these experiments confirm the results seen for the X-band experiments. The lineshapes for the R131C complexes again reflect the least anisotropy overall, while the S180C complexes are a little more anisotropic. The R123C complex spectra again show greater anisotropy than those for the R131C and S180C positions, with the spectra for the K249C reference position showing even more anisotropy. The I197C spectra again show the greatest degree of anisotropy relative to the other positions. Comparing the spectra for the different complexes at each residue, once again for the I197C complexes, the specific complex shows the greatest anisotropy relative to the miscognate and nonspecific complexes. By contrast, the R123C, R131C, S180C, and K249C positions all appear to show greater anisotropy for the nonspecific complex. These results are all in agreement with our observations for the X-band spectra.

Taken together, our results agree with our expectations in terms of the anisotropy and mobility experienced at the different positions within the complex. The I197C data show the expected trend of increased anisotropy in the specific complex relative to noncognate complexes. However, nitroxide-labeled sidechains in the arms showed trends of increased anisotropy for the nonspecific complex. Since the residues in the arm regions are likely to be mobile in all three classes of proteins, these residues are less likely to report on the constraints imposed by the determinants of specificity (protein-base and protein-phosphate contacts) at the interface. That is,

the CW spectra of these nitroxide-labeled arm residues would be less likely to report on differences between the three classes of complexes.

The CW spectra for these residues report on the *local* mobility of the spin label, which may be influenced by the precise orientation of the spin label in the various complexes. For example, an increase in number of conformations accessible by the arms in nonspecific complexes may permit the nitroxide to form interactions (or to experience steric interference) which restrict motion of the spin label. As the CW results may be convoluted by interactions of the spin label in different conformations, the CW spectra cannot currently be fully interpreted. Further deconvolution of these spectra are ongoing, in order to examine the contributions to the differences between these spectra in a more detailed and quantitative way.

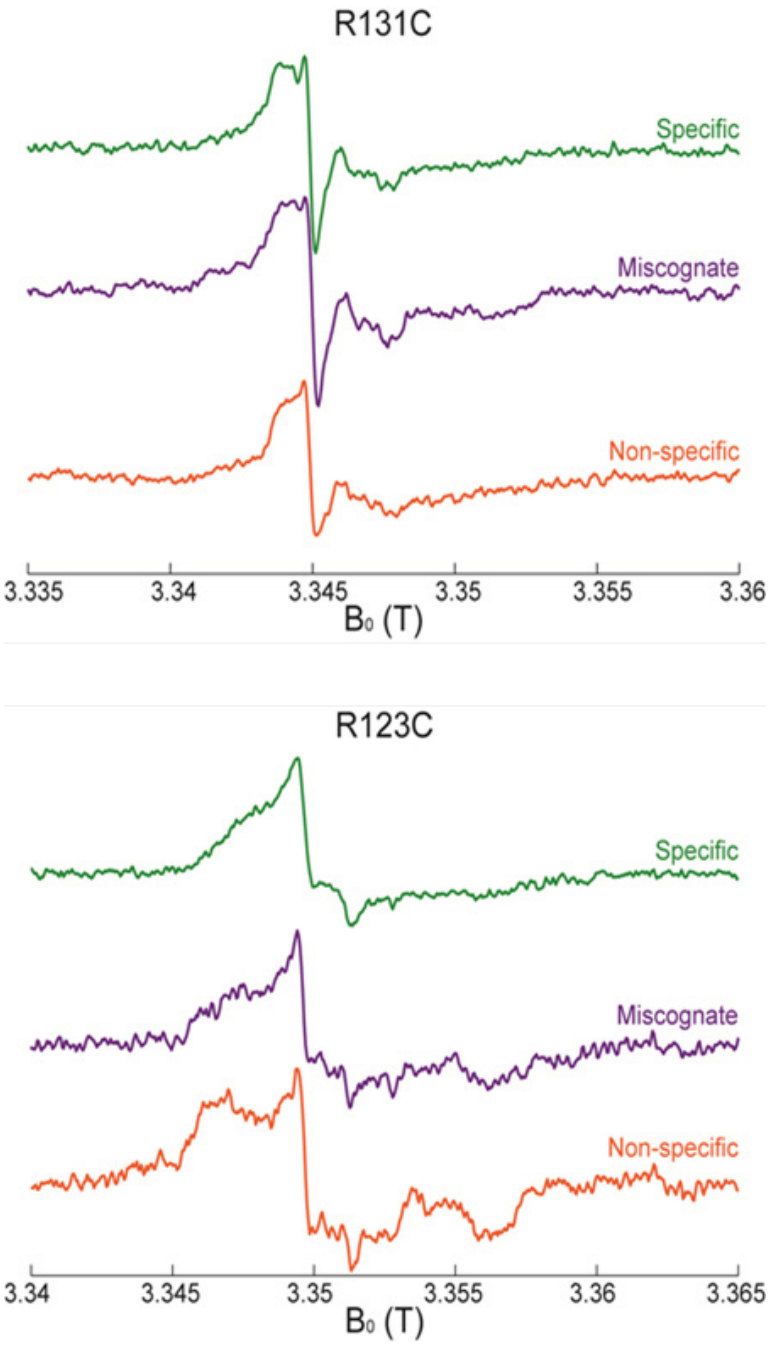


Figure 3.12 W-band continuous wave spectra for R131C and R123C complexes.

Top panel: W-band CW spectra data for R131C specific (green), miscognate (purple), and nonspecific (orange), complexes. Bottom panel: W-band CW spectra data for R123C specific (green) miscognate (purple) and nonspecific (orange) complexes. Data analysis and figure by Jessica Sarver.

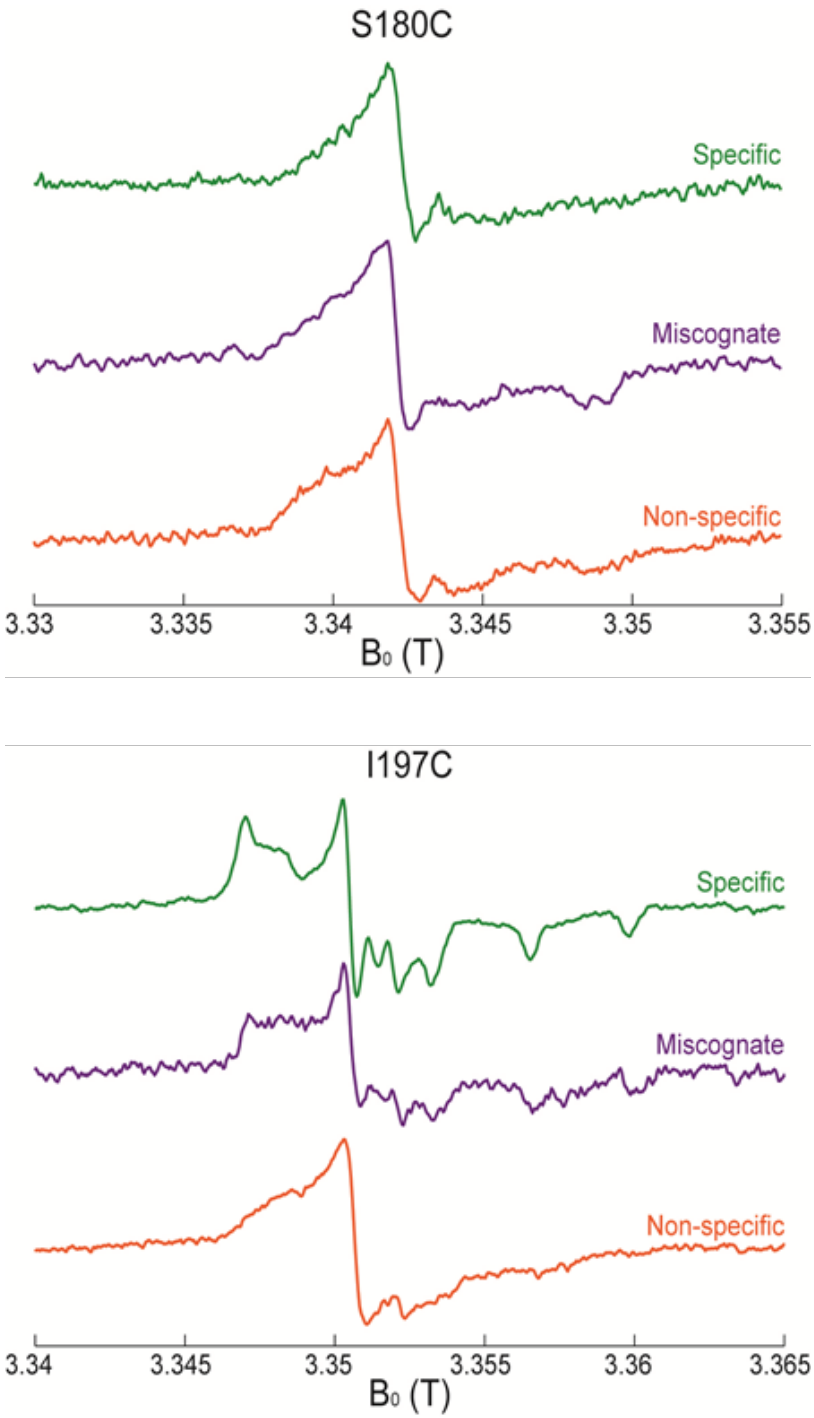


Figure 3.13 W-band continuous wave spectra for S180C and I197C complexes.

Top panel: W-band CW spectra data for S180C specific (green), miscognate (purple), and nonspecific (orange), complexes. Bottom panel: W-band CW spectra data for I197C specific (green) miscognate (purple) and nonspecific (orange) complexes. Data analysis and figure by Jessica Sarver.

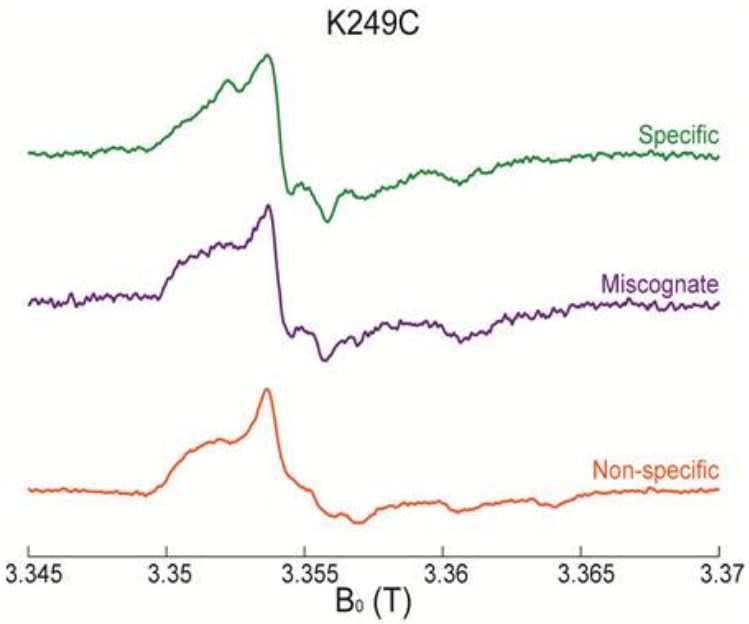


Figure 3.14 W-band continuous wave spectra for K249C complexes.

W-band CW spectra data for K249C specific (green), miscognate (purple), and nonspecific (orange) complexes. Data analysis and figure by Jessica Sarver.

3.6 CONCLUSIONS AND FUTURE DIRECTIONS

In this chapter, I demonstrate the application of CW-ESR to examination of the differences in dynamic behavior of the spin label at specific residues of the protein in the three classes of protein-DNA complex. We show that CW-ESR is capable of detecting differences in the spectra for a protein binding to specific, miscognate, and nonspecific sites. While CW has been used to investigate changes in between bound and unbound states of DNA-binding proteins, this is the first use of CW to investigate the more subtle differences of a protein binding to specific, miscognate, and nonspecific DNA binding sites.

Taken together, the CW results show a trend of dramatically increased anisotropy for the *specific* complex relative to the nonspecific complex for the spin label at the I197C position. In contrast to this, increased anisotropy is seen in the spectra of the *nonspecific* complex for the R123C, R131C, S180C, and K249C positions. These results do not appear to be the result of changes in the global tumbling of the protein, as the same trends are seen in spectra collected at different temperatures, viscosities, and frequencies. However, it is currently not possible to deconvolute the origins of these differences between changes in the behavior of the local protein backbone from changes in the behavior of the spin label.

By conducting computational simulations of the ESR spectrum, it is possible to extract some of the contributions to the observed nitroxide mobility. [153], [167] These simulations are currently being conducted by Jessica Sarver (Saxena lab). Additional future experiments could include the collection of CW spectra at additional frequencies (providing a more thorough investigation of the dynamic behaviors of the spin label at different rates of motion), for additional mutant residues which are adjacent to the DNA binding site, and for EcoRI-specific complexes with different flanking contexts (in order to examine whether flanking context produces effects detectable by CW).

4.0 IDENTIFICATION AND CHARACTERIZATION OF A SECOND METAL BINDING SITE IN ECORI

In this chapter, I briefly review some of what is known about the role of metal ions in restriction enzyme catalysis and discuss investigations into EcoRI structure and metal binding using Cu^{2+} as a probe. We have discovered a previously unknown Cu^{2+} binding site in EcoRI. Herein I discuss the identification and characterization of this binding site and our investigations into the utility of Cu^{2+} as a probe for ESR experiments.

4.1 THE ROLE OF DIVALENT CATIONS IN RESTRICTION ENZYME CATALYSIS

Many enzymes require metal ion cofactors for catalysis. Restriction enzymes are only able to conduct phosphodiester hydrolysis of their DNA substrates in the presence of divalent cations, generally Mg^{2+}. Although the requirement of the divalent cation for cleavage has been well established, the precise nature of the mechanisms driving phosphodiester hydrolysis in restriction enzymes is still under investigation. [168], [169] EcoRI and structurally similar restriction enzymes hydrolyze the phosphodiester bond of the DNA backbone by a S_N2-type nucleophilic substitution mechanism. (As reviewed in [3].) The general mechanism comprises three steps: the deprotonation of the attacking nucleophile, the nucleophilic attack on the phosphorus resulting in

a pentavalent transition state, and the departure of the 3' hydroxyl leaving group. Each of these steps requires an assisting group: the first step requires a basic group to deprotonate the nucleophile, the second requires a Lewis acid to stabilize the transition state, and the third requires an acidic group to protonate the leaving group. While it has long been known that divalent cations are required for type II restriction enzymes to carry out catalysis, the number of divalent cations required for this process for many restriction enzymes is a matter of debate. Mechanisms with one, two, or three metal ions have been proposed. [3] Mechanisms in which one metal is required for catalysis, and a second metal has a modulatory role have also been proposed. [169]

4.1.1 EcoRI metal binding and DNA catalysis

EcoRI cleaves DNA in the presence of the catalytic cofactor Mg^{2+}, but several other divalent cations have been demonstrated to support cleavage at reduced rates, with Mg^{2+} Mn^{2+}>Co^{2+}>>Zn^{2+}>>Cd^{2+}>Ni^{2+}. [170] Sapienza (former lab member) has shown that similar to BamHI, EcoRI cleavage activity in the presence of Mg^{2+} is strongly inhibited by Ca^{2+}, indicating that Ca^{2+} directly competes with Mg^{2+} for binding. [6], [55] The catalytic role of magnesium in EcoRI has been investigated with a combination of thermodynamic studies and molecular dynamics (MD) simulations. [20] The post-reactive complex for EcoRI with manganese has been determined (PDBID: 1QPS) [19]; and in this complex, a Mn^{2+} ion is observed to be coordinated in the active site.

Molecular dynamics studies were performed on the highly refined version of the specific complex, [47] using the location from 1QPS to position the Mg^{2+} ion. In these molecular dynamics simulations the Mg^{2+} is coordinated by six ligands (Figure 4.1). They are the

carboxylate oxygen of E111, two carboxylate oxygens of D91, the backbone carbonyl of A112, the O1P of the scissile phosphate, and a water molecule that may form the attacking nucleophile (W_A). In the presence of Mg^{2+}, (Figure 4.1) this water rotates such that it is oriented for a nucleophilic attack on the scissile phosphate.

Based on the results from the MD simulations, the O2P oxygen of the GApATTC phosphate orients a water (W_C). This water assists in positioning another water (W_A) such that either the Mg^{2+}-bound water itself acts as the nucleophile, or the Mg^{2+} serves as a Lewis acid to promote the dissociation of water to generate a hydroxyl to form the attacking nucleophile. (Figure 4.1-D, [20]) EcoRI therefore uses a single metal mechanism for catalysis. Several other restriction enzymes use a second metal ion to polarize the scissile phosphate, but in EcoRI this role is played by the R145 residue in the active site, which along with K113 coordinates the O2P oxygen of this phosphate. [20], [171]

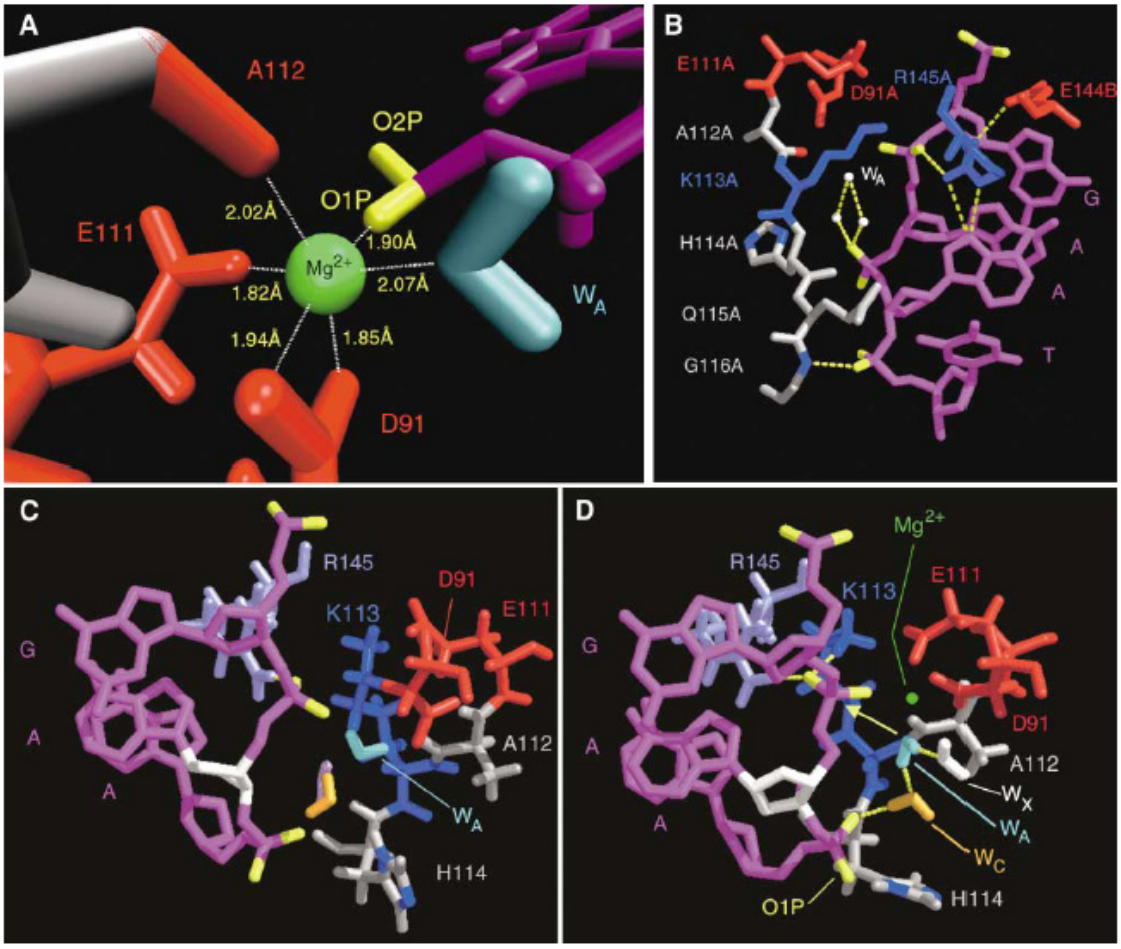


Figure 4.1 Catalytic center of EcoRI-DNA complex.

A) Coordination of Mg^{2+} in the active site. Snapshot taken from an MD simulation. B) Catalytic and recognition elements in the crystal structure of the specific complex. [47] **the A and B designations refer to the A and B chains of the dimer. C) Snapshot taken from an MD simulation of the specific EcoRI-DNA complex in the absence of Mg^{2+}. D) Snapshot taken from an MD simulation of the specific EcoRI-DNA complex in the presence of Mg^{2+}. Figure taken from Kurpiewski 2004 [20].**

4.2 ESR REVEALS THAT CU^{2+} COORDINATES TO ECORI

4.2.1 Use of paramagnetic metals in Electron Spin Resonance experiments

All of the ESR measurements discussed in Chapters 2 and 3 employed the use of nitroxide spin labels attached to cysteine residues of the protein to provide the stable electron spin for ESR detection. Another method of collecting ESR data for biological molecules is through the use of paramagnetic metals that interact with a protein. Paramagnetic ions that are stably coordinated in a metal binding center act as a "spin label" of that location for ESR studies. A great advantage of this approach is the fact that for proteins which already contain a metal-binding site that coordinates a paramagnetic metal, this site can be "labeled" without needing to generate a cysteine mutation at that location. Another advantage of conducting ESR on metal ions is that unlike nitroxide spin labels, there is no "linker" or motion of the spin label itself to convolute the interpretation of the data. The coordinated metal ion reports on a very specific location of the protein.

If more than one metal-binding site is present, paramagnetic metals can also be used in DEER experiments to determine distances between the metals. [113], [172-174] Since the metals are very precisely coordinated and do not have a "linker" contributing to their motion, much narrower distance distributions can be obtained for metal-based DEER experiments. This technique can also be used in combination with site-directed-spin-labeling, such that metal-metal, metal-nitroxide, and nitroxide-nitroxide positions can all potentially be obtained. [175] This permits the position of a nitroxide-labeled residue to be "triangulated" by determining multiple distances, or for a metal-binding site to be identified by triangulation to nitroxide labels.

4.2.2 Copper as a probe for ESR experiments

Among biologically relevant metals, Cu^{2+} is commonly used for DEER measurements. While Mg^{2+}, Zn^{2+}, Ca^{2+} are more commonly found as catalytic cofactors which are coordinated in protein metal binding sites, they are not paramagnetic and therefore cannot be used for ESR. While Mn^{2+} and Fe^{3+} are paramagnetic, they can have as many as five unpaired electrons, which greatly complicates data analysis. Since Cu^{2+} has only one unpaired electron, it is simpler to extract data by using copper as a paramagnetic label for a protein.

