DNA-Mediated Charge Transfer Between [4Fe-4S] Cluster Glycosylases

Thesis by

Christine Anne Romano

In Partial Fulfillment of the Requirements
for the degree of
Doctor of Philosophy in Chemistry

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2011

(Defended March 16, 2011)

Acknowledgments

First and foremost, let me express gratitude to my research advisors, **Professors Jacqueline Barton** and **Dianne Newman**. Jackie was very gracious in accepting me into her lab as a transfer student and giving me a home in the Caltech Chemistry Department. She mentored me through an interdisciplinary research project, and I feel privileged to have learned about so many different fields and experimental techniques. I also appreciate her willingness to let me work so closely with laboratories at both Caltech and MIT, and for her support of my choosing a post-doc in environmental science. My co-advisor, Dianne, was a constant source of inspiration and encouragement, even during the difficult years when she was on the opposite end of the country. I thank her for always remembering me and my potential, and for her unfailing devotion to my personal and academic well-being. Dianne offered both brilliant scientific perspective and tremendous emotional wisdom. She also helped me locate my post-doctoral laboratory, and I hope our paths will continue to cross as I transition into the realm of environmental microbiology.

My committee chairman, **Professor Harry Gray**, is always excited. Whether it's about a minor breakthrough or a major achievement, Harry possessed a keen interest in my progress and my research career that motivated me through the difficult moments of my thesis. I appreciate how he always smiled when I walked into his office and was always willing to meet with me as if I were one of his own students. I thank him for his mentorship, his devotion to my progress and career goals, and his ever-positive presence throughout my Ph.D