Copper is an essential trace element which is associated with a number of enzymes. Enzymes which require copper as a cofactor are generally involved in the carrying or transport of electrons, or in the transport and activation of oxygen. (Reviewed in [176], [177].) However, free copper within the cell is extremely toxic. This toxicity is at least partly because of its redox potential-free copper in the cell can lead to the generation of reactive oxygen species. [177] Cellular concentrations of free copper are therefore extremely low and very tightly controlled, and copper in the cell is generally sequestered by binding to storage proteins. [177]

Divalent copper generally binds to imidazole groups, ie: histidine residues in proteins. [178] Since there are no histidine residues in the Mg^{2+} binding site of EcoRI, we did not expect the Cu^{2+} to be coordinated in the same location. However, since there are several histidine residues in EcoRI, we hoped that at least one of them would be able to coordinate copper, and that we would be able to use Cu^{2+} as an additional probe for investigation of the structural properties of the EcoRI enzyme. Divalent copper has been used in a number of ESR studies of proteins, including serum albumins, [179] multicopper oxidase PcoA, [180] azurin [172] and prion protein. [181-183]

4.2.3 Sample preparation and data collection methods for Cu^{2+}- ESR experiments

Concentrated stocks of wild-type EcoRI-specific complex were prepared by lab member Preeti Mehta in a deuterated buffer of 30mM N-ethylmorpholine, 0.3M NH_4Cl, pH 8 (Cu^{2+}-ESR sample buffer, see Appendix). To this was added a 2:1 molar ratio (two moles of copper salt to each mole of wild-type EcoRI dimer) of isotopically enriched $^{63}CuCl^2$ (Cambridge Isotopes). The EcoRI S180C specific complex was spin labeled as described in Chapter 2 and the Methods section and exchanged into the same Cu^{2+}-ESR sample buffer described above. To this was also added a 2:1 ratio of $^{63}CuCl^2$ to protein-DNA complex. ESR experiments were performed on a Bruker Elexys 580 spectrometer. CW and ESEEM experiments were performed in with a MS3 resonator, at 80K while the DEER experiments were performed with a MD5 resonator at 20K. ESR data presented in this chapter were collected and analyzed by collaborators Zhongyu Yang and Ming Ji (Saxena lab). A more detailed explanation of the data collection and analysis procedures are presented as part of the Ph.D. dissertation of Zhongyu Yang, Department of Chemistry, University of Pittsburgh. [184]

4.2.4 ESEEM reveals that Cu^{2+} is coordinated to a histidine residue

Electron Spin Echo Envelop Modulation (ESEEM) is an ESR technique that determines the coupling between the electron spin being probed and local nuclei. ESEEM has been used for identification of the coordination state of copper in proteins such as NAD glycohydrolase. [185] The three-pulse ESEEM spectrum was collected for the wild-type specific complex with Cu^{2+} at 3369G, shown in Figure 4.2-B. [184] The peaks seen at 0.6 MHz, 1.0 MHz, and 1.6 MHz (indicated in Figure 4.2-B in red) result from the interaction of Cu^{2+} with a remote ^{14}N of an

imidazole ring and are characteristic of histidine coordination. [180], [182], [186], [187] The peak at 14.8MHz (blue) was assigned to the hyperfine interaction between Cu^{2+} and remote protons, which may be from either the solvent or the protein. The ESEEM results therefore indicate that the Cu^{2+} is coordinated by one of the histidine residues of the protein. EcoRI contains five histidine residues in each subunit, at positions 31, 114, 147, 162, 225, shown in Figure 4.3. From this information, we designed further DEER experiments to determine the identity of the Cu^{2+} - coordinating histidine residue.

4.2.1 Continuous wave spectrum of the wild-type specific Cu^{2+} complex

Figure 4.2-A shows the continuous wave spectrum for the wild-type specific complex with Cu^{2+}. (Features of CW spectra are discussed in Chapter 3.) This CW spectrum shows two separate components (indicated with vertical lines in the figure) which are present at an approximately 1:1 ratio. The parameters of these components were deconvoluted by simulation of the spectra. The first component of the CW spectrum has parameters ($g_{\|}$ = 2.289, $A_{\|}$ = 163G) which are characteristic of either Cu^{2+}-2N2O or Cu^{2+}-3N1O binding. The $A_{\|}$ values are consistent with those for type-II Cu^{2+} complexes, which is a coordination geometry in which the copper has four equatorial ligands and two axial ligands. [188] The second component has parameters ($g_{\|}$ = 2.228, $A_{\|}$ = 143G) which do not match any established binding mode and may be the result of an irregular coordination symmetry. [184], [188] These results support the observations from the ESEEM experiment which indicate that the Cu^{2+} is coordinated to a histidine.

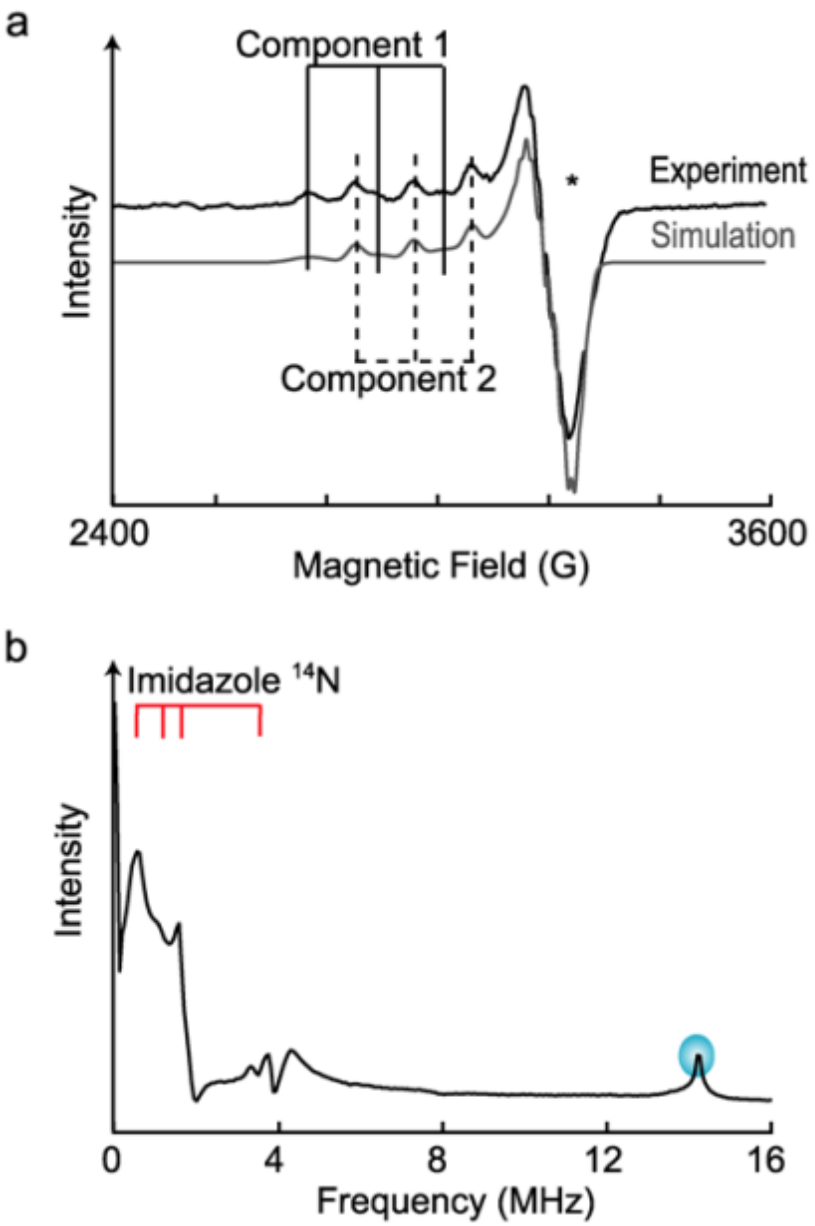


Figure 4.2 CW and ESEEM spectrum of EcoRI specific complex with Cu^{2+}.

Top panel: CW-ESR spectrum of Cu^{2+} bound to EcoRI-specific complex. There are two distinct components to this CW spectrum, indicated as "Component 1" and "Component 2" by the red vertical bars. These two components make roughly equal contributions to the spectrum. Bottom panel: Three-pulse ESEEM spectrum of Cu^{2+} bound to the EcoRI-specific complex. The imidazole peaks are indicated with red bars. The blue circle indicates the proton ESEEM peak. Data collection and analysis by Zhongyu Yang.

4.3 LOCATING THE COPPER BINDING SITE IN ECORI

There are five histidine residues in EcoRI, shown as colored spheres in Figure 4.3. Table 4.1 below shows the distances between these histidine residues when Cu^{2+} is coordinated to the Nε or Nδ. In order to definitively identify the Cu^{2+} binding site, DEER experiments as described in the following sections were used to triangulate the location of the Cu^{2+}.

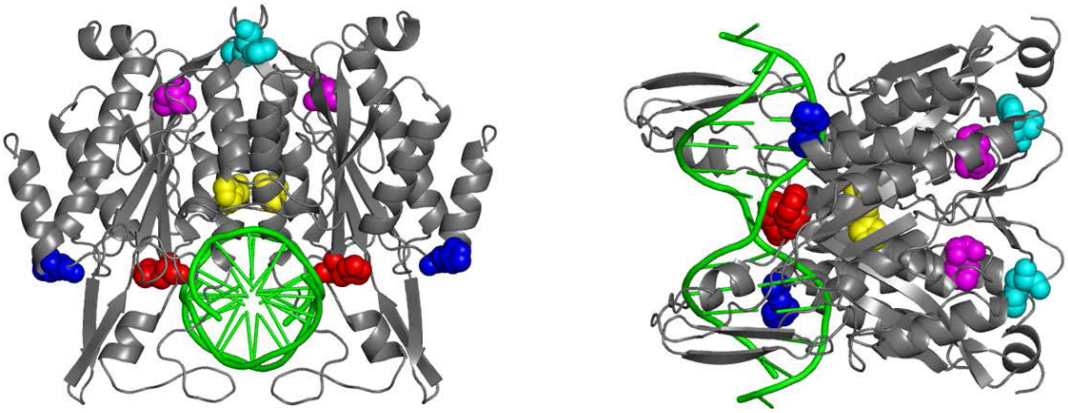


Figure 4.3 Histidine Residues in EcoRI.

The specific EcoRI-DNA complex is shown with DNA in green. Histidine residues are shown as colored spheres. H31 is shown in blue, H114 in red, H147 in yellow, H162 in magenta, and H225 in cyan.

Table 4.1 Distances between histidine residues in EcoRI.

	H31δ	H31ε	H114δ	H114ε	H147δ	H147ε	H162δ	H162ε	H225δ	H225ε
H31δ	67 Å	68 Å	20 Å	19 Å	34 Å	34 Å	32 Å	30 Å	42 Å	41Å
H31ε	68 Å	70 Å	21 Å	20 Å	35 Å	35 Å	32 Å	30 Å	42 Å	41 Å
H114δ	47 Å	49 Å	28 Å	31 Å	17 Å	16 Å	27 Å	26 Å	35 Å	35 Å
H114ε	50 Å	51 Å	31 Å	33 Å	19 Å	18 Å	28 Å	26 Å	36 Å	36 Å
H147δ	37 Å	38 Å	21 Å	22 Å	8 Å	9 Å	20 Å	20 Å	24 Å	25 Å
H147ε	37 Å	38 Å	20 Å	22 Å	9 Å	10 Å	22 Å	22 Å	26 Å	27 Å
H162δ	55 Å	55 Å	39 Å	40 Å	22 Å	22 Å	30 Å	31 Å	12 Å	11 Å
H162ε	55 Å	55 Å	38 Å	40 Å	21 Å	22 Å	31 Å	31 Å	13 Å	12 Å
H225δ	52 Å	52 Å	40 Å	39 Å	22 Å	23 Å	26 Å	28 Å	26 Å	27 Å
H225ε	54 Å	54 Å	39 Å	40 Å	23 Å	24 Å	28 Å	29 Å	27 Å	28 Å

Inter-subunit distances are shown in shaded boxes.

4.3.1 Challenges of Cu^{2+} - based DEER experiments

DEER experiments can be used to determine distances between paramagnetic groups in proteins as discussed in Chapter 2. Paramagnetic metals such as Cu^{2+} therefore can be used for determination of DEER-based distance measurements between Cu^{2+} ions and between Cu^{2+} and nitroxides. [172], [175], [189] These experiments are far from trivial. Due to the comparatively broad absorption of the Cu^{2+} spectrum, the selective pulses of the DEER experiment only excite a small portion of this spectrum. This causes orientation effects in the signal that must be accounted for in order to determine accurate distance measures. Recent experiments by Saxena and coworkers have determined ways to optimize experimental conditions for collecting high quality Cu^{2+}-DEER data . They have demonstrated that by collecting DEER data at multiple field strengths and resonance offsets, accurate distance information may be extracted. [173], [174] These experiments used model peptides in order to demonstrate the validity of these new methods; one goal of our collaborative efforts with the Saxena group is to apply recently developed techniques such as these to experimental questions for more complicated macromolecules as described in this chapter.

4.3.2 DEER Cu^{2+}- Cu^{2+} distance measures in the wild-type EcoRI-specific complex

DEER was used to determine point-to-point distances between Cu^{2+} bound to each of the two subunits of wild-type EcoRI-DNA complex. (Figure 4.3). In order to determine the distances accurately, data were collected at four different field strengths and fit by the molecular modeling procedure described in the doctoral dissertation of Zhongyu Yang (Saxena lab). [184] We obtained a most probable Cu^{2+}-Cu^{2+} distance of 35Å, with a standard deviation of ~1 Å. This is

quite narrow compared to the breadth of distributions obtained for our nitroxide-nitroxide distances (~5 Å for R131 complexes). This is because a coordinated metal is highly restricted in its mobility, as discussed earlier in this chapter. Nitroxide-nitroxide distance distributions, by contrast, include contributions from the spin-label, side chain and backbone mobility.

When the 35 Å distance was compared with the values in Table 4.1, this narrowed down the potential identities of the coordinating histidines. The only residue pairs which could produce distances close to 35Å are the H114-H114, H162-H162, or H225-H225 inter subunit distances, or the H31-H147, H114-H225, or H225-H162 intra-subunit distances. In order to distinguish between these potential sites, we then conducted DEER analysis of Cu^{2+} coordinated to the S180C-specific complex.

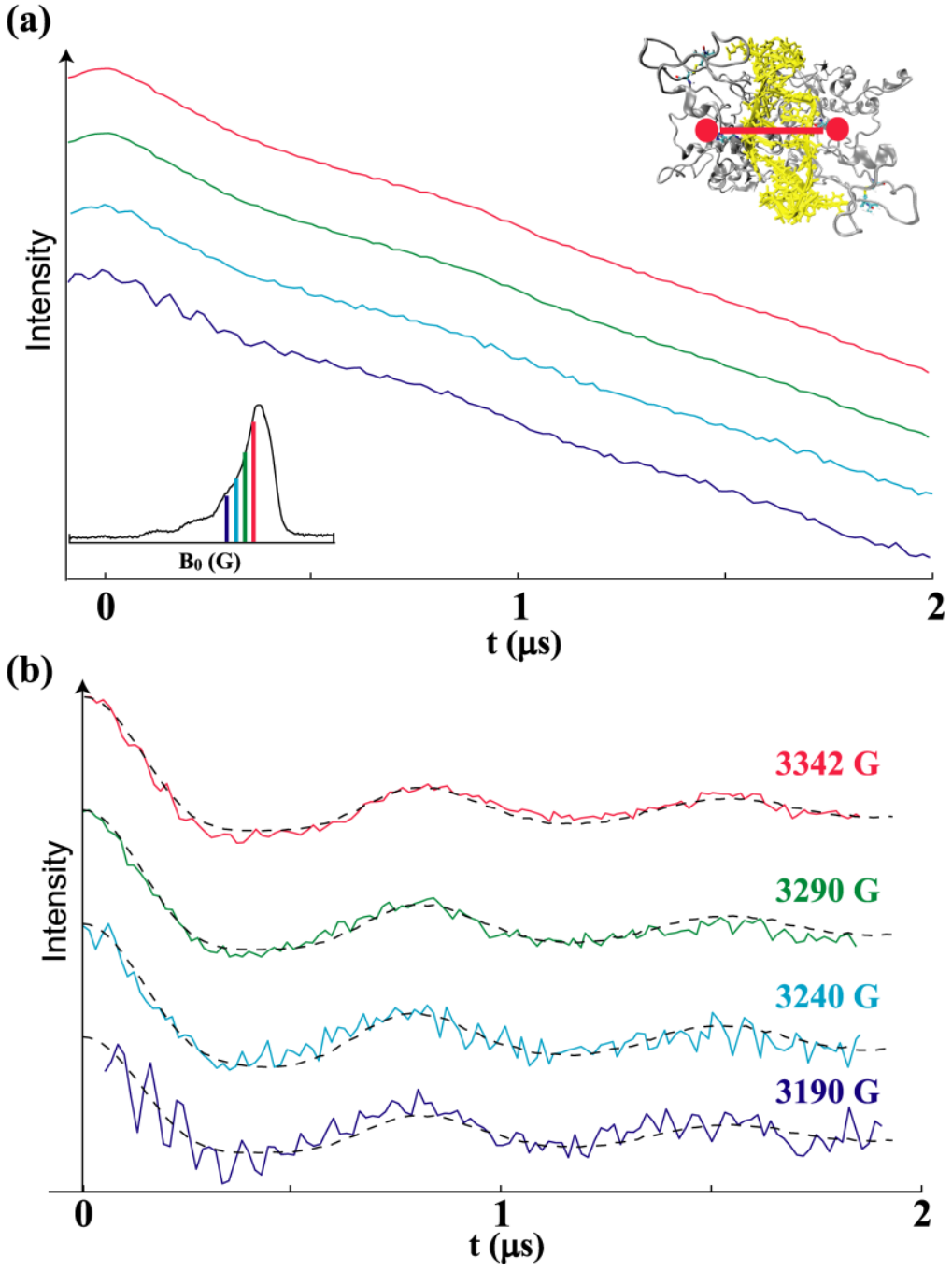


Figure 4.4: WT Cu^{2+}-Cu^{2+} DEER distances.

A) The unprocessed DEER data collected at 20K for four different magnetic fields. The relative position of these fields are color coded on the field-swept echo detected spectrum (left inset) and an illustration of the measured distance is shown (right inset). B) Baseline corrected DEER signal for the four magnetic fields. The dashed lines represent the fit of optimized parameters. Data analysis and figure from [183].

4.3.3 DEER Cu^{2+}- Cu^{2+} distance measures in the S180C EcoRI-specific complex

As discussed in Chapter 2, the S180C mutant of EcoRI can be labeled with MTSSL spin label in order to collect nitroxide based distances, with minimal perturbation to the protein structure. Nitroxide-labeled S180C in complex with specific DNA was prepared as discussed in Chapter 2, except this sample was prepared in a deuterated buffer of 30mM N-ethylmorpholine, 0.3M NH_4Cl, pH 8 (Cu^{2+}- ESR sample buffer) to better promote copper solubility. To this was added a 2:1 molar ratio of $^{63}CuCl$ to protein dimer. The DEER experiments were designed to selectively measure the distance between the S180C-nitroxide and the Cu^{2+}. Data were collected at frequency offsets of 100MHz, 226MHz, 408MHz, and 548MHz; and distance distributions were obtained as described previously. [184] These experiments yielded a bimodal distance distribution, with most probable distances of 22±2Å and 42±3Å (Figure 4.6).

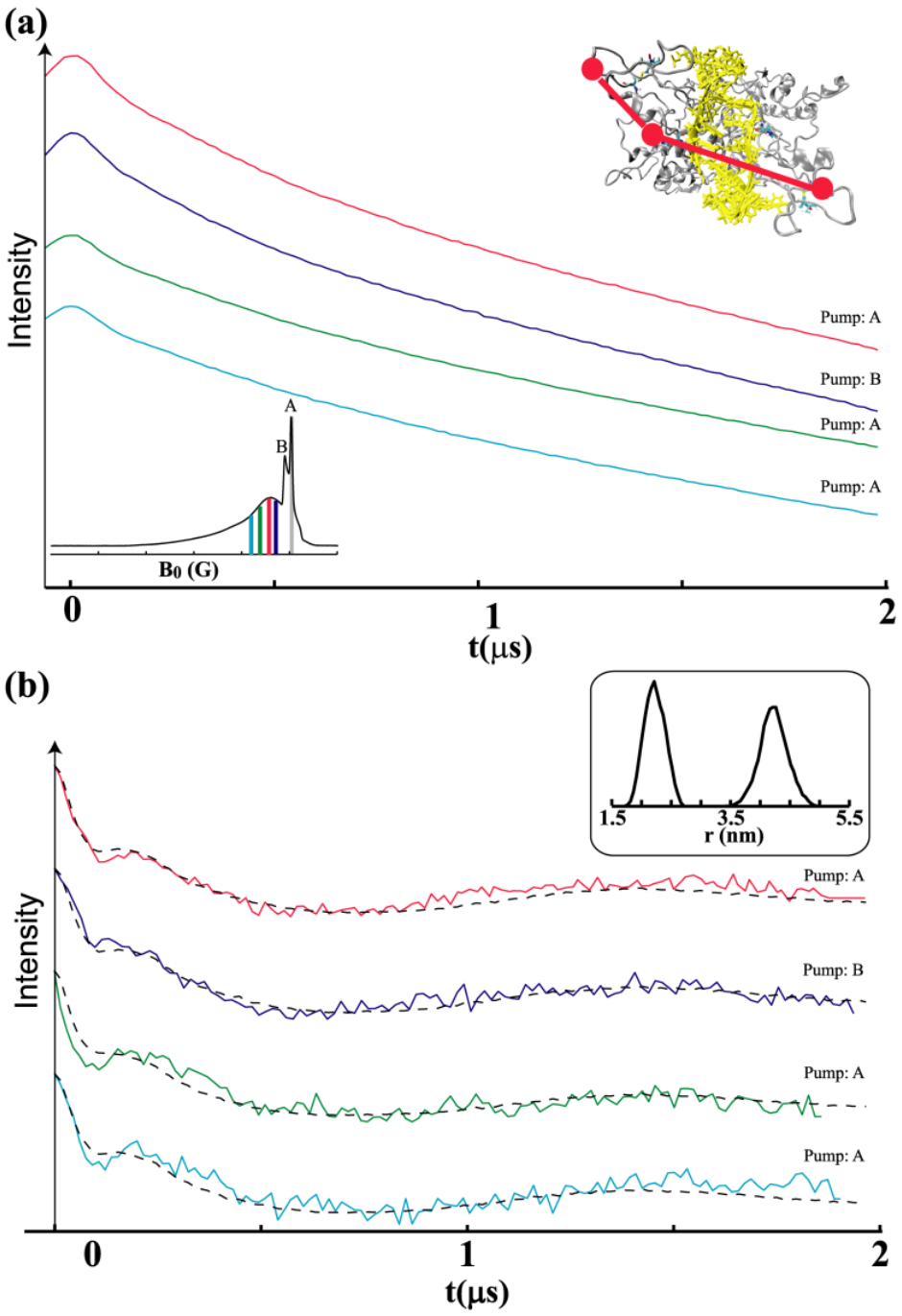


Figure 4.5 DEER data of S180C-EcoRI Cu^{2+}.

Top panel: unprocessed DEER data for Cu^{2+} bound to the S180C EcoRI specific complex collected at magnetic field strengths of 3342G, 3290G, 3190G, and 3090G. The relative positions of these field strengths are color-coded in the spectrum shown in the left inset. Bottom panel: Baseline corrected DEER signal. Data analysis and figure by Zhongyu Yang. [184]

4.3.4 DEER results show that H114 is the residue that coordinates Cu^{2+} to EcoRI

The wild-type Cu^{2+}-Cu^{2+} distance of 35Å could only have arisen if Cu^{2+} was coordinated to the histidine residues shown in Table 4.2. The Cu^{2+}-nitroxide distances that would be observed for these potential binding sites are shown in Table 4.2. Of these possibilities, the observed S180C-Cu^{2+} DEER distances of 22Å and 42Å are only consistent with coordination to the H114 binding site. Combining the results from ESEEM and the two DEER experiments, the Cu^{2+} binding site is identified as H114 (Figure 4.7). More specifically, the Cu^{2+} to Cu^{2+} distances are most compatible with coordination at the Nε of H114. Molecular dynamics experiments are currently being performed by collaborator Ming Ji (Saxena lab) to further refine the details of copper coordination in this metal binding site.

Table 4.2 S180Cβ-Histidine distances.

	Intra-subunit distances	Inter-subunit distances
H114δ-S180C	**23 Å**	**41 Å**
H114ε-S180C	**22 Å**	**43 Å**
H162δ-S180C	43 Å	48 Å
H162ε-S180C	42 Å	48 Å
H225δ-S180C	52 Å	45 Å
H225ε-S180C	52 Å	46 Å
H31δ-S180C, H147δ-S180C	33 Å, 31 Å	55 Å, 38 Å
H31δ-S180C, H147ε-S180C	33 Å, 30 Å	55 Å, 38 Å
H31ε-S180C, H147δ-S180C	34 Å, 31 Å	56 Å, 38 Å
H31ε-S180C, H147ε-S180C	34 Å, 30 Å	56 Å, 38 Å
H114δ -S180C, H225δ-S180C	23 Å, 52 Å	41 Å, 45 Å
H114δ-S180C, H225ε-S180C	23 Å, 52 Å	41 Å, 46 Å
H114ε-S180C, H225δ-S180C	22 Å, 52 Å	43 Å, 45 Å
H114ε-S180C, H225ε-S180C	22 Å, 52 Å	43 Å, 46 Å
H162δ-S180C, H225δ-S180C	42 Å, 52 Å	48 Å, 52 Å
H162δ-S180C, H225ε-S180C	42 Å, 52 Å	48 Å, 52 Å
H162ε -S180C, H225δ-S180C	42 Å, 52 Å	48 Å, 52 Å
H162ε-S180C, H225ε-S180C	42 Å, 52 Å	48 Å, 52 Å

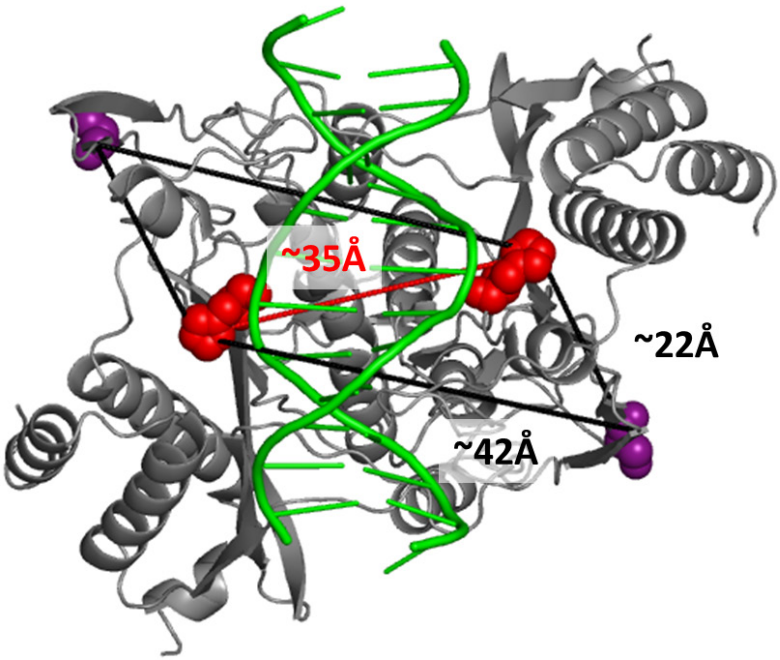


Figure 4.6 Identification of the EcoRI copper binding site by triangulation.

EcoRI-specific complex crystal structure with S180C shown in purple and H114 shown in red.

4.3.5 Supporting evidence that Cu^{2+} is coordinated to the H114 residue of EcoRI

Upon discovering that EcoRI does in fact coordinate Cu^{2+}, we wished to investigate further the properties of the previously uncharacterized EcoRI- Cu^{2+} binding interaction. Mike Kurpiewski (Jen-Jacobson lab member) initially determined the binding affinity of Cu^{2+} to EcoRI. The binding of Cu^{2+} to EcoRI results in an enhancement of EcoRI binding to DNA. This is likely because the introduction of two positive charges from Cu^{2+} at H114 in each half-site partially mitigates the electrostatic repulsion between the DNA phosphate backbone and the negatively charged carboxylate side chains in the active site. The enhancement of binding is only eight-fold, however, in contrast to the 300-fold binding enhancement observed for divalent cations such as Ca^{2+} which directly coordinate with the active site charge cluster and compete with Mg^{2+} binding. This is because H114 is proximal to, but not a part of, the active site cluster. (Reviewed in [23].)

Since Cu^{2+} enhances the binding of EcoRI to DNA, the binding of Cu^{2+} to EcoRI was initially measured by determining the enhancement of the affinity constant of EcoRI for the specific DNA sequence (K_A/K_{A0}) at increasing Cu^{2+} concentrations (Figure 4.6). By this method, Kurpiewski (Jen-Jacobson lab, unpublished data) demonstrated that Cu^{2+} binds to EcoRI with an apparent K_d ~5 μM. (Figure 4.6) In agreement with our predictions from the DEER experiments that the Cu^{2+} binds to H114, Kurpiewski found that Cu^{2+} binds to the H114Y mutant with ~1600-fold lower ($K_{d,\ apparent}$ ~8mM) than it does to the wild-type enzyme (Figure 4.6) He also found that Cu^{2+} does not support cleavage of the DNA substrate as would be expected, since it does not bind to the negatively charged cluster of acidic residues at the active site. Cu^{2+} coordination enhances binding to DNA, but completely inhibits Mg^{2+}-catalyzed cleavage. Even at saturating

concentrations of Mg^{2+} (8mM), it was found that Cu^{2+} has an inhibitory effect on cleavage, indicating that Cu^{2+} inhibits the catalytic rate of EcoRI rather than just competing with Mg^{2+} binding (Figure 4.7). Importantly, whereas 100 μM Cu^{2+} completely inhibits wild type EcoRI cleavage, Cu^{2+} does not inhibit the Mg^{2+}-catalyzed H114Y cleavage reaction even at 200 μM Cu^{2+}. Taken together, these results strongly support our ESR results, indicating that Cu^{2+} is coordinated to H114.

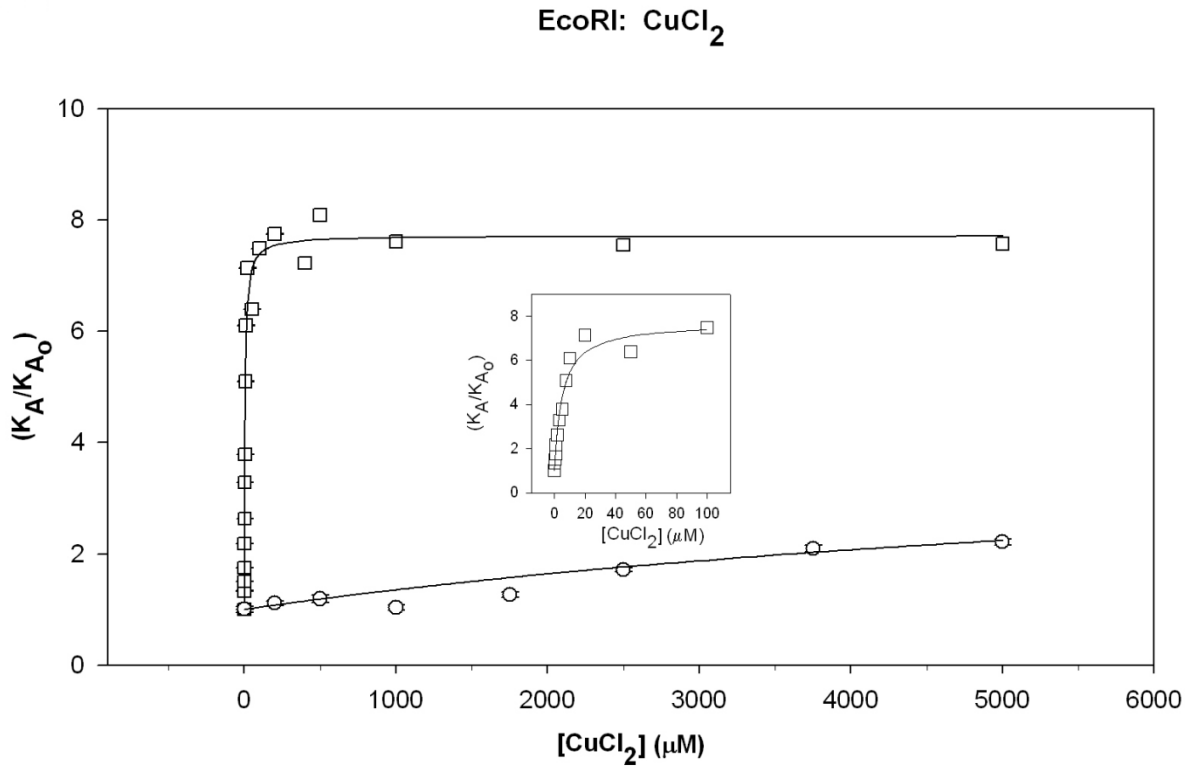


Figure 4.7 Cu^{2+}Binding to wild-type and H114Y EcoRI.

The binding of Cu^{2+} to wild-type (squares) and H114Y (circles) EcoRI is shown. Binding of Cu^{2+} was measured by determining the enhancement of the affinity constant of EcoRI for the specific DNA sequence (K_A/K_{A0}) at increasing Cu^{2+} concentrations. The inset shows a zoomed-in view of the enhancement of formation of wild-type EcoRI-DNA complexes by concentrations of Cu^{2+} from 0 to 100 μM.

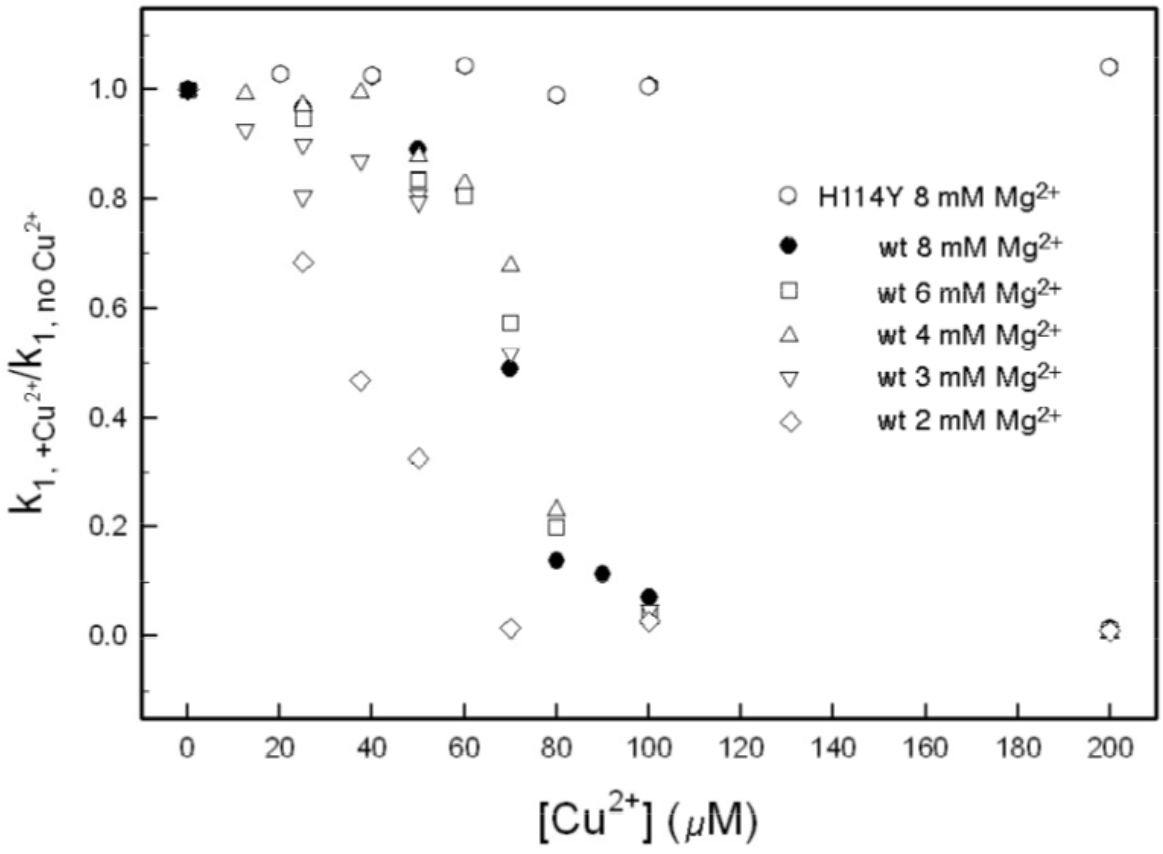


Figure 4.8 Inhibition of EcoRI cleavage by Cu^{2+}.

The Y-axis is the ratio of first-order cleavage rate constants with and without Cu^{2+}, the X-axis is Cu^{2+} concentration. Data are shown as means of at least 3 determinations for each point at 8 mM Mg^{2+}. Error bars (SD) are too small to be seen at this scale. Data and figure by Mike Kurpiewski.

4.4 ITC EXPERIMENTS REVEAL STOCHIOMETRY OF COPPER BINDING

The experiments described previously in this chapter show that EcoRI binds to Cu^{2+} by coordination to H114 and provide a tantalizing glimpse into the thermodynamic consequences of Cu^{2+} binding. One issue none of the above experiments address is stoichiometry-while we can establish the binding site and an indirect measure of the binding affinity of EcoRI for Cu^{2+}, these experiments do not provide a stoichiometry of binding. I have investigated the stoichiometry and thermodynamics of Cu^{2+} binding to EcoRI-specific complexes via isothermal titration calorimetry (ITC). In the following sections, I will briefly introduce ITC and discuss the characteristics of Cu^{2+} binding to EcoRI-specific complexes.

4.4.1 Isothermal titration calorimetry

Calorimetry is a technique that directly measures the heat associated with a chemical reaction triggered by the mixing of two components. At constant pressure, this heat is equal to ΔH, or enthalpy change of the process. In a typical isothermal titration calorimetry (ITC) experiment, a solution containing ligand is added to a matching solution containing macromolecule as a series of injections. (This is described in more detail later in this section.) The absorbed or released heat (q) that accompanies each increment of ligand addition is proportional to the extent of macromolecule-ligand binding. From these experimental data, the ΔH_{obs}, stoichiometry (n), and equilibrium constant (K_{obs}) for the reaction can be directly determined by fitting the experimental data. (Reviewed in [190], [191].) By using these experimentally determined values, the $\Delta G°_{bind}$ and ΔS_{obs} for the reaction can be calculated using the following thermodynamic relationships:

$$\Delta G°_{bind} = -RT\ln K_{Obs}$$

$$\Delta G°_{bind} = \Delta H_{obs}° - T\Delta S_{obs}°$$

Isothermal titration calorimetry therefore has the unique advantage of permitting the complete thermodynamic characterization of a giving binding process: n, ΔG, ΔH_{obs}, and ΔS_{obs}, are determined in a single experiment. [192] By repeating the experiment at multiple temperatures, ΔCp_{obs} may also be determined. (Reviewed in [191].) For these reasons, ITC has become an increasingly popular method for analyzing macromolecular binding reactions. [190] A number of valuable ITC techniques have also been developed for further dissecting the fundamental thermodynamic components of a reaction (ΔG, ΔH, and ΔS) into their contributing factors. [191] It should be noted, however, that the ΔH_{obs} value obtained from any single experiment includes contributions from buffer ionization enthalpy and any side reactions that occur as a consequence of ligand binding. Therefore, the ΔH_{obs} value often includes a number of artifactual contributions to heat which must be corrected for in order to obtain the true molecular binding enthalpy and entropy changes. (Reviewed in [193])

In an isothermal titration calorimetry experiment, there are two identical cells contained within an adiabatic shield in the calorimeter (Figure 4.8). One cell is a reference cell filled with buffer and the other cell is the sample cell, which contains the macromolecule. The temperature differences between the cells ($\Delta T1$) and between the sample cell and jacket ($\Delta T2$) are constantly monitored. The cells are constantly cooled while a temperature-controlled heating element maintains the cells at constant temperature.

In a typical ITC experiment, a syringe periodically injects equal volumes of solution containing the ligand into the sample cell containing the macromolecule. This causes either the absorption (endothermic reactions) or evolution (exothermic reactions) of heat. Consequently, this alters the amount of power input required to keep the sample cell at constant temperature. This time-dependent power input (measured in μCal/sec) is measured. By integrating the power input over time, the heat change resulting from a single injection can be determined. (Figure 4.8-A). As the experiment progresses, the peaks diminish in size as binding sites become saturated, until all sites are saturated, and only the heat of dilution is observed. [192]

After the experiment is completed, data analysis software [194] is used to integrate the area under each peak. The software is then used to plot the total heat per injection versus the molar ratio of ligand to macromolecule, referred to as the binding isotherm (Fig 4.8-B). Non-linear regression is used to fit the curve. The shape of the binding isotherm provides direct information on the binding association constant, stoichiometry, and enthalpy of binding. From these values, the entropy of binding can also be calculated, using the basic thermodynamic relationships described previously in this chapter.

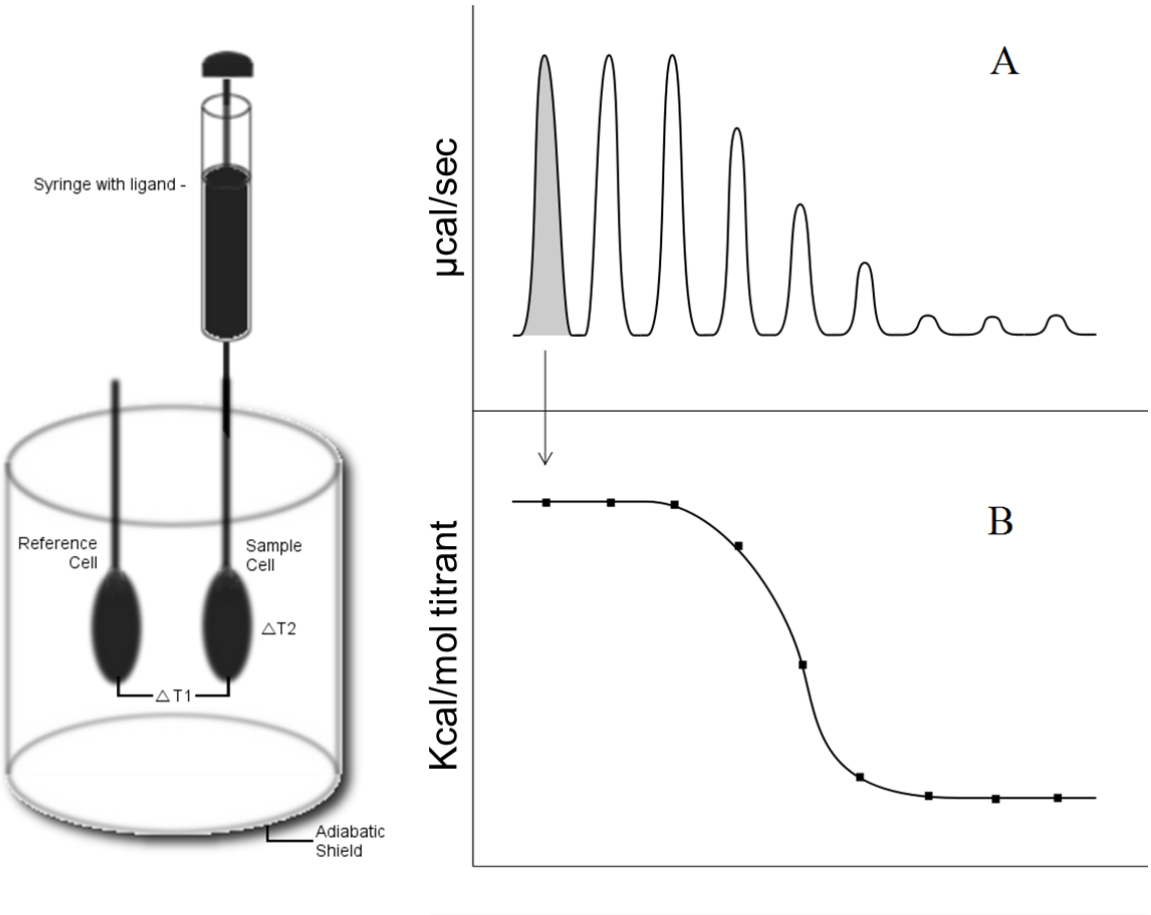


Figure 4.9 ITC experimental setup.

Left panel: schematic of an isothermal titration calorimeter. Within an adiabatic shield are contained a reference and sample cell. The difference in temperature between these cells (T1) and between the sample cell and the jacket (T2) are constantly monitored. The syringe contains the ligand, which is periodically injected into the sample cell during the course of the experiment. Right panel: A) An idealized representation of the data from an ITC experiment. Power input (µcal/sec) is measured over time. Each sample injection results in a peak. By integrating the area of each peak (shaded example), the ΔH for each injection is obtained. Right panel: B) After integrating each peak, the results are plotted as ΔH (kcal/mol titrant) versus the molar ratio of reactants. By fitting these points to a curve (solid line), the thermodynamic parameters can be obtained.

4.4.2 Experimental design for EcoRI-Cu^{2+} ITC experiments

ITC has been used for characterization of metal ligands as well as other biological ligands, but there are additional complications that must be taken in to consideration for experiments which use metals as a ligand for ITC. [195] Since most metals have competing interactions with buffers, in order to obtain clean ITC data, it is necessary to choose a buffer for which there is known to be one dominant metal-buffer species in solution. For copper this includes imidazole and Tris buffers. It is also ideal to choose a buffer for which the stoichiometry, stability, and enthalpy of the metal-binding complex are known, so that the thermodynamics of these interactions can potentially be subtracted out when attempting to determine molecular enthalpies. [195] For copper these buffers are imidazole and Tris. [196]

4.4.3 Low c-value ITC

The curvature of the binding isotherm is dependent on the binding association constant (K_A) and the concentration of binding sites n[M], where n is the number of ligand molecules that bind and [M] is the molar concentration of the molecule in the sample cell. The c-parameter [192] is a dimensionless experimental parameter that describes the shape of the binding isotherm:

$$c = n[M]\ K_A$$

For c values > 1000, a rectangular curve with height of ΔH is obtained. These curves provide little information besides the ΔH value. For c values between 10 and 1000, the shape of the curve is clearly sigmoidal and highly sensitive to small changes in c value. These curves are optimal for determination of all parameters. As the c value decreases below 10, the curve

becomes more and more shallow and the experimental parameters become more challenging to obtain accurately.

In theory, once the binding constant for a given association is known, the ITC experiment can be designed such that the c value is within the optimal range. In reality, for many systems, including copper binding to EcoRI, the concentrations that would be required are often in excess of the solubility of the macromolecule. [197] In these cases, referred to as "low c-value titrations," accurate values can still be obtained if the experiment is carefully designed. Since the shape of the curve is much more shallow and the upper plateau is not well defined, it is essential to obtain as many data points as possible so that there enough data points to accurately define the curve during the fitting procedure. Therefore, instead of the usual 20-30 data points suggested for ITC experiments, it is crucial to collect as many data points as possible for each experiment. [197] All ITC experiments presented in this dissertation were collected over at least 80 injections in order to obtain accurate results. Since each experiment is divided into so many data points, the signal to noise ratio must be maximized by very careful matching of syringe and cell solutions, degassing of solutions, and precise concentration determination for all reactants. A further caveat of "low c-value" titrations is that the accuracy of the ΔH parameter suffers from the loss of the upper part of the plateau, and thus the results must be carefully interpreted. [197]

4.4.4 Stoichiometry of Cu^{2+} coordination by the EcoRI-DNA complex

A representative binding isotherm is shown in Figure 4.9. As noted above, all of the ITC experiments were performed as "low c-value" titrations, so the full sigmoid curve is not apparent. However, due to the large number of points collected, a good-fitting curve (solid line in

bottom panel) can still be obtained from the data. All of the ITC experiments performed under stable conditions are reported in the Appendix, and summarized in Table 4.3.

Table 4.3 Summary of ITC data.

Temp (°C)	Stoichiometry[A]	$K_d(M^{-1})$[B]	ΔH_{obs} (kcal/mol)[C]	$T\Delta S_{obs}$ (kcal/mol)[D]
Imidazole[E]				
10°C	2.17(± 0.24)	1.2(±0.7) $x10^{-5}$	24.7 (±8.0)	31.3 (±8.5)
15°C	1.99(±0.04)	2.3(±0.1) $x10^{-5}$	20.8 (±1.5)	27.0 (±1.5)
21°	1.94(±0.48)	2.8(±3.0) $x10^{-5}$	34.8 (±17.5)	41.6 (±17.6)
Tris[F]				
15°C	2.12(±0.56)	3.0(±0.2) $x10^{-5}$	27.8 (±10.5)	25.6 (±4.5)

All numbers represent the mean ± standard deviation of at least three experimental determinations (experimental data shown in Appendix). Typical titrations consisted of 80 or 130 injections of 2.5ul or 5ul (respectively) of Cu^{2+} (0.375-1mM) into the overfilled sample cell (1.4ml) containing WT EcoRI (4.8-9um) complexed with specific DNA (7.5-20um) in Tris or imidazole buffer (see Appendices for DNA sequence and buffer composition)

A: Stoichiometry of Cu^{2+}: EcoRI-DNA complex determined directly from data fitting of binding isotherm in Origin software.

B: Kd directly determined from data fitting of binding isotherm in Origin software.

C: ΔHobs directly determined from data fitting of binding isotherm in Origin software. These data are background corrected for buffer dilution enthalpy but not for other potential contributions such as buffer ionization enthalpy; thus observed values do not represent the true molecular enthalpy change upon binding as discussed in section 4.4.5.

D: Determined from the following equations:

$$\Delta G = -RT\ln K_A$$

$$T\Delta S = \Delta H - \Delta G$$

E and F: Buffers at pH 7.8. ITC buffers in Appendix.

The data for each experimental condition consistently show a stoichiometry of two copper atoms per dimer (one Cu^{2+} binding per subunit) (For example, 1.99±0.04 for the 15°C imidazole data.) Since a stoichiometry of two was obtained, I also attempted to fit the data using two-site interacting and sequential models. None of the curves were able to be fit with these models, thus the data are consistent with two identical, non-interacting binding sites. The apparent K_d is in the range of ~1-3x10^{-5} (M^{-1}) which is good agreement with the constant obtained by Kurpiewski, given that the experiments were performed under different buffer conditions.

4.4.5 Thermodynamic contributions to Cu^{2+} binding to EcoRI-DNA complex

For all experimental conditions, the observed enthalpy change ΔH_{obs} was positive, (unfavorable enthalpy change) and the calculated observed entropy change, $T\Delta S_{obs,}$ was positive (favorable entropy change). The overall coordination of copper to EcoRI-DNA complexes therefore appears to be entropy-driven under these experimental conditions. However, the enthalpy change directly measured in the ITC experiment is an "observed" enthalpy change (reviewed in [193]), which includes contribution from both the true molecular enthalpy change as well as from buffer ionization enthalpy (ΔH_{HB}), and interactions between the metal and the buffer (ΔH_{MB}).

For reactions where protons are absorbed or released, a major contributing factor to the observed binding enthalpy is the ionization enthalpy of the buffer solution used for the experiment. [198] In these systems, the observed enthalpy is comprised of the sum of the enthalpy change that would occur in a buffer with zero ionization enthalpy ($\Delta H°_0$), and the ionization enthalpy of the buffer multiplied by the number of protons exchanged ($n_H\Delta H_{HB}$).

Metal ions often displace protons upon binding to proteins, therefore the number of protons displaced must be determined in order to gain a more accurate view of the binding enthalpy. A further complication of examining data obtained from protein-metal binding interactions is that metal ions generally interact with the buffer such that other heat-generating events occur in coupled equalibria with the protein-metal interaction. [199] When protein-metal binding interactions are investigated by ITC, therefore, the additional metal-buffer contributions to enthalpy (ΔH_{MB}) must be considered in addition to the buffer enthalpy of ionization (ΔH_{HB}). The following scheme summarizes the reactions occurring in the ITC sample cell, where M represents the metal, m the number of metals binding the ligand, L the ligand, H a proton, and n the number of protons released [196]:

$$mMB + nHL \leftrightarrow mML + nHB$$

Table 4.4 Contributions to enthalpy changes in metal-ligand ITC experiments

Reaction	*Enthalpic Designation*
$mM + L^- \leftrightarrow mML$	$m\Delta H_{ML}$
$mMB \leftrightarrow mM + B$	$-m\Delta H_{MB}$
$n\ (H+B \leftrightarrow HB)$	$n\Delta H_{ion}$
$H_nL \leftrightarrow L + nH$	$-n\Delta H_{HL}$

These reactions can be condensed as follows:

$$\Delta H_{obs} = m*(\Delta H_{ML} - \Delta H_{MB}) + n*(\Delta H_{HB} - \Delta H_{HL})$$

If ΔH_{MB} and ΔH_{HB} are known for two different buffers, it is possible to conduct the ITC experiments under two different buffer conditions and use the difference in the observed enthalpy change to determine the number of protons displaced and thus to extract the molecular binding enthalpy for the system. I designed my ITC experiments in this way with the aim of

extracting the true molecular enthalpy change of copper coordination from the observed values. Unfortunately, the ΔH_{obs} values obtained have deviations which are too large to permit such analysis. (Table 4.3) This is likely to be because the experiments were performed at low c-values as discussed in the previous section. Therefore, it is unfortunately not possible to conduct a more detailed thermodynamic dissection of the data.

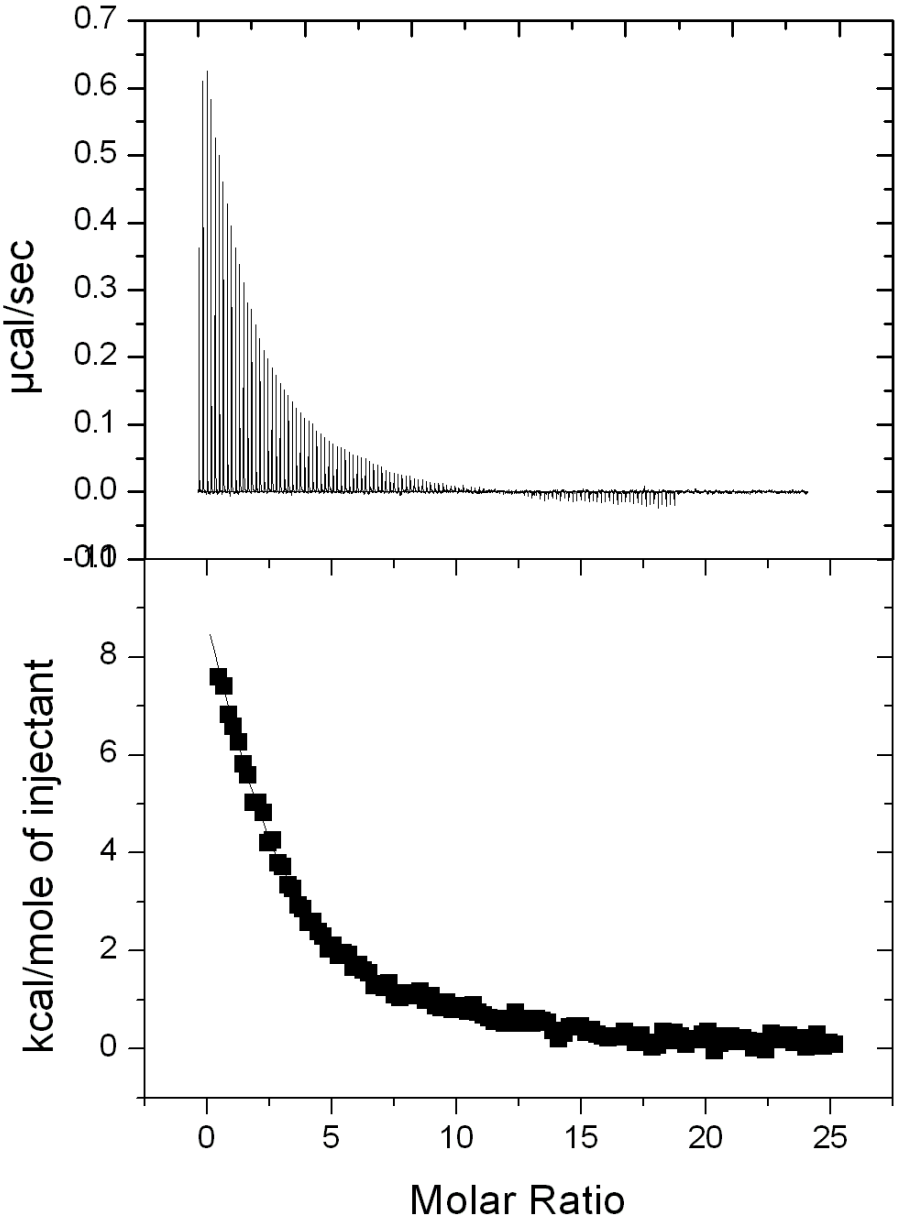


Figure 4.10 Representative thermogram of ITC data.

Experimental data from 125 injections of 2.5ul Cu^{2+}(1mM) into sample cell containing EcoRI (9um) and DNA (20um) in imidazole buffer (10mM Imidazole pH 7.8, 0.24 M NaCl) at 15C. The top panel represents the baseline corrected raw data, while the bottom panel represents the peak-integrated reaction heats plotted versus the molar ratio of copper to EcoRI-DNA complex. The squares indicate the experimental data for each injection, the solid line represents the best fit curve for the data.

4.5 CONCLUSIONS AND FUTURE DIRECTIONS

We have identified the location of a previously unknown metal binding site in EcoRI. The ESEEM experiment determined that Cu^{2+} coordinates to a histidine, and by triangulation of distances by DEER experiments, that histidine was identified as H114. This residue also interacts with the GApATTC phosphate of the binding site, and several mutations of this residue are known to be promiscuous. Characterization of Cu^{2+}-EcoRI binding by ITC has determined that two Cu^{2+} atoms bind to two identical sites in an independent manner, confirming the hypothesis that one Cu^{2+} binds to each monomer of the enzyme. When Cu^{2+} is bound to EcoRI, binding to DNA is enhanced, but cleavage activity is reduced. Further biochemical assays and molecular dynamics simulations are being conducted by Mike Kurpiewski (Jen-Jaconson lab) and Ming Ji (Saxena lab) in order to characterize the binding of Cu^{2+} to the EcoRI-DNA complex.

5.0 DISCUSSION

In this work I demonstrate the application of SDSL-ESR to examine differences in the solution structure and dynamic behaviors of EcoRI bound to different classes of DNA site. Using DEER, we were able to determine that the arms envelop the DNA and are similarly oriented in specific, miscognate, and nonspecific complexes. Differences in the distance distributions indicate that the noncognate complexes may experience greater conformational freedom than the specific complex. We also were able to use CW to detect differences in the mobility of residues at different positions in the protein, and at these positions in the different classes of complex. Our results show that the residues likely experience a hierarchy of mobility where the core domain residues are less mobile, residues in the middle of the arms are more mobile, and residues on the loops at the ends of the arms are highly mobile. The spin label attached to residue I197C at the interface between the protein and DNA shows spectra with decreased apparent mobility in the miscognate complex relative to the nonspecific complex, and further reduced mobility in the specific complex. Additionally, we used ESR to identify a novel Cu^{2+} binding site in EcoRI, and to determine that the Cu^{2+} coordinates to H114. Furthermore, we were able to use ITC to determine the binding affinity and stoichiometry of Cu^{2+} to EcoRI.

Work is still in progress to further deconvolute and complement the information obtained from our ESR results. Molecular dynamics simulations are being used to examine the behavior of the nitroxide spin label in order to increase the resolution of our DEER results. Spectral

simulations will be used in order to extract additional parameters from the CW spectra; aiding in deconvolution of the contributions to the motion of the spin label. Molecular dynamics simulations and additional biochemical experiments are in progress in order to further characterize the coordination of the Cu^{2+} ion and to investigate the effects of Cu^{2+} binding to the protein.

As discussed in the previous chapter, Cu^{2+} highly toxic and therefore sequestered in the cell, so Cu^{2+} binding to EcoRI is likely not physiologically relevant. However, it still provides an extremely useful label near the active site. We have identified this previously unknown Cu^{2+} binding site, and established that we can determine Cu^{2+}-nitroxide distances in this system. With this information, we can now prepare protein labeled with nitroxide at different positions; addition of Cu^{2+} to these spin-labeled proteins enables determination of additional distances. This permits the measurement of nitroxide-nitroxide distances as well as the measurement of intra- and inter-subunit distances between the Cu^{2+} (which is in a fixed position in the protein) and the nitroxide. Such experiments can provide multiple distance measures with a single sample, which will enable better determination of residue positions.

The experiments described herein are the first to demonstrate that ESR methods such as DEER and CW can detect measurable differences for a protein bound to different classes of DNA complex. Since noncognate complex structures are rare, these techniques could be applied to investigate the noncognate structures of other DNA binding proteins such as restriction enzymes and repressor proteins. Additionally, these methods can be used to investigate the differences in solution dynamics of complexes for proteins which are not amenable to the conditions required for NMR. The “arms” or arm-like projections are seen in a number of

different restriction enzymes, [5] and it would be interesting to determine if other restriction enzymes also enfold the DNA in all three classes of complex.

As discussed in the Introduction, the binding affinity and $\Delta C°p$ differ for EcoRI binding to the specific site in different flanking contexts. As the binding affinity improves, the $\Delta C°p$ becomes more negative, indicating that the interface is more intimate and the configurational freedom of the complex is more restricted. [26] It would therefore be intriguing to investigate whether CW spectra are able to detect differences for the specific complex in the context of different flanking sequences. It would also be interesting to investigate whether differences could be detected among CW spectra for various miscognate sequences (only the AAATTC miscognate site has been examined thus far). The I197C mutant, which permits a spin label to be attached right at the protein-DNA interface, would be ideal for these experiments. Additionally other mutations could be engineered close to the protein-DNA interface for these investigations.

Although we were able to obtain a great deal of information about the arms of the enzyme in the protein-DNA complexes, the positions and behavior of the arms in the apo enzyme are still unknown. While the crystal structure for the free enzyme has been determined, the arms did not have sufficient electron density to model their positions; they are "invisible" in the free enzyme structural model. Determining the positions of the arms in the free enzyme would provide valuable insight to the binding process- it would enable an investigation of the changes that occur in the arms between the bound and unbound states of the enzyme. For example, it is unknown whether the arms have a more "open" conformation in the free enzyme that closes upon binding the DNA, or how the protein initially associates with the DNA. Determining the positions of the arms in the free enzyme could help answer these questions. Originally, the concentrations required for ESR were in excess of what we were able to achieve

for the free enzyme without resulting in aggregation and precipitation of the enzyme. However, we have recently determined that 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) is an excellent non-denaturing detergent that enhances the solubility of EcoRI. This may allow us to produce spin-labeled apo EcoRI at concentrations high enough for ESR experiments to be performed.

Another intriguing avenue of research is direct observation of facilitated diffusion by EcoRI along DNA. We have described elsewhere that the K249C mutant of EcoRI is ideal for labeling. [200] In addition to providing a core domain reference point for SDSL-ESR, this mutant EcoRI protein may also be biotinylated and still retain binding activity. [200] We have shown that biotinylated K249C can be attached to a fluorescent nanosphere and used for detection of EcoRI specific sites on λ DNA both when stretched on polylysine-coated slides and when trapped at a stagnation point under microfluidic planar extension. [200], [201] This mutant would be an ideal candidate for other single-molecule studies which require the attachment of a fluorescent label.

The use of single-molecules studies to directly observe the diffusion of proteins along DNA has become increasingly popular in recent years. [202] For example, Total Internal Reflection Florescent Microscopy (TIRFM) has been used to observe facilitated diffusion of RNA polymerase, *lac* repressor, Rad51 DNA recombinase, and other proteins. (Reviewed in [202]) In order to investigate facilitated diffusion by EcoRV, Biebricher and colleagues generated EcoRV protein labeled with quantum dots. [203] They used this labeled protein to directly observe the diffusion of EcoRV along DNA that was held by a pair of optical tweezers. They found that both sliding motions along the DNA, as well as “jumps” (fast 3D translocations)

were observed. [203] Similar techniques could be used with labeled K249C EcoRI protein in order to directly observe and investigate the movements of this protein along DNA.

6.0 METHODS

6.1 CONSTRUCTION OF MUTANTS

6.1.1 Plasmid Background

The parent plasmid for the mutants discussed in this work is pPS12. This plasmid was constructed by Paul Sapienza [6] by amplifying the wild-type EcoRI endonuclease gene from plasmid pMB3 [204], and inserting it between the NheI and BamHI restriction sites of pET24a (Novagen). In this plasmid the transcription of the endonuclease gene is directed by the T7 RNA polymerase promoter. All genetic manipulations were carried out in strain DH5α (F-$^{\varphi}$80 Δ(lacZ)ΔM15 Δ(lacZYA-argF) U169 *deoR recA1 endA1 hsdR17 (rk- mk+) phoA supE44 λ- thi-1 gyrA96 relA1).* Since this strain lacks the T7 RNA polymerase, this minimizes any potential toxic effects of EcoRI expression during cloning.

6.1.2 Site-Directed Mutagenesis

The first two mutants created in this work, S180C and R131C, were cloned using the QuikChange^{TM} protocol developed by Stratagene (now Agilent). In this method, point mutations are introduced into the plasmid by using a pair of complementary primers containing the desired point mutations to amplify the parent plasmid within a PCR reaction. (See Appendix.) Any

remaining template plasmid is then eliminated by digestion with *Dpn*I, a restriction enzyme which cleaves methylated or hemi-methylated GATC sites. (Figure 5.1) Subsequently, the mutant plasmid is transformed into the cloning strain. Putative mutant colonies are then confirmed by sequence analysis of the EcoRI gene. Subsequent mutants were cloned using modifications of the Quickchange protocol as described below.

6.1.2.1 Primer Design

Due to the fairly low G+C content of the EcoRI gene, it is generally impossible to design primers with characteristics that follow the guidelines set forth in the QuikChange protocol, namely that of primer pairs whose primer-template [Tm] values are at or above 78°C, while maintaining length between 24-45 bases. All mutagenic primers (See Appendix) were designed according to my revised design "priorities":

- Complementary primer pair containing the desired mutation. (By definition)
- Mutation has at least 10, preferably 15 bases of correct sequence on both sides (I found this to be essential.)
- Mutation centered in primer- I found 'asymmetric' primers to be unsuccessful.
- Primers terminate in one or more G or C on both ends
- Tm must be at least 76.
- Above priorities are essential, even if the primer length exceeds 45bp.
- If above priorities are fulfilled and primer length would not exceed 45bp, also try to design the primer such that :
- Tm> 78°C
- G+C content is at least 40%

In order to design primers that could successfully anneal with the EcoRI gene to incorporate the mutant bases, most primers were longer than the suggested 45bp- generally 50-60 bp. Long primers greatly impair mutagenesis efficiency, because the 100% complementary primer dimer formation is much more favorable than primer-template annealing, due to the mismatch in bases

between the primer and the template. The S180C and R131C mutants were cloned after many attempts, for subsequent mutants a more efficient protocol was developed.

6.1.2.2 Two-Stage Mutagenesis Protocol

In order to address the problem of primer-dimer formation, I developed an adaptation of the "two stage PCR mutagenesis protocol". [205] In this method, the primer-dimer problem is avoided by setting up two separate PCR reactions, one containing only the forward primer and the other containing only the reverse primer. (Figure 5.1) After 1-5 cycles of amplification (in my system 5-10 gave a better success rate), the two reactions are then combined for an additional 16 (in my system 20) cycles. Separating the primers for the first few amplifications prevents the primer-dimer problem and allows the accumulation of template now containing the point mutation (and thus 100% complimentary to the other primer). Upon combining the reactions the primer to mutated template hybridization can compete more effectively with the primer-dimer hybridization. This protocol is outlined in the Appendix.

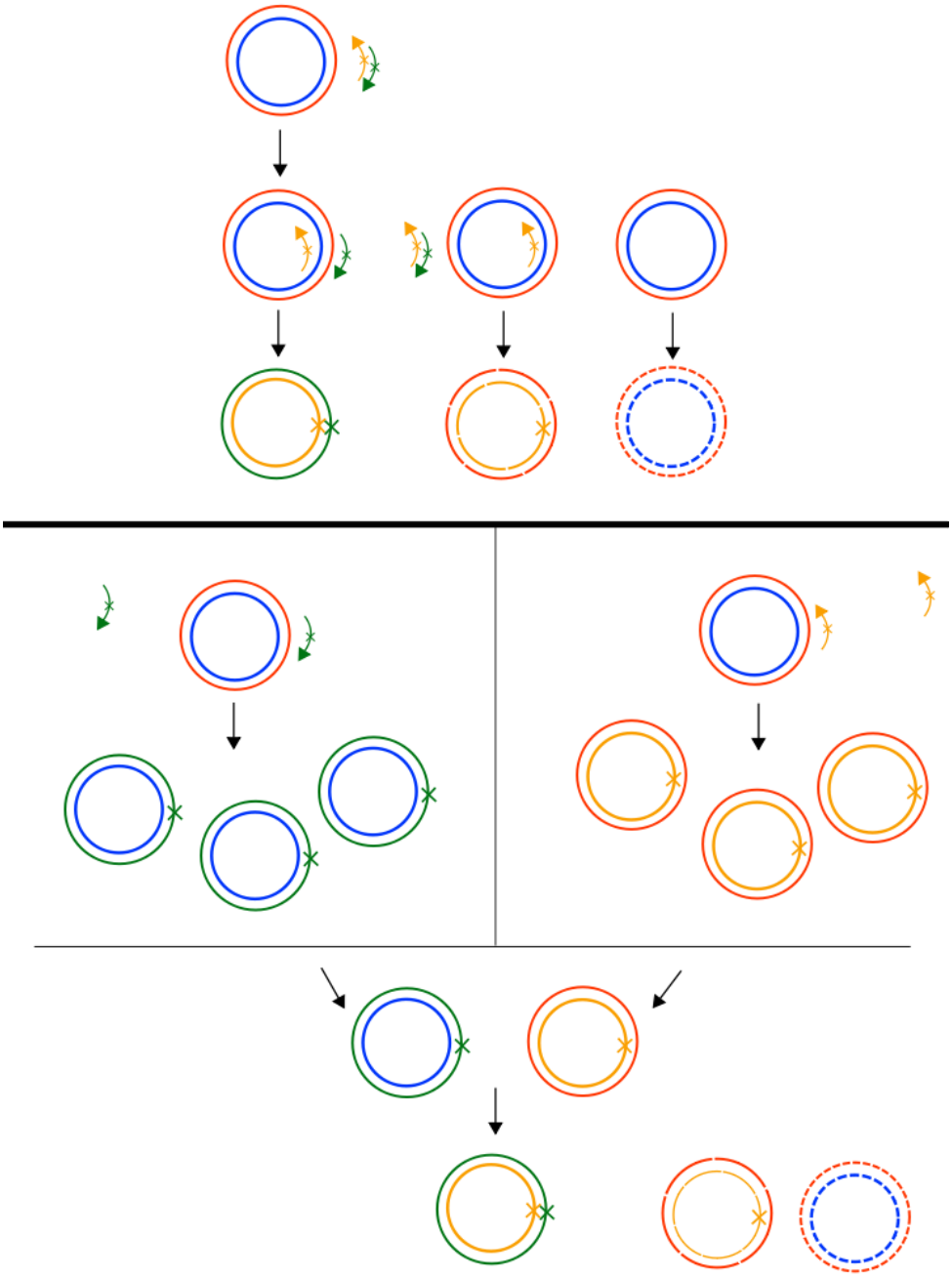


Figure 6.1 Schematic of mutagenesis methods.

Parent plasmid template (red and blue circles) undergoes PCR in the presence of mutagenic primer (green and orange arrows with X's). Fully mutant plasmid (green and orange circles with X's) survives DpnI digestion, while remaining unmethylated parent plasmid is fully degraded (dashed circles), and half-mutant (hemi-methylated) plasmid is degraded at a reduced rate. Top panel: Primer dimers are more complimentary to each other than then to the parent template, and compete for binding. Bottom panel: In the two-stage protocol the forward and reverse primers are used in separate PCR reactions in order to generate a population of mutant template in the absence of the potential for primer-dimer competition. The reactions are subsequently combined and the mutagenized template undergoes further amplification before digestion.

6.1.2.3 Reduction of false positives

Even with the optimization of primer design and the two stage protocol, a significant source of false positive colonies (50% of screened colonies or more) originated from hemi-methylated template. In this case an original parent template plasmid has hybridized to one of the mutagenized templates; these are cleaved by DpnI 60x more slowly (NEB supplemental information) and upon transformation result in a colony with a mix of plasmid types. These are identified upon sequencing as a mix of parent and mutant sequence. I found that a two-hour DpnI digestion at 37°C followed by adding fresh DpnI and a second sixteen-hour digestion at room temperature drastically reduced this source of false positives to 5-10% of screened colonies. All putative mutants were screened by sequencing before proceeding.

6.1.3 Construction of MBP-Fusion proteins

Most of the EcoRI mutant proteins I generated had low expression or were highly insoluble in cell lysate, therefore they produced very low yields of purified protein (discussed further in subsequent sections). In order to address this issue, I constructed fusions of our wild-type and mutant EcoRI genes to maltose-binding protein (MBP). These fusions are known to result in greatly enhanced solubility and yield of purified protein. [206] I constructed fusions by digesting pMAL^{TM}-c4x vector [207] with XmnI and BamHI, resulting in a 5' blunt end overlapping the Factor Xa protease cleavage site in the polylinker downstream from the *malE* gene, and a 3' sticky end. The cloned gene is prepared by PCR amplification from the parent plasmid, followed by blunting with T4 polymerase, and digestion with BamHI, this results in a product with a 5'

blunt end and a 3' sticky end. The cloned gene is then ligated into the pMAL™-c4x vector to form the MBP-EcoRI fusion construct. This cloning strategy was based on the suggestions outlined in "Strategy I" of the NEB pMAL™ manual [207] and the protocol is detailed in the Appendix.

6.2 METHODS DEVELOPMENT IN ECORI PROTEIN EXPRESSION AND PURIFICATION

6.2.1 Purification methods in pPS12 background

Initially, the mutant proteins were produced and purified according to previously established protocols in our lab. [6] Briefly, plasmid encoding the EcoRI endonuclease is transformed into the *E. coli* strain ER2566 (=*fhuA2 [lon] ompT lacZ::T7 genel gal sulAll* Δ*(mcrC-mrr)l 14::IS10R (mcr-73::miniTn10—Tet^S)2 R(zgb-210::Tn10-Tet^S)end A1 [dcm]* (NEB) containing the plasmid (pAXU22-8, gift of NEB) which encodes expression of the EcoRI methylase. Cultures are grown to mid-log phase at 37°C in LB media containing glucose to provide catabolite repression of the EcoRI gene prior to induction. The cultures are then induced by addition of IPTG and grown for two hours at 21°C. Subsequently, cells are frozen down, sonicated, and the resulting lysate is purified by ion exchange chromatography over columns of phosphocellulose, biorex 70, and heparin. (Figure 5.2) This method was used by Paul Sapienza and Lance Mabus to prepare the protein we used for the R131C DEER samples. However, while the DNA binding capability of the cysteine mutants is not hindered, (as discussed in Chapter 2) the expression and solubility of these mutant proteins are affected to varying degrees.

For the wild-type EcoRI protein, the above purification method yields an average of 0.5mg of purified protein per liter of cell culture. However, the highest yielding mutant protein, S180C, only yields an average of 0.1- 0.2 mgs per liter using this method. For the long distance DEER measurements, approximately 4mgs of protein were needed for each sample preparation, thus each sample required the processing of over 40 liters of cell culture. The other mutant proteins demonstrated even worse yields, and some of the mutants could not be purified at all using the original method. In some cases, this was because the overall expression of the mutant protein was too low, in others the protein formed insoluble inclusion bodies and could not be recovered for purification. Exhaustive attempts to improve yields via altering growth temperatures during induction and varying length of induction time failed to produce a satisfactory improvement in yields.

To attempt to address this problem, we initially changed our protein expression system from IPTG induction of mid-log growth cultures to the auto-induction system based on the methods described by Studier. [208] Plasmid pPS12 or mutant derivatives were transformed into the BL21(DE3) strain in the presence of the methylase plasmid (pAXU22-8) and grown for 16 hours in ZYP-5052 media (Appendix). This is a rich buffered media containing a mix of glucose, α-lactose, and glycerol, resulting in suppression followed by slow induction of protein expression as the culture approaches saturation. This method improved the expression levels for some of our mutants, however this did not resolve our solubility problems. The purification method used with this method is the same as described previously above. This method was used to produce the protein for the DEER experiments conducted on the S180C/K249C complexes and the S180C specific and nonspecific complexes.

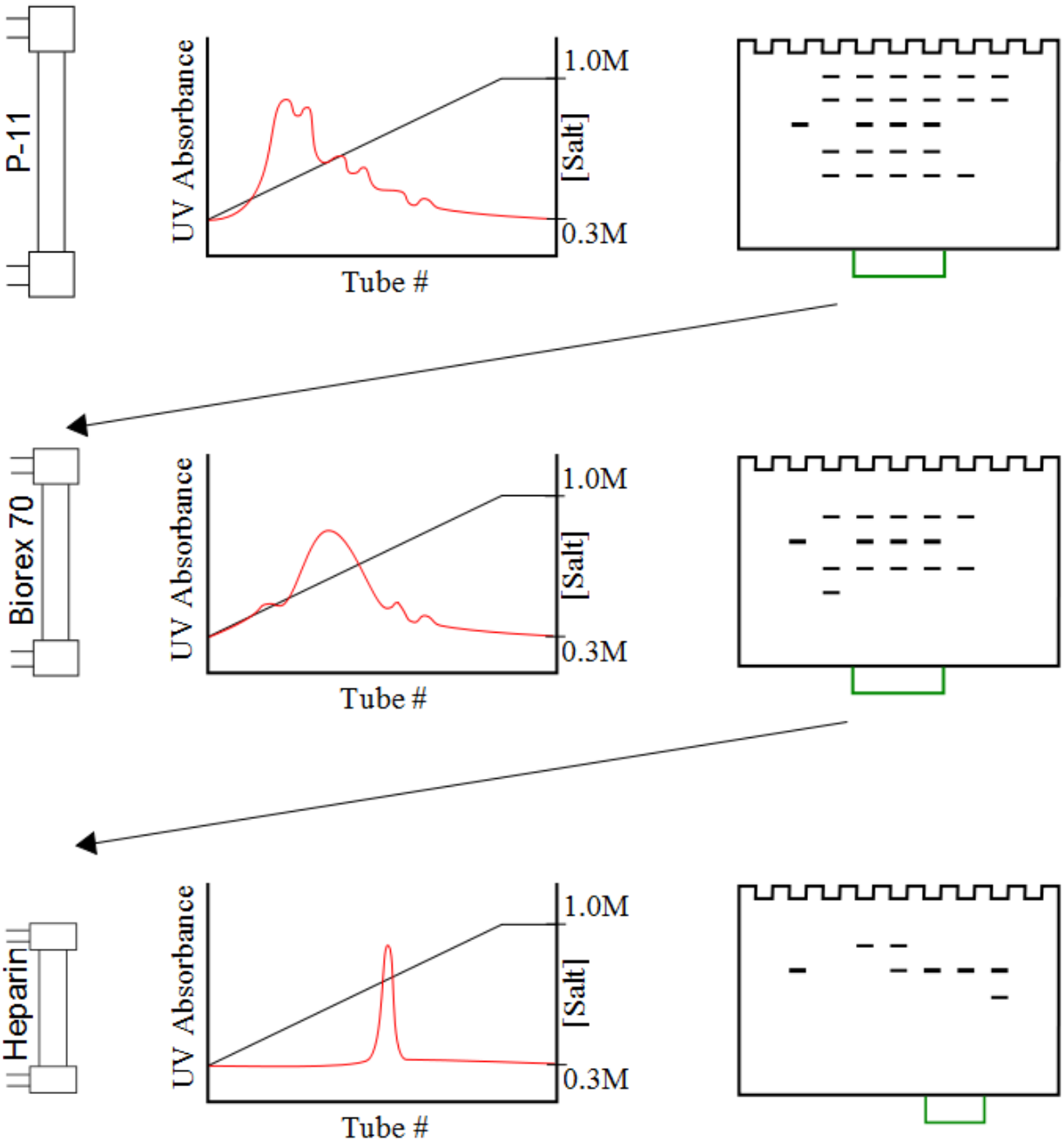


Figure 6.2 Schematic of standard protein prep.

Lysate from cells induced to overexpress EcoRI is run over phosphocellulose resin, washed, and eluted with a gradient of NaCl via HPLC. Peak fractions (determined by UV absorbance, red trace in center graph) are examined by gel electrophoresis (the band in the second lane indicates pure EcoRI used as a marker) The purest fractions (green bars on 'gels' on right) are run over Biorex-70 resin, and the purest fractions from this separation are further separated by binding and elution from heparin resin. Pure EcoRI fractions are pooled and quantified.

6.2.1.1 Purification methods in MBP background

The previously discussed protein purification methods were ultimately impractical for achieving our experimental goals. As mentioned above, dozens of liters of cell culture and weeks of purification were required for each labeling attempt. Additionally, some of the mutants had extremely low expression or solubility such that it was impossible to obtain pure protein for experiments. In order to address these issues, fusion constructs with Maltose Binding Protein (MBP) were employed. This system was chosen because of three important characteristics; first, the system is designed so that proteolytic cleavage of the fusion construct will produce the exact wild-type protein (I was able to design the construct such that there are no linkers or tails of any kind). Second, the MBP-fusion is noted for improving the solubility of most attached proteins. [206] Third, this system allowed simplification of the previous purification procedure. With the fusion construct I was able to design a simpler purification protocol as described below. In a preliminary test with wild-type EcoRI, purified protein yields increased from 0.5mg/L to ~1.3 mg/L- a nearly three-fold improvement. In this system, I was able to consistently obtain yields of 0.5 mg/L for a variety of EcoRI mutants, even for mutant proteins that I had previously been unable obtain sufficient soluble protein to purify.

In order to purify EcoRI protein from the MBP- fusion construct, I developed a protocol that was a hybrid between our existing protein purification protocol and suggestions from the pMAL manual. [207] In this protocol, cultures are grown to mid-log phase at 37°C in LB media containing glucose to provide catabolite repression of the EcoRI gene prior to induction. The cultures are then induced by addition of IPTG and grown for two more hours at 37°C. Subsequently, cells are frozen down, sonicated, and the lysate run over an amylose column. (See the schematic in Figure 5.3) This column binds to the maltose binding protein of the fusion. The

column is then washed extensively (Protein Prep Buffer A, 0.35M NaCl, pH 7) and the fusion protein eluted with a 100ml gradient of buffer containing maltose (Protein Prep Buffer A+M). This provides a fairly pure pool of MBP-EcoRI fusion protein. This fusion protein is then exhaustively digested in the eluent buffer for several days with Factor Xa protease at 4°C. CHAPS is added to 0.1% to prevent precipitation of protein during this process, as once the fusion protein is cleaved the EcoRI becomes less soluble. This pool is then run over a 5ml heparin column. As heparin is highly negatively charged and the MPB has a slight net negative charge, only the EcoRI binds to the heparin column. After washing the column, (Protein Prep Buffer A, 0.35M NaCl, pH 7) the EcoRI can then be eluted with a 100ml salt gradient (0.35MNaCl Protein Prep Buffer A to 100% Protein Prep Buffer B) as pure protein. All EcoRI protein was determined to be >99% pure by SDS gel electrophoresis.

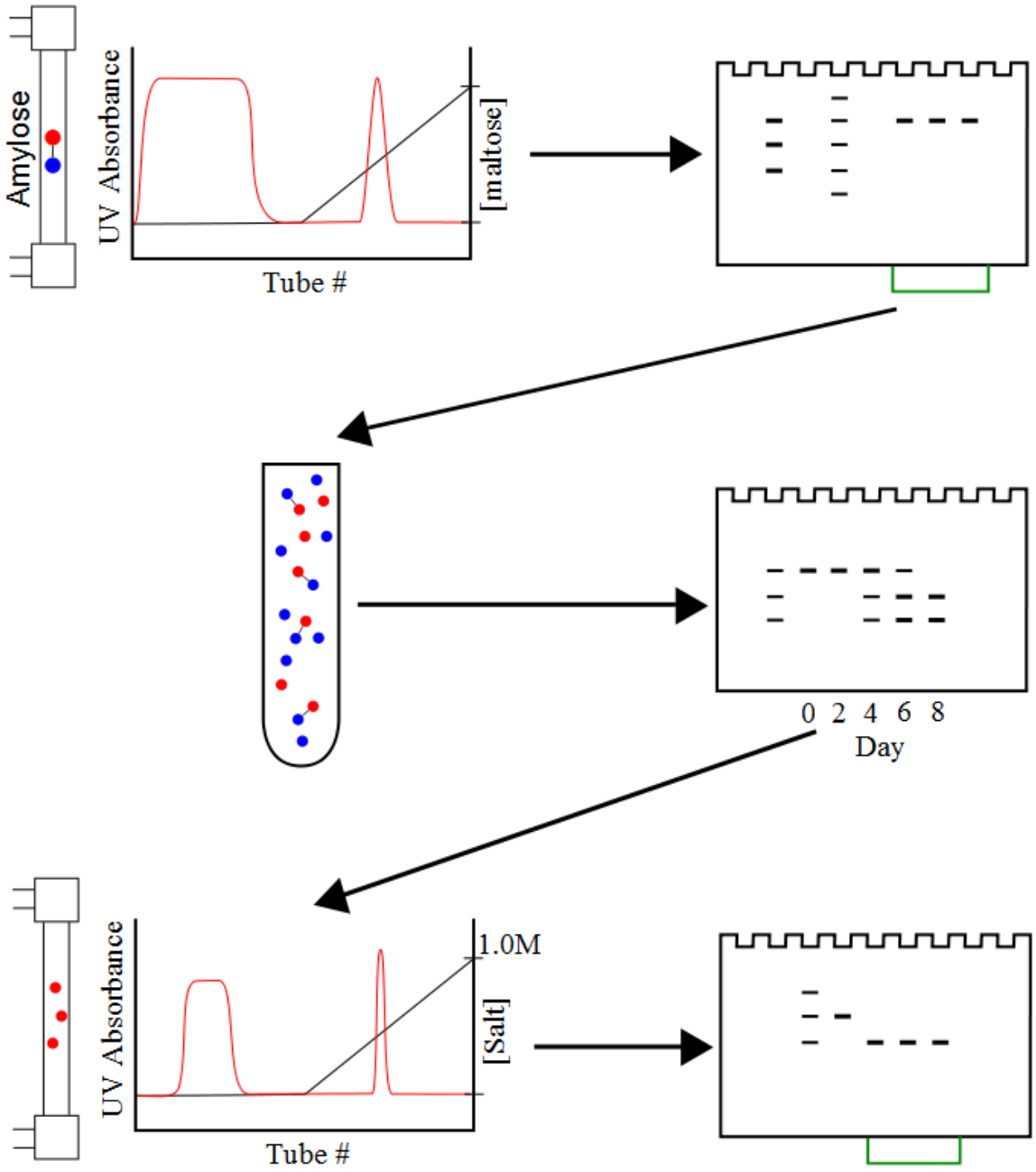


Figure 6.3 Schematic of MBP-EcoRI fusion protein purification method.

Lysate from cells induced to overexpress EcoRI (red dot) – MBP (blue-dot) fusion protein is run over amylose resin and rinsed. Only MBP-EcoRI fusion stays bound to this resin. The fusion is eluted with a maltose gradient. Peak fractions (determined by UV absorbance, red trace in center graph) are examined by gel electrophoresis (the bands in the second lane indicate fusion (top) MBP (middle) and EcoRI (lower) used as a marker). The pure fusion is digested with factor Xa and the reaction is monitored until the fusion is fully digested (center). The digest is run over a heparin column- MBP flows through, while EcoRI binds and is eluted with a NaCl gradient. Pure EcoRI fractions are pooled and quantified.

6.3 PROTEIN QUANTIFICATION DETERMINATION

Protein concentrations were determined by loading an SDS-gel (5-15% gradient) with known mass of EcoRI reference sample (concentration determined by direct analysis using norluecine as an internal standard [136]) with known volumes of pure EcoRI of unknown concentration. The gels were stained with Coomassie Blue and scanned. Optical densities of the bands from the reference protein sample are used to generate a standard curve from which the concentration of the unknown samples is determined. (See Appendix).

6.4 DNA PURIFICATION, QUANTIFICATION AND DUPLEX CONFIRMATION

Oligonucleotides were obtained from Integrated DNA Technologies (IDT). DNA for direct binding experiments was further purified by electrophoresis separation on 16% polyacrylamide gels. UV shadowing was used to locate the DNA bands, which were excised from the gels. The DNA was eluted from the gel slices by incubation in C-18 buffer at 37°C overnight with shaking. The DNA solution was adsorbed to Alltech C-18 columns, washed with 5 volumes of C-18 buffer and 5 volumes of water, then eluted with 30% ethanol.

Concentrations of single-stranded DNA was determined by UV spectroscopy as previously described. [6] Stochiometric amounts of single stranded DNA were combined in filter buffer with 0.22M NaCl. The strands were annealed by heating in a water bath to >95°C for 15 minutes and being allowed to cool to room temperature overnight. The duplex formation was confirmed by nondenaturing electrophoresis on 16% polyacrylamide gels examined by staining with GelRed.

6.5 EQUALIBRIUM BINDING

Equilibrium binding constants (K_A) were determined by the direct binding filter binding method as has been described previously. [6], [33] Briefly, fixed amounts of P^{32} labeled DNA are titrated with a series of EcoRI concentrations. The reactions are filtered through nitrocellose, and the amount of bound complex retained on the filter after washing is determined by scintillation counter. The resulting data are fit to a binding isotherm, which yields the observed equilibrium constant.

Because the spin-labeled ESR samples contained a large molar excess of DNA, in order to determine the binding affinity for spin labeled EcoRI it was necessary to eliminate the competing substrate. The DNA in the sample was cleaved by the EcoRI by the addition of Mg^{2+}, then this Mg^{2+} was chelated with the addition of excess EDTA. Since the EcoRI substrates were short (12-14 base pairs) the resulting cleavage products are too short to compete effectively with the full-length specific substrate. This is confirmed by the fact that the combined effects of the "mock" labeling procedure on wild-type EcoRI only result in a modest decrease in binding activity. After the chelation step, the binding assay was then carried out with the 32P-labeled DNA substrate as described above.

6.6 SITE DIRECTED SPIN LABELING OF ECORI

In the ESR experiments, a methane thiosulfonate spin label (MTSSL) must be covalently attached to the cysteine residues of the protein in a process referred to as site-directed-spin-labeling (SDSL). The specific MTSSL used in this work is (1-Oxyl-2,2,5,5-tetramethyl-3-

pyrrolin-3-methyl-) methane thiosulfonate. [107] The specifics of this protocol have been modified numerous times during the course of this dissertation in order to optimize the yield of functional EcoRI-MTSSL protein. However, the overall scheme has remained the same as described below. Unless otherwise noted, all steps are performed at 4°C:

1) EcoRI protein is incubated at 4°C in the presence of a large excess (10-20 fold molar excess) of DNA (specific, miscognate, or nonspecific as appropriate to the sample). In all cases, physiological salt concentrations (0.22M) and pH (7.3) were used. These conditions ensure that virtually all of the protein in these samples exists in the bound form. [115] This both generates the protein-DNA complex of interest and also ensures that the protein-DNA equilibrium is able to compete effectively against the dimer->tetramer equilibrium in subsequent steps.
2) The MTSSL is dissolved in DMSO and added to the EcoRI-DNA complexes to a 100-fold molar excess. DMSO is added to a final concentration of 10%, which aids in maintaining the solubility of the spin label. The spin-labeling reaction is allowed to proceed at 10°C for at least four hours.
3) The excess unbound spin label is removed via a combination of exhaustive dialysis and concentration/dilution exchange cycles in an Icon spin concentrator. A DMSO concentration of 10% is maintained for the first couple of exchanges to aid in maintaining the solubility of the spin label as the excess is removed.
4) For the S180C, K249C, R123C, and I197C DEER experiments, deuterated buffer is used during the later steps of the removal of excess spin label. I found that attempting to increase the percent of deuterated water to higher than 95% adversely affected the

solubility of our protein, so the buffers and exchanges were planned such that at least 5% water was present.

5) The EcoRI-MTSSL-DNA sample is concentrated to at least 75um using an Icon protein concentrator.
6) Remaining cosolutes are added. Glycerol is added to a final concentration of 30% for DEER samples. For S180C, K249C, I197C, and R123C samples, Deuterated glycerol (DOCD2)2CDOD or "D8" glycerol was used. For CW samples, appropriate concentrations of glycerol or Ficoll 70 were added.

It should be noted that all steps are conducted at 4°C, with the exception of the labeling reaction which is incubated at 10°C. Careful handling of the protein is required at all steps, as EcoRI precipitates easily at the concentrations required for ESR.

6.7 ISOTHERMAL TITRATION CALORIMETRY

Wild-type EcoRI and specific DNA (1.5-fold molar excess of protein) were combined and exhaustively dialyzed in ITC buffer at 4°C with either imidazole or tris as the buffering agent. (See buffer appendix). The copper ligand solutions were prepared from the same buffer batch used for dialysis of a given sample so that the solutions were otherwise identical. ($CuCl_2$ was added to the dialysis buffer and the pH carefully re-matched as necessary). All samples were degassed on ice for ten minutes before loading into the calorimeter. All ITC data were collected using a MCS Microcal calorimeter. All copper solutions were used in a control experiment titrating the copper solution into dialysis buffer in the absence of macromolecule in order to confirm the matching of solutions in the syringe and the sample cell. The heat of dilution was

subtracted from the experimental data by subtracting the control experiment in Origen. The ITC data were fit using Origin software. [194] The nonlinear least squares method was used to fit the binding isotherm and iterated until there was no further reduction in χ^2.

Appendix A

DNA SUBSTRATES

Table 6.1 DNA Substrates used in experiments in this work.

Sequence	*Length*	*Studies used in*
5'-TCGCGAATTCCG AGCGCTTAAGCG-5'	12bp	ESR, ITC
5'- TCGCAAATTCGCG GCGTTTAAGCGCT-5'	12bp	ESR
5'-GTGCCTTAAGCGCG CACGGAATTCGCGC-5'	14bp	ESR
5'-GGGCGGGCGCGAATTCGCGGGCGC CCCGCCGGCGCTTAAGCGCCCGCG-5'	24bp	Equilibrium Binding

Appendix B

PRIMERS USED FOR SITE-DIRECTED MUTAGENESIS

Table 6.2 Sequences of primers used for sites-directed mutagenesis.

Sequence [A]	Mutant
5'- GCCAAACACCAAGGTAAAGATATTATAAATATATGTAATGGTTTGTTAGTTGGG CGGTTTGTGGTTCCATTTCTATAATATTTATATACATTACCAAACAATCAACCC -5'	R123C
5'- GGTTTGTTAGTTGGGAAATGTGGAGATCAAGATTTAATGGCTGC CCAAACAATCAACCCTTTACACCTCTAGTTCTAAATTACCGACG-5'	R131C
5'-GGGGTCTAACTTTTTAACAGAAAATATCTGCATAACAAGACCAGATGGAAGGG . ….CCCCAGATTGAAAAATTGTCTTTTATAGACGTATTGTTCTGGTCTACCTTCCC-5'	S180C
5' GGGTTGTTAATCTTGAGTATAATTCTGGTTGTTTAAATAGGTTAGATCGACTAACTGC CCCAACAATTAGAACTCATATTAAGACCAACAAATTTATCCAATCTAGCTGATTGACG -5'	I197C
5' GATGGGAGGGAGTGGGATTCGTGTATCATGTTTGAAATAATGTTTG CTACCCTCCCTCACCCTAAGCACATAGTACAAACTTTATTACAAAC-5'	K249C

[A] Primer duplexes shown. Bases that are mismatched with the wild-type sequence to generate the mutant sequence are shown in red.

Appendix C

OTHER PRIMERS

Table 6.3 Sequences of non-mutagenic primers used in this work.

5’ ATGTCTAATAAAAAACAGTCAAATAGGCTAACTGAACAACATAAG	EcoRI-MBP amplification forward primer
5’CCGCTGAGGATCCTCACTTAGATGTAAGCTGTTCAAACAAGTCACGCCCC	EcoRI-MBP amplification reverse primer plus BamHI site in red
5’- GGTCGTCAGACTGTCGATGAAGCC	MBP-fusion confirmation primer. Binds 60bp upstream of XmnI cloning site

Appendix D

BUFFER RECIPIES

C18 Buffer: Buffer for reverse-phase purification of DNA over Alltech C-18 columns

- 100 mM Bis-Tris-Propane (BTP), pH 7.5 at 21°C
- 10 mM Triethylammonium actetate (TEA)
- 1mM ethylenediamenetetraacetate (EDTA)

Filter Buffer: Buffer for washing filters in equilibrium binding experiments

- 20mM cacodylic acid, pH 7.3 at 21°C
- 1mM EDTA
- 0.01% sodium azide
- pH and salt as listed for experiment

Binding Buffer: Reaction buffer for equilibrium binding experiments

- 20mM cacodylic acid, pH 7.3 at 21°C
- 1mM EDTA
- 0.01% sodium azide
- 100mg/ml bovine serum albumin
- 100μm dithiothreitol (DTT)
- pH and salt as listed for experiment

Cell Lysis Buffer:

- 30mM Na_2HPO_4, pH 7.8 at 21°C
- 700mM NaCl*
- 1mM EDTA
- 0.01% sodium azide
- 100µM DTT
- 10% glycerol

Protein Prep Buffer A:

- 20mM Na_2HPO_4, pH 7* at 21°C
- 300mM NaCl*
- 1mM EDTA
- 0.01% sodium azide
- 100µM DTT
- 10% glycerol
- *-optimum pH and salt vary
- Buffer is filter-sterilized and de-gassed until no bubbles appear

Protein Prep Buffer B:

- 20mM Na_2HPO_4, pH 7* at 21°C
- 1M NaCl*
- 1mM EDTA
- 0.01% sodium azide
- 100µM DTT
- 10% glycerol
- *-optimum pH varies
- Buffer is filter-sterilized and de-gassed until no bubbles appear

Protein Prep Buffer M (maltose):

- 20mM Na_2HPO_4, pH 7.3 at 21°C
- 350mM NaCl
- 10mM maltose
- 1mM EDTA
- 0.01% sodium azide
- 100µM DTT
- 10% glycerol
- Buffer is filter-sterilized and de-gassed until no bubbles appear

Protein Prep Ionic Adjustment Buffer:

- 20mM Na_2HPO_4, pH 7* at 21°C
- 1mM EDTA
- 0.01% sodium azide
- 100μM DTT
- 10% glycerol
- *pH to match that of sample being adjusted
- Buffer is filter-sterilized and de-gassed until no bubbles appear

ESR Ionic Adjustment Buffer:

- 20mM Na_2HPO_4, pH 7.3
- 1mM EDTA
- 0.01% sodium azide

ESR Spin Labeling/Early Stage Nitroxide Removal Buffer:

- 0.22M NaCl
- 20mM Na_2HPO_4, pH 7.3
- 1mM EDTA
- 0.01% sodium azide
- 10% glycerol
- 10% DMSO

ESR Exchange Buffer:

- 0.22M NaCl
- 20mM Na_2HPO_4, pH 7.3
- 1mM EDTA
- 0.01% sodium azide
- 10% glycerol

*Note: for deuterated samples prepare in 95% deuterated water and 10% D8 glycerol

Cu^{2+}-ESR Buffers:

Sample Dilution buffer:

- 30mM NEM pH 8
- 20um EDTA
- 0.01% sodium azide

Nitroxide Removal Buffer:

- 0.3M NH_4Cl
- 30mM NEM pH 8
- 20um EDTA
- 0.01% sodium azide
- 10% Dioxane

ITC Buffers:

Imidizole Buffer

- 10mM Imidazole pH 7.8
- 0.24 M NaCl

Tris Buffer

- 10mM Tris pH 7.8
- 0.24M NaCl

Appendix E

TWO-STAGE MUTAGENESIS PROTOCOL

Tube Setup:

Set up two sets of tubes, one containing forward primer, the other containing reverse primer. The primers are designed as described in the section above. Each set contains tubes with plasmid templates in amounts of 10ng, 20ng, 50ng, and one tube with 50ng of template but No Primer. This is the negative control for subsequent gel analysis, digestion, and transformation steps.

Each tube contains (add in this order):

- dd H_2O to 24µl
- 2.5µl of 10xPfu Buffer
- 1µl of 5mM dNTPs (Use a Fresh aliquot each time, do not freeze-thaw)
- 1µl of primer (at ~125ng/µl) Except No Primer Control!
- 1, 2, or 5 µl of template (at 10ng/µl)

Sample Setup:

Table 6.4 PCR setup for site-directed mutagenesis.

	10ng Forward*	20ng Forward	50ng Forward	No Primer
Buffer	2.5 µl	2.5 µl	2.5 µl	2.5 µl
dNTP	1 µl	1 µl	1 µl	1 µl
Primer	1 µl Forward	1 µl Forward	1 µl Forward	NONE
Template	1 µl	2 µl	5 µl	5 µl
dd H_2O	18.5 µl	17.5 µl	14.5 µl	15.5 µl

*Repeat entire setup with Reverse primer.

Spin down tubes briefly to settle contents and incubate on ice for 5 minutes, then add 0.75µl of PfuUltra immediately before starting PCR reaction. (Pfu polymerases have exonuclease activity and may degrade primers if added to the reaction too far ahead of time)

PCR Cycles:

Stage 1: Use the following 'default' conditions and optimize accordingly if needed based on gel analysis of the resulting DNA.

- Initial Melt: 95°C 1min
- 6 Cycles of:
  -Melt 95°C 1min
  -Anneal 60°C 1 min
  -Extend 70°C 15min
- Finally: Extend 70°C 15 minutes (to ensure completion of ongoing replication)
- Hold at 4°C

After completion set up Stage 2.

Stage 2: Spin down all reaction tubes briefly, then combine reaction mixture of 'partner' tubes (10ng F with 10ngR, no primer with no primer, etc). Then incubate on ice ~5min and add 1 µl of fresh PfuUltra using a positive displacement pipette.

Stage 2 Program:

- Initial Melt: 95°C 1min
- 20 Cycles of:
  -Melt 95°C 1min
  -Anneal 60°C 1 min
  -Extend 70°C 15min
- Finish: Extend 70°C 15 minutes (to finish any ongoing replication)
- Hold at 4°C

An aliquot of the resulting PCR products are analyzed over a 1% agarose gel, if clean product is observed (a solid band running as the pPS12 nicked circle, clear no primer controls) then the product is digested as described in the next section and transformed for screening.

Appendix F

MBP-FUSION PROTOCOL

Part A: Preparation of MBP plasmid

1) Prepare MBP plasmid using the standard Qiagen protocols on MBP-C4X (NEB) plasmid that has been transformed into *E. coli* DH5α. DO NOT use strains which have been pre-transformed with the methylase or any other plasmid, this will contaminate your prep. At the end of the Qiagen protocol elute into water, not elution buffer.

2) Digest with XmnI:

Reactions in .2ml PCR tube, each reaction:

- 5ul plasmid DNA
- 2ul NEB 2
- 13ul H20

Chill on ice, then add XmnI (0.75µl each individual tube), digest using the PCR program as follows:

- 35°Cx 1hr
- 65°Cx 25 minutes (heat kill)
- Hold 4°C

Check aliquot of plasmid over 1% agarose gel, it should be totally linearized.

3) Digestion with BamHI. BamHI is functional in the same buffer as XmnI but cannot be heat-killed. 1µl of BamHI is added to the reaction from the previous step, and incubated at 35°C for 1 hour. An additional tube should be set up as in the previous step to serve as a positive control for BamHI digestion and checked for linearization as above. The reaction should then be immediately (to stop the reaction) purified with a Wizard PCR cleanup kit and eluted in water. The plasmid is now ready for part C (ligation).

Part B: Cloning of EcoRI fragment

Steps-

1) PCR amplification
2) Gel Purification
3) T4 polymerase to blunt ends
4) Spin Purification
5) BamHI Digestion
6) Spin Purification
7) Ligation

1) PCR amplification of fragment:

The fragment is amplified using the iTaq kit, which provides high yields. Since the iTaq kit uses polymerases that leave overhangs, it is necessary to blunt the fragment ends prior to ligation (see T4 step). Add to each of 8 tubes of an iTaq strip: 1ul of plasmid prep (100ng), 10pmol (1ul) each of Forward and Reverse primer (Appendix C) 17ul sterile H20. Mix gently but thoroughly and run the following PCR cycle.

Cycle MBP-1:

- 95°C x 1minute
- 30 cycles of:
  -95°C x 1 minute
  -70°C x 1minute
  -72°C x 2 minutes

- 72°C x 5 minutes to finish extensions
- Hold at 4C

2) Gel-Purification of fragment: Samples are electrophoresed over a 0.7% agarose gel precast with 1x GelRed™ until the bromophenol blue dye front is approximately ¾ the way down the gel. The bands are visualized by UV shadowing and excised with a razor. The DNA fragments are then purified from the gel by use of a Wizard™ gel cleanup kit and eluted in 100ul water.

3) T4 Blunting of fragment:

Make up the following "master mix"

- 100ul of gel purified fragment (prepared as above)
- 20ul of 10x NEB 2
- 8 ul of 5mM DNTPs
- 2ul of BSA
- 70ul H20

Mix thoroughly, then incubate on ice (~15'). Also put an 8-well PCR strip on ice and turn on the PCR machine. When mix is ice cold, add 2ul of T4 polymerase to the master mix, mix gently but well, aliquot into cold PCR strip (25ul each tube) and start PCR program. For this protocol it is important to work quickly once the enzyme has been added.

Blunting PCR Program:

- 4°Cx30 seconds
- 12°Cx15minutes
- hold at 4°C

Immediately after the program finishes, all the reactions are quenched by pipetting (combine samples) into an Eppendorf tube containing 45ul of .1M EDTA. Vortex, then add 250ul of Membrane Binding Solution (Wizard Spin Kit) and Spin-Purify as per standard kit instructions.

4) BamHI Digestion of fragment:

"Master Mix"

- 72ul purified fragment
- 8ul 10x NEB 2
- 1ul BamHI

Incubate at 37°Cx1hr, purify immediately over wizard kit column as discussed previously. The whole "master mix" tube can be incubated in a 37C water bath, but I seem to get better results if I aliquot ~25ul each into PCR tubes and use the PCR machine cycle as described for BamHI in part A-3. The digested fragment is then purified via Wizard PCR cleanup kit.

Part C: Ligation of fragment

Per tube:

- 8ul of fragment
- 2ul of plasmid
- 10ul of 2x Quick ligation mix

Bring to room temperature, add 1ul Quick ligase. Incubate 5-10' at room temperature, then purify over spin column, working GENTLY with sample. Ligated plasmid is then transformed into DH5α.

PREPARATION OF RESINS FOR PROTEIN PURIFICATION

F.1 PREPARATION OF PHOSPHOCELLULOSE (P11) RESIN

For one prep (6-12L of cell culture) a 100ml resin volume is appropriate.

The P11 resin has a packing density of .17 dg/ml (dg = dry grams)

Therefore, a 100ml column requires 100ml* .17dg/ml = 17dg

However, quite a bit of resin is lost during the defining process, so make up at least* 1.5x the needed volume. (In this example, 17*1.5 = 25.5g)

For all steps, the phosphocellulose must be stirred gently to prevent generating fines. Fines are tiny bits of resin that pass through the filters into the HPLC lines and cause pressure problems in the AKTA system.

The following solutions are required for this protocol:

- 2L of 0.5M NaOH (Prepared in a 4L beaker)
- 2L of 0.5M HCl
- 2L of 0.5M Na2HPO4, pH adjusted to 7.0
- 2L of Protein Prep Buffer A (See Appendix)

Steps:

1) Weigh out the resin, and carefully stir into 2L of 0.5M NaOH and leave for exactly 5 minutes (Too long and the base will start to hydrolyze the resin)

2) Carefully decant the base solution, and re-suspend the resin in dH_2O. Let settle for ~5 minutes, then repeat (decant the resin, and re-suspend in dH_2O etc.) After 3x cycles of this, then check pH (by pH strip or use a beaker to collect some supernatant to use pH electrode). If the pH is below 11, decant H_2O and proceed to next step, otherwise repeat until pH is below 11.

3) Carefully add 2L of 0.5M HCl into the resin and leave for exactly 5 minutes. (Too long and the acid will start to hydrolyze the resin)

4) Carefully decant the resin, and re-suspend in dH_2O. Let settle for ~5 minutes, then repeat (decant the resin, and re-suspend in dH_2O etc.) After 3x cycles of this, then check pH (by pH strip or use a beaker to collect some supernatant to use pH electrode). If pH is above 3, decant H2O and proceed to next step, otherwise repeat until pH is above 3.

5) Add 1L 0.5 Na_2HPO_4, pH 7, mix gently, and leave for at least 10 minutes. Decant supernatant and repeat once.

6) Decant supernatant and re-suspend in ~3x volume of Buffer A. Leave for 5min, then check pH. (Decant supernatant and use electrode). If pH is not very close to that of Buffer A, repeat Buffer A exchanges the pH matches.

7) Transfer resin to a side-arm flask, adjust Buffer A to settled resin volume + 20%

8) De-Gas resin for 15 minutes and transfer to cold room.

F.2 PREPARATION OF BIOREX RESIN

For one prep (6-12L of cell culture) a 20ml resin volume is appropriate.

The BioRex 70 resin has a packing density of .5 dg/ml (dg = dry grams)

Therefore, a 20ml column is 20ml* .5 dg/ml = 10 dg

However, quite a bit is lost during the defining process, so make up at least* 2x what is needed. (In this example, 10g *2 = 20g)

For all steps, the resin must be stirred gently to prevent generating fines. Fines are tiny bits of resin that pass through the filters into the HPLC lines and cause pressure problems in the AKTA system.

The following solutions are required for this protocol:

- Protein Prep Buffer A (See Appendix X)
- Phosphoric acid for pH adjustment

Steps:

1) Suspend resin in 10x volumes of Buffer A and allow to equilibrate for 30 minutes

2) Adjust pH down to 7.0 with phosphoric acid, stirring gently

3) Equilibrate for 30 minutes, Measure pH, adjust to 7.0 if necessary

4) Repeat above step until pH is still 7.0 after 30 minute equilibration

5) Decant buffer and replace with fresh Buffer A. Repeat above steps until no pH change is noticed when fresh buffer is used.

Appendix G

PROTEIN QUANTIFICATION PROTOCOL

Protein Concentration Determination

In this protocol you will run a SDS gel comprising a protein standard of known concentration alongside a sample of unknown concentration and use optical density to calculate the concentration of the sample.

Part 1: Gel setup

- Materials:
  Protein standard of known concentration (provided sample is 50ng/ul in 1xSDS)
  Sample
  1x SDS loading buffer with Bromophenol Blue
  SDS Mini Gel (recommend 10 well commercial gradient gel of 4-15% for optimal results with EcoRI)
  Hamilton syringes (recommend a 10ul and a 20ul with flat-tip for gel loading)

Setting up the standard:

Note: The linear range for Comassie staining is ~50-500ng of protein/well.

Set up five tubes with known concentration across the linear range, using a 10ul Hamilton. Make sure that the standard is thoroughly thawed and mixed. I will generally use two different standard stock tubes, using one to set up the "odd" tubes and one for the "even". (This is not strictly necessary, but they do occasionally get contaminated and degrade or have other

problems- using two different tubes in this way makes it really obvious if this happens). The total volume for each tube will be 16ul, you will load ¾ (12ul) of this on the gel. This is necessary because you can never quantitatively retrieve the entire sample for loading. The configuration I usually use is this:

Table 6.5 Sample setup for protein standards.

Tube #	Amount 1xSDS	Amount 50ng/ul Std.	Total ng	Ng loaded
1	14ul	2ul	100	75
2	12ul	4ul	200	150
3	10ul	6ul	300	225
4	8ul	8ul	400	300
5	6ul	10ul	500	375

Mix gently by pipette and spin down. **Do NOT boil** samples prior to loading. It is important to run the gel right away (ie don't let them sit more than an hour) or evaporation will alter the protein concentration. This occurs even in tubes that have been sealed because the sample volume is so small.

Setting up the sample:

Set up 5 tubes of the sample, also with total volume 16u for 12ul loading volume in 1xSDS. You can do this two ways. 1) Initial "ball park" estimations you may want to try multiple dilutions- say a 1:20, 1:40, 1:60, 1:80 and 1:100 and set up each as per tube#3 above, for example. 2) When I think I am close to a true estimate I will confirm by making two dilutions (A and B) of the sample in 1xSDS to what should be 50ng/ul and set up exactly as for the standard using the odd/even pattern described above.

Important Tips: Do all dilutions using Hamiltons, directly into 1xSDS loading buffer (not water), mix gently and thoroughly by pipette, and if doing dilutions > 1:20 it is best to do serial dilutions. (ie: 1:10-> 1:10 instead of one 1:100).

Part 2: Running/Staining the Gel:

Load 12ul from each tube onto a 4-15% SDS acrylamide gel using a 20ul Hamilton with flat-tip. Use best gel-running behavior (wash wells, moderate voltage, etc). The sharper the bands are the easier your life will be in the next steps. Ideally the bands are sharpest when they wind up about halfway down the gradient, which for EcoRI is when the BPB is about ¾-4/5 down. Stain following my acrylamide gel-staining protocol following the steps for quantification gels.

Part 3: Optical Density

For our lab, ask someone to demonstrate the SPARK system or follow the SPARK system protocol. If not using the SPARK system- different scanners have different quirks but the final results should be the same. Once the gel is stained as in the gel staining protocol, measure the integrated optical density (IOD) for each band and analyze as below.

Part 4: Analysis

Plot out the IOD versus known ng for the standard. This should result in a nice straight line with an intercept near zero. If the intercept is a little high, this is usually the result of too much background staining and can be subtracted out by filling in the "Y int" field in the template with the Y intercept. Then enter the slope of the standard line into the "Std Slope" field. Also input the dilution factor for the sample (this is easy to modify for multiple dilutions as needed). This template will then automatically calculate a concentration estimate for the sample based on the average of data points and give you a standard deviation for your calculation. Note again that

the linear range is 50-500ng, any points outside that range (ng Sam field) should be deleted from the final column or they can skew your calculations. If all or most sample points are outside the correct range, go ahead and get an estimate from the average of all points, adjust your dilution factor accordingly, and re-iterate this protocol. I like to have at least three points in the correct range sometimes it takes a couple of iterations to get this to work out. I prefer to run at least two or three gels with the correct dilution and average the results for my final concentration estimate for samples which require a high degree of precision.

Appendix H

ITC DATA

Table 6.6- ITC Data-All imidazole curves.

Temp (°C)	Stoichiometry	K_d	ΔH_{obs} (kcal/mol)	$T\Delta S_{obs}$ (kcal/mol)
Imidazole				
10°C-A	1.88	3.95×10^{-6}	33.58	40.78
10°C-B	2.33	1.79×10^{-6}	18.10	24.44
10°C-C	2.29	1.45×10^{-5}	22.40	28.80
15°C-A	2.02	2.24×10^{-5}	21.60	27.72
15°C-B	2.04	2.53×10^{-5}	22.20	28.30
15°C-C	1.95	2.20×10^{-5}	22.00	28.21
15°C-D	1.99	2.25×10^{-5}	19.40	25.53
15°C-E	2.03	2.21×10^{-5}	19.00	25.18
21°-A	2.15	8.93×10^{-6}	41.40	48.24
21°-C	1.99	1.82×10^{-5}	27.20	34.12
21°-D	2.36	1.32×10^{-5}	15.00	21.56

Table 6.7 All ITC data- all Tris Buffer curves

Temp (°C)	Stoichiometry	K_d	ΔH_{obs} (kcal/mol)	$T\Delta S_{obs}$ (kcal/mol)
15°C-A	2.65	2.82E x10^{-5}	15.60	21.64
15°C-B	2.16	3.03 x10^{-5}	33.00	24.61
15°C-C	1.54	3.22 x10^{-5}	34.60	30.54

BIBLIOGRAPHY

[1] P. H. von Hippel, "From 'simple' DNA-protein interactions to the macromolecular machines of gene expression.," *Annual review of biophysics and biomolecular structure*, vol. 36, pp. 79-105, Jan. 2007.

[2] S. Harrison, "DNA Recognition By Proteins With The Helix-Turn-Helix Motif," *Annual Review of Biochemistry*, vol. 59, no. 1, pp. 933-969, Jan. 1990.

[3] A. Pingoud, M. Fuxreiter, V. Pingoud, and W. Wende, "Type II restriction endonucleases: structure and mechanism," *Cellular and molecular life sciences*, vol. 62, no. 6, pp. 685–707, 2005.

[4] J. J. Perona, "Type II restriction endonucleases.," *Methods (San Diego, Calif.)*, vol. 28, no. 3, pp. 353-64, Nov. 2002.

[5] A. Pingoud and A. Jeltsch, "Structure and function of type II restriction endonucleases," *Nucleic Acids Research*, vol. 29, no. 18, p. 3705, 2001.

[6] P. J. Sapienza, "Thermodynamic, kinetic, and structural basis for the relaxed dna sequence specificity of 'promiscuous' mutant EcoRI Endonucleases," University of Pittsburgh, 2005.

[7] L. Jen-Jacobson, "Protein-DNA recognition complexes: conservation of structure and binding energy in the transition state.," *Biopolymers*, vol. 44, no. 2, pp. 153-80, Jan. 1997.

[8] D. R. Lesser, M. R. Kurpiewski, and L. Jen-Jacobson, "The energetic basis of specificity in the Eco RI endonuclease--DNA interaction.," *Science (New York, N.Y.)*, vol. 250, no. 4982, pp. 776-86, Nov. 1990.

[9] R. J. Roberts, "How restriction enzymes became the workhorses of molecular biology," *PNAS*, vol. 102, no. 17, pp. 5905-5908, 2005.

[10] T. Bickle and D. Krüger, "Biology of DNA restriction.," *Microbiological reviews*, vol. 57, no. 2, pp. 434-50, Jun. 1993.

[11] R. J. Roberts, T. Vincze, J. Posfai, and D. Macelis, "REBASE--a database for DNA restriction and modification: enzymes, genes and genomes.," *Nucleic acids research*, vol. 38, no. Database issue, pp. D234-6, Jan. 2010.

[12] S.-H. Chan, B. L. Stoddard, and S.-Y. Xu, "Natural and engineered nicking endonucleases--from cleavage mechanism to engineering of strand-specificity.," *Nucleic acids research*, vol. 39, no. 1, pp. 1-18, Aug. 2010.

[13] L. Jen-Jacobson, L. E. Engler, and L. A. Jacobson, "Structural and thermodynamic strategies for site-specific DNA binding proteins.," *Structure (London, England : 1993)*, vol. 8, no. 10, pp. 1015-23, Oct. 2000.

[14] L. Jen-Jacobson, "Structural-perturbation approaches to thermodynamics of site-specific protein-DNA interactions.," *Methods in enzymology*, vol. 259, pp. 305-44, Jan. 1995.

[15] S. E. Halford and N. P. Johnson, "Single turnovers of the EcoRI restriction endonuclease.," *The Biochemical journal*, vol. 211, no. 2, pp. 405-15, May. 1983.

[16] D. A. Hiller, A. M. Rodriguez, and J. J. Perona, "Non-cognate enzyme-DNA complex: structural and kinetic analysis of EcoRV endonuclease bound to the EcoRI recognition site GAATTC.," *Journal of molecular biology*, vol. 354, no. 1, pp. 121-36, Nov. 2005.

[17] L. E. Engler, K. K. Welch, and L. Jen-Jacobson, "Specific binding by EcoRV endonuclease to its DNA recognition site GATATC.," *Journal of molecular biology*, vol. 269, no. 1, pp. 82-101, May. 1997.

[18] U. Siebenlist and W. Gilbert, "Contacts between Escherichia coli RNA polymerase and an early promoter of phage T7," *Proceedings of the National Academy of Sciences*, vol. 77, no. 1, p. 122, 1980.

[19] A. Grigorescu, M. Horvath, K. Willkoz, J. Chandrasekhar, and J. M. Rosenberg, "The Integration of Recognition and Cleavage: X-Ray Structures of Pre-Transition State Complex, Post-Reactive Complex, and the DNA-Free Endonuclease," in *Restriction Endonucleases*, Springer-Verlag, 2004.

[20] M. R. Kurpiewski et al., "Mechanisms of coupling between DNA recognition specificity and catalysis in EcoRI endonuclease.," *Structure*, vol. 12, no. 10, pp. 1775-88, Oct. 2004.

[21] J. D. McGhee and P. H. von Hippel, "Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice.," *Journal of molecular biology*, vol. 86, no. 2, pp. 469-89, Jun. 1974.

[22] G. M. Clore, A. M. Gronenborn, and R. W. Davies, "Cooperative non-specific DNA binding of the N-terminal core of the cyclic AMP receptor protein of Escherichia coli and its modulation by cyclic AMP.," *FEBS letters*, vol. 164, no. 1, pp. 57-62, Nov. 1983.

[23] L. Jen-Jacobson and L. A. Jacobson, "Role of Water and Effects of Small Ions in Site-specific Protein-DNA Interactions," in *Structural Biology of Protein-Nucleic Acid Interaction*, 2008, pp. 13-48.

[24] J. H. Ha, R. S. Spolar, and M. T. Record, "Role of the hydrophobic effect in stability of site-specific protein-DNA complexes.," *Journal of molecular biology*, vol. 209, no. 4, pp. 801-16, Oct. 1989.

[25] R. S. Spolar and M. Record, "Coupling of local folding to site-specific binding of proteins to DNA," *Science*, vol. 263, no. 5148, p. 777, 1994.

[26] L. Jen-Jacobson, L. E. Engler, J. Ames, M. R. Kurpiewski, and A. Grigorescu, "Thermodynamic Parameters of Specific and Nonspecific Protein-DNA Binding," *Supramolecular Chemistry*, vol. 12, no. 2, pp. 143-160, Oct. 2000.

[27] N. V. Prabhu and K. A. Sharp, "Heat capacity in proteins.," *Annual review of physical chemistry*, vol. 56, pp. 521-48, Jan. 2005.

[28] J. M. Sturtevant, "Heat Capacity and Entropy Changes in Processes Involving Proteins," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 74, no. 6, pp. 2236-2240, 1977.

[29] R. S. Spolar, J. R. Livingstone, and M. T. Record, "Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water.," *Biochemistry*, vol. 31, no. 16, pp. 3947-55, Apr. 1992.

[30] K. P. Murphy and E. Freire, "Thermodynamics of structural stability and cooperative folding behavior in proteins.," *Advances in protein chemistry*, vol. 43, pp. 313-61, Jan. 1992.

[31] G. I. Makhatadze and P. L. Privalov, "Energetics of protein structure.," *Advances in protein chemistry*, vol. 47, pp. 307-425, Jan. 1995.

[32] A. D. Robertson and K. P. Murphy, "Protein Structure and the Energetics of Protein Stability.," *Chemical reviews*, vol. 97, no. 5, pp. 1251-1268, Aug. 1997.

[33] S. P. Hancock, "The role of phosphate neutralization in EcoRV-induced DNA bending," University of Pittsburgh, 2009.

[34] Y. Takeda, P. D. Ross, and C. P. Mudd, "Thermodynamics of Cro protein-DNA interactions.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 17, pp. 8180-4, Sep. 1992.

[35] E. Merabet and G. K. Ackers, "Calorimetric analysis of lambda cI repressor binding to DNA operator sites.," *Biochemistry*, vol. 34, no. 27, pp. 8554-63, Jul. 1995.

[36] J. E. Ladbury, J. G. Wright, J. M. Sturtevant, and P. B. Sigler, "A Thermodynamic Study of the trp Repressor–Operator Interaction," *Journal of molecular biology*, vol. 238, no. 5, pp. 669–681, 1994.

[37] V. Petri, M. Hsieh, and M. Brenowitz, "Thermodynamic and kinetic characterization of the binding of the TATA binding protein to the adenovirus E4 promoter.," *Biochemistry*, vol. 34, no. 31, pp. 9977-84, Aug. 1995.

[38] C. Berger, I. Jelesarov, and H. R. Bosshard, "Coupled folding and site-specific binding of the GCN4-bZIP transcription factor to the AP-1 and ATF/CREB DNA sites studied by microcalorimetry.," *Biochemistry*, vol. 35, no. 47, pp. 14984-91, Nov. 1996.

[39] T. Lundbaeck, C. Cairns, J. A. Gustafsson, J. Carlstedt-Duke, and T. Haerd, "Thermodynamics of the glucocorticoid receptor-DNA interaction: binding of wild-type GR DBD to different response elements," *Biochemistry*, vol. 32, no. 19, pp. 5074–5082, 1993.

[40] C. J. Morton and J. E. Ladbury, "Water-mediated protein-DNA interactions: the relationship of thermodynamics to structural detail.," *Protein science : a publication of the Protein Society*, vol. 5, no. 10, pp. 2115-8, Oct. 1996.

[41] M. Karplus, T. Ichiye, and B. M. Pertirr, "Configurational Entropy of Native Proteins," *Biophysical Journal*, vol. 52, no. December, pp. 1083-1085, 1987.

[42] B. Brooks and M. Karplus, "Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 80, no. 21, pp. 6571-5, Nov. 1983.

[43] C. L. Brooks, M. Karplus, and B. M. Pettitt, *Proteins: A theoretical perspective of dynamics, structure, and thermodynamics*. John Wiley and Sons, 1988, p. 274.

[44] F. K. Winkler et al., "The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments.," *The EMBO Journal*, vol. 12, no. 5, p. 1781, 1993.

[45] B. Luisi, W. X. Xu, Z. Otwinowski, L. Freedman, K. Yamamoto, and P. B. Sigler, "Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA," *Nature*, vol. 352, pp. 497-505, 1991.

[46] Y. Duan, P. Wilkosz, and J. M. Rosenberg, "Dynamic contributions to the DNA binding entropy of the EcoRI and EcoRV restriction endonucleases.," *Journal of molecular biology*, vol. 264, no. 3, pp. 546-55, Dec. 1996.

[47] A. Grigorescu, "Structural and Energetic Determinants of the DNA Binding Specificity of EcoRI Endonuclease," University of Pittsburgh, 2003.

[48] M. R. Kurpiewski, M. Koziolkiewicz, A. Wilk, W. J. Stec, and L. Jen-Jacobson, "Chiral phosphorothioates as probes of protein interactions with individual DNA phosphoryl oxygens: essential interactions of EcoRI endonuclease with the phosphate at pGAATTC.," *Biochemistry*, vol. 35, no. 27, pp. 8846-54, Jul. 1996.

[49] J. L. Leroy, E. Charretier, M. Kochoyan, and M. Gueron, "Evidence from base-pair kinetics for two types of adenine tract structures in solution: their relation to DNA curvature," *Biochemistry*, vol. 27, no. 25, pp. 8894-8898, Dec. 1988.

[50] M. R. Eftink, A. C. Anusiem, and R. L. Biltonen, "Enthalpy-entropy compensation and heat capacity changes for protein-ligand interactions: general thermodynamic models and data for the binding of nucleotides to ribonuclease A.," *Biochemistry*, vol. 22, no. 16, pp. 3884-96, Aug. 1983.

[51] T. M. Lohman, P. L. deHaseth, and M. T. Record, "Pentalysine-deoxyribonucleic acid interactions: a model for the general effects of ion concentrations on the interactions of proteins with nucleic acids.," *Biochemistry*, vol. 19, no. 15, pp. 3522-30, Jul. 1980.

[52] J. A. McClarin et al., "Structure of the DNA-Eco RI endonuclease recognition complex at 3 A resolution.," *Science (New York, N.Y.)*, vol. 234, no. 4783, pp. 1526-41, Dec. 1986.

[53] Y. C. Kim, J. C. Grable, R. Love, P. J. Greene, and J. M. Rosenberg, "Refinement of Eco RI endonuclease crystal structure: a revised protein chain tracing.," *Science (New York, N.Y.)*, vol. 249, no. 4974, pp. 1307-9, Sep. 1990.

[54] R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig, and R. S. Mann, "Origins of specificity in protein-DNA recognition.," *Annual review of biochemistry*, vol. 79, pp. 233-69, Jan. 2010.

[55] L. E. Engler, P. Sapienza, L. F. Dorner, R. Kucera, I. Schildkraut, and L. Jen-Jacobson, "The energetics of the interaction of BamHI endonuclease with its recognition site GGATCC1," *Journal of molecular biology*, vol. 307, no. 2, pp. 619–636, 2001.

[56] S. P. Hancock, D. A. Hiller, J. J. Perona, and L. Jen-Jacobson, "The energetic contribution of induced electrostatic asymmetry to DNA bending by a site-specific protein.," *Journal of molecular biology*, vol. 406, no. 2, pp. 285-312, Feb. 2011.

[57] Z. Otwinowski et al., "Crystal structure of trp represser/operator complex at atomic resolution," *Nature*, vol. 355, pp. 321-329, 1988.

[58] D. R. Lesser, M. R. Kurpiewski, T. Waters, B. A. Connolly, and L. Jen-Jacobson, "Facilitated distortion of the DNA site enhances EcoRI endonuclease-DNA recognition.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 16, pp. 7548-52, Aug. 1993.

[59] R. E. Dickerson and T. K. Chiu, "Helix bending as a factor in protein/DNA recognition.," *Biopolymers*, vol. 44, no. 4, pp. 361-403, Jan. 1997.

[60] W. K. Olson, A. A. Gorin, X. J. Lu, L. M. Hock, and V. B. Zhurkin, "DNA sequence-dependent deformability deduced from protein–DNA crystal complexes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 19, p. 11163, 1998.

[61] S. Schultz, G. Sheilds, and T. Steitz, "Crystal Structure of a CAP-DNA complex: The DNA is bent by 90," *Science*, vol. 1, no. 3, pp. 235-246, Nov. 1991.

[62] G. Parkinson, "Structure of the CAP-DNA Complex at 2.5 Å Resolution: A Complete Picture of the Protein-DNA Interface," *Journal of Molecular Biology*, vol. 260, no. 3, pp. 395-408, Jul. 1996.

[63] J. L. Kim, D. B. Nikolov, and S. K. Burley, "Co-crystal structure of TBP recognizing the minor groove of a TATA element," *Nature*, vol. 365, pp. 520-527, 1993.

[64] U. Siebenlist, R. B. Simpson, and W. Gilbert, "E. coli RNA polymerase interacts homologously with two different promoters," *Cell*, vol. 20, no. 2, pp. 269–281, 1980.

[65] S. Jones, P. van Heyningen, H. M. Berman, and J. M. Thornton, "Protein-DNA interactions: A structural analysis.," *Journal of molecular biology*, vol. 287, no. 5, pp. 877-96, Apr. 1999.

[66] C. J. Tsai, B. Ma, Y. Y. Sham, S. Kumar, and R. Nussinov, "Structured disorder and conformational selection.," *Proteins*, vol. 44, no. 4, pp. 418-27, Sep. 2001.

[67] C. Lee, S. H. Park, M. Y. Lee, and M. H. Yu, "Regulation of protein function by native metastability.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 14, pp. 7727-31, Jul. 2000.

[68] J. Heitman and P. Model, "Substrate recognition by the EcoRI endonuclease.," *Proteins*, vol. 7, no. 2, pp. 185-97, Jan. 1990.

[69] P. J. Sapienza, C. A. Dela Torre, W. H. McCoy, S. V. Jana, and L. Jen-Jacobson, "Thermodynamic and kinetic basis for the relaxed DNA sequence specificity of 'promiscuous' mutant EcoRI endonucleases.," *Journal of molecular biology*, vol. 348, no. 2, pp. 307-24, Apr. 2005.

[70] P. J. Sapienza, J. M. Rosenberg, and L. Jen-Jacobson, "Structural and thermodynamic basis for enhanced DNA binding by a promiscuous mutant EcoRI endonuclease.," *Structure*, vol. 15, no. 11, pp. 1368-82, Nov. 2007.

[71] S. A. Townson, J. C. Samuelson, Y. Bao, S.-Y. Xu, and A. K. Aggarwal, "BstYI bound to noncognate DNA reveals a 'hemispecific' complex: implications for DNA scanning.," *Structure (London, England : 1993)*, vol. 15, no. 4, pp. 449-59, Apr. 2007.

[72] H. Viadiu and A. K. Aggarwal, “Structure of BamHI bound to nonspecific DNA: a model for DNA sliding.,” *Molecular cell*, vol. 5, no. 5, pp. 889-95, May. 2000.

[73] R. A. Albright, M. C. Mossing, and B. W. Matthews, “Crystal structure of an engineered Cro monomer bound nonspecifically to DNA: possible implications for nonspecific binding by the wild-type protein.,” *Protein science : a publication of the Protein Society*, vol. 7, no. 7, pp. 1485-94, Jul. 1998.

[74] C. G. Kalodimos et al., “Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes.,” *Science (New York, N.Y.)*, vol. 305, no. 5682, pp. 386-9, Jul. 2004.

[75] J. R. Horton, K. Liebert, S. Hattman, A. Jeltsch, and X. Cheng, “Transition from nonspecific to specific DNA interactions along the substrate-recognition pathway of dam methyltransferase,” *Cell*, vol. 121, no. 3, pp. 349–361, 2005.

[76] R. A. Albright and B. W. Matthews, “How Cro and lambda-repressor distinguish between operators: the structural basis underlying a genetic switch.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 7, pp. 3431-6, Mar. 1998.

[77] C. Surridge, “Advancing in a dynamic field,” *Nature structural & molecular biology*, vol. 1, no. 10, pp. 664-666, 1994.

[78] L. Columbus and W. L. Hubbell, “A new spin on protein dynamics.,” *Trends in biochemical sciences*, vol. 27, no. 6, pp. 288-95, Jun. 2002.

[79] A. K. Gardino and D. Kern, “Functional dynamics of response regulators using NMR relaxation techniques.,” *Methods in enzymology*, vol. 423, no. 7, pp. 149-65, Jan. 2007.

[80] B. F. Volkman, D. Lipson, D. E. Wemmer, and D. Kern, “Two-state allosteric behavior in a single-domain signaling protein.,” *Science (New York, N.Y.)*, vol. 291, no. 5512, pp. 2429-33, Mar. 2001.

[81] R. Ishima and D. A. Torchia, “Protein dynamics from NMR.,” *Nature structural biology*, vol. 7, no. 9, pp. 740-3, Sep. 2000.

[82] D. Kern, E. Z. Eisenmesser, and M. Wolf-Watz, “Enzyme dynamics during catalysis measured by NMR spectroscopy.,” *Methods in enzymology*, vol. 394, pp. 507-24, Jan. 2005.

[83] C. Altenbach, A. K. Kusnetzow, O. P. Ernst, K. P. Hofmann, and W. L. Hubbell, “High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 21, pp. 7439-44, May. 2008.

[84] W. L. Hubbell, C. Altenbach, C. M. Hubbell, and H. G. Khorana, "Rhodopsin structure, dynamics, and activation: a perspective from crystallography, site-directed spin labeling, sulfhydryl reactivity, and disulfide cross-linking.," *Advances in protein chemistry*, vol. 63, pp. 243-90, Jan. 2003.

[85] H. S. Mchaourab, K. J. Oh, C. J. Fang, and W. L. Hubbell, "Conformation of T4 lysozyme in solution. Hinge-bending motion and the substrate-induced conformational transition studied by site-directed spin labeling.," *Biochemistry*, vol. 36, no. 2, pp. 307-16, Jan. 1997.

[86] H. S. Mchaourab, M. A. Lietzow, K. Hideg, and W. L. Hubbell, "Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics.," *Biochemistry*, vol. 35, no. 24, pp. 7692-704, Jun. 1996.

[87] M. Karplus, T. Ichiye, B. M. Pettitt, and B. M. Pertirr, "Configurational Entropy of Native Proteins," *Biophysical journal*, vol. 52, no. 6, pp. 1083-5, Dec. 1987.

[88] A. Cooper, C. M. Johnson, J. H. Lakey, and M. Nöllmann, "Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions.," *Biophysical chemistry*, vol. 93, no. 2-3, pp. 215-30, Nov. 2001.

[89] J. A. McCammon and S. C. Harvey, *Dynamics of proteins and nucleic acids*. Cambridge University Press, 1988, p. 248.

[90] K. K. Frederick, M. S. Marlow, K. G. Valentine, and A. J. Wand, "Conformational entropy in molecular recognition by proteins.," *Nature*, vol. 448, no. 7151, pp. 325-9, Jul. 2007.

[91] P. J. Sapienza and A. L. Lee, "Using NMR to study fast dynamics in proteins: methods and applications.," *Current opinion in pharmacology*, vol. 10, no. 6, pp. 723-730, Oct. 2010.

[92] J. R. Bolton, "Electron Spin Resonance Theory," in *Biological Applications of Electron Spin Resonance*, 1972, pp. 11-60.

[93] J. Strancar, "Advanced ESR spectroscopy in membrane biophysics," in *ESR spectroscopy in membrane biophysics*, Springer, 2007, pp. 49–93.

[94] C. S. Klug and J. B. Feix, "Methods and Applications of Site-Directed Spin Labeling EPR Spectroscopy," *Methods in Cell Biology*, vol. 84, no. 7, pp. 617-658, 2008.

[95] J. R. Bolton, D. C. Borg, and H. M. Swartz, "Experimental Aspects of Biological Electron Spin Resonance Studies," in *Biological Applications of Electron Spin Resonance*, 1972, pp. 64-118.

[96] F. Tombolato, A. Ferrarini, and J. H. Freed, "Dynamics of the nitroxide side chain in spin-labeled proteins.," *The journal of physical chemistry. B*, vol. 110, no. 51, pp. 26248-59, Dec. 2006.

[97] P. G. Fajer, L. Brown, and L. Song, "Practical pulsed dipolar ESR (DEER)," in *ESR spectroscopy in membrane biophysics*, Springer, 2007, pp. 95–128.

[98] G. E. Fanucci and D. S. Cafiso, "Recent advances and applications of site-directed spin labeling.," *Current opinion in structural biology*, vol. 16, no. 5, pp. 644-53, Oct. 2006.

[99] W. L. Hubbell, H. S. Mchaourab, C. Altenbach, and M. A. Lietzow, "Watching proteins move using site-directed spin labeling," *Structure*, vol. 4, no. 7, pp. 779-783, Jul. 1996.

[100] W. L. Hubbell, D. S. Cafiso, and C. Altenbach, "Identifying conformational changes with site-directed spin labeling.," *Nature structural biology*, vol. 7, no. 9, pp. 735-9, Sep. 2000.

[101] R. Biswas, H. Kühne, G. W. Brudvig, and V. Gopalan, "Use of EPR spectroscopy to study macromolecular structure and function.," *Science progress*, vol. 84, no. 1, pp. 45-67, Jan. 2001.

[102] K. M. Stone, J. E. Townsend, J. Sarver, P. J. Sapienza, S. Saxena, and L. Jen-Jacobson, "Electron spin resonance shows common structural features for different classes of EcoRI-DNA complexes.," *Angewandte Chemie (International ed. in English)*, vol. 47, no. 52, pp. 10192-4, Jan. 2008.

[103] A. P. Todd, J. Cong, F. Levinthal, C. Levinthal, and W. L. Hubbell, "Site-directed mutagenesis of colicin E1 provides specific attachment sites for spin labels whose spectra are sensitive to local conformation.," *Proteins*, vol. 6, no. 3, pp. 294-305, Jan. 1989.

[104] C. Altenbach, K. J. Oh, R. J. Trabanino, K. Hideg, and W. L. Hubbell, "Estimation of inter-residue distances in spin labeled proteins at physiological temperatures: experimental strategies and practical limitations.," *Biochemistry*, vol. 40, no. 51, pp. 15471-82, Dec. 2001.

[105] W. L. Hubbell, A. Gross, R. Langen, and M. A. Lietzow, "Recent Advances in Site-Directed Spin labeling of proteins," *Current Opinion in Structural Biology*, pp. 649-656, 1998.

[106] L. J. Berliner, J. Grunwald, H. O. Hankovszky, and K. Hideg, "A novel reversible thiol-specific spin label: Papain active site labeling and inhibition• 1," *Analytical Biochemistry*, vol. 119, no. 2, pp. 450–455, 1982.

[107] L. Columbus and W. L. Hubbell, "Mapping backbone dynamics in solution with site-directed spin labeling: GCN4-58 bZip free and bound to DNA.," *Biochemistry*, vol. 43, no. 23, pp. 7273-87, Jun. 2004.

[108] E. Perozo, D. M. Cortes, P. Sompornpisut, A. Kloda, and B. Martinac, "Open channel structure of MscL and the gating mechanism of mechanosensitive channels.," *Nature*, vol. 418, no. 6901, pp. 942-8, Aug. 2002.

[109] E. Perozo, D. M. Cortes, and L. G. Cuello, "Three-dimensional architecture and gating mechanism of a K+ channel studied by EPR spectroscopy," *Nature Structural & Molecular Biology*, vol. 5, no. 6, pp. 459–469, 1998.

[110] A. D. Milov, A. Ponomarev, and Y. D. Tsvetkov, "Electron-electron double resonance in electron spin echo: model biradical systems and the sensitized photolysis of decalin," *Chemical physics letters*, vol. 110, no. 1, pp. 67–72, 1984.

[111] G. Jeschke, "Distance measurements in the nanometer range by pulse EPR.," *Chemphyschem : a European journal of chemical physics and physical chemistry*, vol. 3, no. 11, pp. 927-32, Nov. 2002.

[112] P. G. Fajer, "Site directed spin labelling and pulsed dipolar electron paramagnetic resonance (double electron–electron resonance) of force activation in muscle," *Journal of Physics: Condensed Matter*, vol. 17, no. 18, p. S1459-S1469, May. 2005.

[113] B. E. Bode, D. Margraf, J. Plackmeyer, G. Dürner, T. F. Prisner, and O. Schiemann, "Counting the monomers in nanometer-sized oligomers by pulsed electron-electron double resonance.," *Journal of the American Chemical Society*, vol. 129, no. 21, pp. 6736-45, May. 2007.

[114] G. Jeschke, "DeerAnalysis2009 User Manual." pp. 1-38, 2009.

[115] Y.-W. Chiang, P. P. Borbat, and J. H. Freed, "The determination of pair distance distributions by pulsed ESR using Tikhonov regularization.," *Journal of magnetic resonance (San Diego, Calif. : 1997)*, vol. 172, no. 2, pp. 279-95, Feb. 2005.

[116] K. M. Stone, J. E. Townsend, J. Sarver, P. J. Sapienza, S. Saxena, and L. Jen-Jacobson, "Electron spin resonance shows common structural features for different classes of EcoRI-DNA complexes- Supporting Information.," *Angewandte Chemie (International ed. in English)*, vol. 47, no. 52, pp. 10192-4, Jan. 2008.

[117] S. Steigmiller, M. Börsch, P. Gräber, and M. Huber, "Distances between the b-subunits in the tether domain of F(0)F(1)-ATP synthase from E. coli.," *Biochimica et biophysica acta*, vol. 1708, no. 2, pp. 143-53, Jun. 2005.

[118] I. Smirnova, V. Kasho, J.-Y. Choe, C. Altenbach, W. L. Hubbell, and H. R. Kaback, "Sugar binding induces an outward facing conformation of LacY.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 42, pp. 16504-9, Oct. 2007.

[119] D. Hilger et al., "Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na+/H+ antiporter of E. coli.," *Biophysical journal*, vol. 89, no. 2, pp. 1328-38, Aug. 2005.

[120] D. Hilger, Y. Polyhach, E. Padan, H. Jung, and G. Jeschke, "High-resolution structure of a Na+/H+ antiporter dimer obtained by pulsed electron paramagnetic resonance distance measurements.," *Biophysical journal*, vol. 93, no. 10, pp. 3675-83, Nov. 2007.

[121] J. E. Banham, C. R. Timmel, R. J. M. Abbott, S. M. Lea, and G. Jeschke, "The characterization of weak protein-protein interactions: evidence from DEER for the trimerization of a von Willebrand Factor A domain in solution.," *Angewandte Chemie (International ed. in English)*, vol. 45, no. 7, pp. 1058-61, Feb. 2006.

[122] M. Bennati, J. H. Robblee, V. Mugnaini, J. Stubbe, J. H. Freed, and P. P. Borbat, "EPR distance measurements support a model for long-range radical initiation in E. coli ribonucleotide reductase.," *Journal of the American Chemical Society*, vol. 127, no. 43, pp. 15014-5, Nov. 2005.

[123] N. Piton, Y. Mu, G. Stock, T. F. Prisner, O. Schiemann, and J. W. Engels, "Base-specific spin-labeling of RNA for structure determination.," *Nucleic acids research*, vol. 35, no. 9, pp. 3128-43, Jan. 2007.

[124] O. Schiemann, N. Piton, Y. Mu, G. Stock, J. W. Engels, and T. F. Prisner, "A PELDOR-based nanometer distance ruler for oligonucleotides.," *Journal of the American Chemical Society*, vol. 126, no. 18, pp. 5722-9, May. 2004.

[125] G. Sicoli, G. Mathis, O. Delalande, Y. Boulard, D. Gasparutto, and S. Gambarelli, "Double Electron–Electron Resonance (DEER): A Convenient Method To Probe DNA Conformational Changes," *Angewandte Chemie*, vol. 120, no. 4, pp. 747-749, Jan. 2008.

[126] A. Godt, M. Schulte, H. Zimmermann, and G. Jeschke, "How Flexible Are Poly(para-phenyleneethynylene)s?," *Angewandte Chemie*, vol. 118, no. 45, pp. 7722-7726, Nov. 2006.

[127] S. Pornsuwan, G. Bird, C. E. Schafmeister, and S. Saxena, "Flexibility and lengths of bis-peptide nanostructures by electron spin resonance.," *Journal of the American Chemical Society*, vol. 128, no. 12, pp. 3876-7, Mar. 2006.

[128] K. Sugata, M. Nakamura, S. Ueki, P. G. Fajer, and T. Arata, "ESR reveals the mobility of the neck linker in dimeric kinesin," *Biochemical and Biophysical Research Communications*, vol. 314, no. 2, pp. 447-451, Feb. 2004.

[129] K. Sugata, L. Song, M. Nakamura, S. Ueki, P. G. Fajer, and T. Arata, "Nucleotide-induced flexibility change in neck linkers of dimeric kinesin as detected by distance Measurements using spin-labeling EPR.," *Journal of molecular biology*, vol. 386, no. 3, pp. 626-36, Feb. 2009.

[130] P. G. Fajer, M. Gyimesi, A. Málnási-Csizmadia, C. R. Bagshaw, K. I. Sen, and L. Song, "Myosin cleft closure by double electron–electron resonance and dipolar EPR," *Journal of Physics: Condensed Matter*, vol. 19, no. 28, p. 285208, Jul. 2007.

[131] J. C. Klein et al., "Actin-binding cleft closure in myosin II probed by site-directed spin labeling and pulsed EPR.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 35, pp. 12867-72, Sep. 2008.

[132] K. I. Sen, T. M. Logan, and P. G. Fajer, "Protein dynamics and monomer-monomer interactions in AntR activation by electron paramagnetic resonance and double electron-electron resonance.," *Biochemistry*, vol. 46, no. 41, pp. 11639-49, Oct. 2007.

[133] K. M. Stone, "Structural Insights into Oligomeric Protein Complexes by Electron Spin Resonance," University of Pittsburgh, 2009.

[134] G. Jeschke, A. Koch, U. Jonas, and A. Godt, "Direct conversion of EPR dipolar time evolution data to distance distributions.," *Journal of magnetic resonance (San Diego, Calif. : 1997)*, vol. 155, no. 1, pp. 72-82, Mar. 2002.

[135] G. Jeschke, A. Bender, H. Paulsen, H. Zimmermann, and A. Godt, "Sensitivity enhancement in pulse EPR distance measurements.," *Journal of magnetic resonance (San Diego, Calif. : 1997)*, vol. 169, no. 1, pp. 1-12, Jul. 2004.

[136] L. Jen-Jacobson et al., "Coordinate ion pair formation between EcoRI endonuclease and DNA.," *The Journal of biological chemistry*, vol. 258, no. 23, pp. 14638-46, Dec. 1983.

[137] M. Pannier, S. Veit, A. Godt, G. Jeschke, and H. W. Spiess, "Dead-time free measurement of dipole-dipole interactions between electron spins.," *Journal of magnetic resonance (San Diego, Calif. : 1997)*, vol. 142, no. 2, pp. 331-40, Feb. 2000.

[138] G. Jeschke et al., "Applied Magnetic Resonance Analyzing Pulsed ELDOR Data," *Applied Magnetic Resonance*, vol. 498, pp. 473-498, 2006.

[139] L. Columbus, T. Kálai, J. Jekö, K. Hideg, and W. L. Hubbell, "Molecular motion of spin labeled side chains in alpha-helices: analysis by variation of side chain structure.," *Biochemistry*, vol. 40, no. 13, pp. 3828-46, Apr. 2001.

[140] K. Sale, C. Sar, K. A. Sharp, K. Hideg, and P. G. Fajer, "Structural Determination of Spin Label Immobilization and Orientation: A Monte Carlo Minimization Approach," *Journal of Magnetic Resonance*, vol. 156, no. 1, pp. 104-112, May. 2002.

[141] K. Sale, L. Song, Y. Liu, and E. Perozo, "Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER," *Journal of the American*, vol. 127, no. 26, pp. 9334-5, Jul. 2005.

[142] B. J. Terry, W. E. Jack, and P. Modrich, “Facilitated diffusion during catalysis by EcoRI endonuclease. Nonspecific interactions in EcoRI catalysis.,” *The Journal of biological chemistry*, vol. 260, no. 24, pp. 13130-7, Oct. 1985.

[143] H. J. Ehbrecht, A. Pingoud, C. Urbanke, G. Maass, and C. Gualerzi, “Linear diffusion of restriction endonucleases on DNA.,” *The Journal of biological chemistry*, vol. 260, no. 10, pp. 6160-6, May. 1985.

[144] D. J. Wright, W. E. Jack, and P. Modrich, “The kinetic mechanism of EcoRI endonuclease.,” *The Journal of biological chemistry*, vol. 274, no. 45, pp. 31896-902, Nov. 1999.

[145] R. B. Winter, O. G. Berg, and P. H. von Hippel, “Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor--operator interaction: kinetic measurements and conclusions.,” *Biochemistry*, vol. 20, no. 24, pp. 6961-77, Nov. 1981.

[146] S. E. Halford and J. F. Marko, “How do site-specific DNA-binding proteins find their targets?,” *Nucleic acids research*, vol. 32, no. 10, pp. 3040-52, Jan. 2004.

[147] O. G. Berg, R. B. Winter, and P. H. von Hippel, “Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory.,” *Biochemistry*, vol. 20, no. 24, pp. 6929-48, Nov. 1981.

[148] M. Karplus and J. Kuriyan, “Molecular dynamics and protein function.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 19, pp. 6679-85, May. 2005.

[149] F. Parak and H. Frauenfelder, “Protein dynamics,” *Physica A: Statistical Mechanics and its Applications*, vol. 201, no. 1-3, pp. 332-345, Dec. 1993.

[150] T. Prisner, M. Rohrer, and F. MacMillan, “Pulsed EPR spectroscopy: biological applications.,” *Annual review of physical chemistry*, vol. 52, pp. 279-313, Jan. 2001.

[151] D. C. Harris and M. D. Bertolucci, *Symmetry and spectroscopy: an introduction to vibrational and electronic spectroscopy*. Courier Dover Publications, 1989, p. 550.

[152] T. I. Smirnova and A. I. Smirnov, “High-field ESR spectroscopy in membrane and protein biophysics,” *ESR Spectroscopy in Membrane Biophysics*, pp. 165–251, 2007.

[153] S. Stoll and A. Schweiger, “EasySpin, a comprehensive software package for spectral simulation and analysis in EPR.,” *Journal of magnetic resonance (San Diego, Calif. : 1997)*, vol. 178, no. 1, pp. 42-55, Jan. 2006.

[154] K. Mobius, A. Schnegg, M. Plato, M. Fuchs, and A. Savitsky, "High-Field EPR Spectroscopy on Transfer Proteins in Biological Action," in *XXI International Meeting on Radio and Microwave Spectroscopy*, 2005, vol. 108, no. 2, pp. 215-234.

[155] J. P. Barnes, Z. Liang, H. S. Mchaourab, J. H. Freed, and W. L. Hubbell, "A multifrequency electron spin resonance study of T4 lysozyme dynamics.," *Biophysical journal*, vol. 76, no. 6, pp. 3298-306, Jun. 1999.

[156] M. a Lietzow and W. L. Hubbell, "Motion of spin label side chains in cellular retinol-binding protein: correlation with structure and nearest-neighbor interactions in an antiparallel beta-sheet.," *Biochemistry*, vol. 43, no. 11, pp. 3137-51, Mar. 2004.

[157] E. Bordignon and H. J. Steinhoff, "Membrane protein structure and dynamics studied by site-directed spin-labeling ESR," in *ESR Spectroscopy in Membrane Biophysics*, Springer, 2007, pp. 129–164.

[158] Z. Zhang et al., "Multifrequency electron spin resonance study of the dynamics of spin labeled T4 lysozyme.," *The journal of physical chemistry. B*, vol. 114, no. 16, pp. 5503-21, Apr. 2010.

[159] C. J. López, M. R. Fleissner, Z. Guo, A. K. Kusnetzow, and W. L. Hubbell, "Osmolyte perturbation reveals conformational equilibria in spin-labeled proteins.," *Protein science : a publication of the Protein Society*, vol. 18, no. 8, pp. 1637-52, Aug. 2009.

[160] K. A. Earle, D. E. Budil, and J. H. Freed, "250-GHz EPR of nitroxides in the slow-motional regime: models of rotational diffusion," *The Journal of Physical Chemistry*, vol. 97, no. 50, pp. 13289-13297, Dec. 1993.

[161] K. Möbius et al., "Combining high-field EPR with site-directed spin labeling reveals unique information on proteins in action.," *Magnetic resonance in chemistry*, vol. 43, p. S4-S19, Nov. 2005.

[162] G. Jeschke, "Instrumentation and Experimental Setup," in *ESR spectroscopy in membrane biophysics*, vol. 19, no. 1, 2007, pp. 17-47.

[163] Z. Liang and J. H. Freed, "An Assessment of the Applicability of Multifrequency ESR to Study the Complex Dynamics of Biomolecules," *The Journal of Physical Chemistry B*, vol. 103, no. 30, pp. 6384-6396, Jul. 1999.

[164] D. Sezer, J. H. Freed, and B. Roux, "Multifrequency electron spin resonance spectra of a spin-labeled protein calculated from molecular dynamics simulations.," *Journal of the American Chemical Society*, vol. 131, no. 7, pp. 2597-605, Feb. 2009.

[165] Y. E. Nesmelov, C. B. Karim, L. Song, P. G. Fajer, and D. D. Thomas, "Rotational dynamics of phospholamban determined by multifrequency electron paramagnetic resonance.," *Biophysical journal*, vol. 93, no. 8, pp. 2805-12, Oct. 2007.

[166] S. Pornsuwan, “Measurements of Dynamics, Structure, and Flexibility of Macromolecules by Pulsed Electron Spin Resonance,” 2007.

[167] D. E. Budil, S. Lee, S. Saxena, and J. H. Freed, “Nonlinear-least-squares analysis of slow-motion EPR spectra in one and two dimensions using a modified Levenberg-Marquardt algorithm,” *Journal of Magnetic Resonance Series A*, vol. 120, pp. 155–189, 1996.

[168] T. Mordasini, A. Curioni, and W. Andreoni, “Why do divalent metal ions either promote or inhibit enzymatic reactions? The case of BamHI restriction endonuclease from combined quantum-classical simulations.,” *The Journal of biological chemistry*, vol. 278, no. 7, pp. 4381-4, Feb. 2003.

[169] V. Pingoud et al., “On the divalent metal ion dependence of DNA cleavage by restriction endonucleases of the EcoRI family.,” *Journal of molecular biology*, vol. 393, no. 1, pp. 140-60, Oct. 2009.

[170] J. L. Woodhead, N. Bhave, and A. Malcolm, “Cation dependence of restriction endonuclease EcoRI activity.,” *European journal of biochemistry / FEBS*, vol. 115, no. 2, pp. 293-6, Apr. 1981.

[171] N. C. Horton, K. J. Newberry, and J. J. Perona, “Metal ion-mediated substrate-assisted catalysis in type II restriction endonucleases.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 23, pp. 13489-94, Nov. 1998.

[172] I. van Amsterdam, M. Ubbink, G. W. Canters, and M. Huber, “Measurement of a Cu□ Cu Distance of 26 Å by a Pulsed EPR Method,” *Angewandte Chemie International Edition*, vol. 42, no. 1, pp. 62–64, 2003.

[173] Z. Yang, M. Ji, and S. Saxena, “Practical Aspects of Copper Ion-Based Double Electron Electron Resonance Distance Measurements,” *Applied Magnetic Resonance*, pp. 487-500, 2010.

[174] Z. Yang, D. Kise, and S. Saxena, “An approach towards the measurement of nanometer range distances based on Cu2+ ions and ESR.,” *The journal of physical chemistry. B*, vol. 114, no. 18, pp. 6165-74, May. 2010.

[175] E. Narr, A. Godt, and G. Jeschke, “Selective measurements of a nitroxide-nitroxide separation of 5 nm and a nitroxide-copper separation of 2.5 nm in a terpyridine-based copper(II) complex by pulse EPR spectroscopy.,” *Angewandte Chemie (International ed. in English)*, vol. 41, no. 20, pp. 3907-10, Oct. 2002.

[176] K. Yamamoto and A. Ishihama, “Transcriptional response of Escherichia coli to external copper.,” *Molecular microbiology*, vol. 56, no. 1, pp. 215-27, Apr. 2005.

[177] C. Rensing and G. Grass, “Escherichia coli mechanisms of copper homeostasis in a changing environment,” *FEMS Microbiology Reviews*, vol. 27, no. 2-3, pp. 197-213, Jun. 2003.

[178] R. P. Bonomo, F. Riggi, and A. J. Di Bilio, “EPR reinvestigation of the copper (II)-imidazole system,” *Inorganic Chemistry*, vol. 27, no. 14, pp. 2510–2512, 1988.

[179] G. Rakhit and B. Sarkar, “Electron spin resonance study of the copper(II) complexes of human and dog serum albumins abd some peptide analogs.,” *Journal of inorganic biochemistry*, vol. 15, no. 3, pp. 233-41, Nov. 1981.

[180] D. L. Huffman et al., “Spectroscopy of Cu(II)-PcoC and the multicopper oxidase function of PcoA, two essential components of Escherichia coli pco copper resistance operon.,” *Biochemistry*, vol. 41, no. 31, pp. 10046-55, Aug. 2002.

[181] E. Aronoff-Spencer et al., “Identification of the Cu2+ binding sites in the N-terminal domain of the prion protein by EPR and CD spectroscopy.,” *Biochemistry*, vol. 39, no. 45, pp. 13760-71, Nov. 2000.

[182] C. S. Burns et al., “Molecular features of the copper binding sites in the octarepeat domain of the prion protein.,” *Biochemistry*, vol. 41, no. 12, pp. 3991-4001, Mar. 2002.

[183] C. S. Burns et al., “Copper coordination in the full-length, recombinant prion protein.,” *Biochemistry*, vol. 42, no. 22, pp. 6794-803, Jun. 2003.

[184] Z. Yang, “Copper Ion-Based Electron Spin Resonance Spectroscopic Rulers,” 2010.

[185] S. D. Wu, X. Xu, Y. Xie, L. Chen, Z. Chen, and Q. Liu, “Cu(II) Coordination Structure in NAD Glycohydrolase from Agkistrodon acutus Venom Studied by EPR Spectroscopy,” *Chemistry Letters*, no. 3, pp. 354-354, 2002.

[186] F. Jiang, J. Mccracken, and J. Peisach, “Nuclear Quadrupole Interactions in Copper (II)-Diethylenetriamine-Substituted Imidazole Complexes and in Copper (II) Proteins,” *Journal of American Chemical Society*, vol. 112, no. 12, pp. 9035-9044, 1990.

[187] W. B. Mims and J. Peisach, “The nuclear modulation effect in electron spin echoes for complexes of Cu2+ and imidazole with 14N and 15N,” *Journal of Chemical Physics*, vol. 69, no. 11, pp. 4921-4930, 1978.

[188] J. Peisach and W. E. Blumberg, “Structural Implications Derived from the Analysis Resonance Spectra of Natural of Electron and Paramagnetic Artificial Copper Proteins,” *Archives of Biochemistry and Biophysics*, vol. 708, pp. 691-708, 1974.

[189] J. Voss, L. Salwiński, H. R. Kaback, and W. L. Hubbell, “A method for distance determination in proteins using a designed metal ion binding site and site-directed spin

labeling: evaluation with T4 lysozyme.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 26, pp. 12295-9, Dec. 1995.

[190] A. Velázquez Campoy and E. Freire, "ITC in the post-genomic era...? Priceless.," *Biophysical chemistry*, vol. 115, no. 2-3, pp. 115-24, Apr. 2005.

[191] S. Leavitt and E. Freire, "Direct measurement of protein binding energetics by isothermal titration calorimetry.," *Current opinion in structural biology*, vol. 11, no. 5, pp. 560-6, Oct. 2001.

[192] T. Wiseman, S. Williston, J. F. Brandts, and L. N. Lin, "Rapid measurement of binding constants and heats of binding using a new titration calorimeter.," *Analytical biochemistry*, vol. 179, no. 1, pp. 131-7, May. 1989.

[193] M. L. Doyle, "Titration microcalorimetry.," *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]*, vol. 20, p. Unit20.4, May. 2001.

[194] Microcal, *ITC Data Analysis in Origin ®*, no. January. 2004, p. 121.

[195] Y. Zhang, S. Akilesh, and D. E. Wilcox, "Isothermal titration calorimetry measurements of Ni(II) and Cu(II) binding to His, GlyGlyHis, HisGlyHis, and bovine serum albumin: a critical evaluation.," *Inorganic chemistry*, vol. 39, no. 14, pp. 3057-64, Jul. 2000.

[196] C. Quinn, *Analyzing ITC Data for the Enthalpy of Binding Metal Ions to Ligands*. 2010, pp. 1-6.

[197] W. B. Turnbull, "Divided We Fall Studying low affinity fragments of ligands by ITC," *Microcal*.

[198] B. M. Baker and K. P. Murphy, "Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry.," *Biophysical journal*, vol. 71, no. 4, pp. 2049-55, Oct. 1996.

[199] N. E. Grossoehme, S. B. Mulrooney, R. P. Hausinger, and D. E. Wilcox, "Thermodynamics of Ni2+, Cu2+, and Zn2+ binding to the urease metallochaperone UreE.," *Biochemistry*, vol. 46, no. 37, pp. 10506-16, Sep. 2007.

[200] R. Dylla-Spears, J. E. Townsend, L. L. Sohn, L. Jen-Jacobson, and S. J. Muller, "Fluorescent marker for direct detection of specific dsDNA sequences.," *Analytical chemistry*, vol. 81, no. 24, pp. 10049-54, Dec. 2009.

[201] R. Dylla-Spears, J. E. Townsend, L. Jen-Jacobson, L. L. Sohn, and S. J. Muller, "Single-molecule sequence detection via microfluidic planar extensional flow at a stagnation point.," *Lab on a chip*, vol. 10, no. 12, pp. 1543-9, Jun. 2010.

[202] J. Gorman and E. C. Greene, "Visualizing one-dimensional diffusion of proteins along DNA.," *Nature structural & molecular biology*, vol. 15, no. 8, pp. 768-74, Aug. 2008.

[203] A. Biebricher, W. Wende, C. Escudé, A. Pingoud, and P. Desbiolles, "Tracking of single quantum dot labeled EcoRV sliding along DNA manipulated by double optical tweezers.," *Biophysical journal*, vol. 96, no. 8, pp. L50-2, Apr. 2009.

[204] M. Betlach, V. Hershfield, L. Chow, W. Brown, H. Goodman, and H. W. Boyer, "A restriction endonuclease analysis of the bacterial plasmid controlling the ecoRI restriction and modification of DNA.," *Federation proceedings*, vol. 35, no. 9, pp. 2037-43, Jul. 1976.

[205] W. Wang and A. Malcolm, "Two-stage polymerase chain reaction protocol allowing introduction of multiple mutations, deletions, and insertions, using QuikChange site-directed mutagenesis.," *Methods in molecular biology (Clifton, N.J.)*, vol. 182, pp. 37-43, Jan. 2002.

[206] R. B. Kapust and D. S. Waugh, "Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused.," *Protein science : a publication of the Protein Society*, vol. 8, no. 8, pp. 1668-74, Aug. 1999.

[207] NEB, *pMAL ™ Protein Fusion and Purification System, Version 5.3*. 2007, p. 56.

[208] F. Studier, "Protein production by auto-induction in high-density shaking cultures," *Protein Expression and Purification*, vol. 41, no. 1, pp. 207-234, May. 2005.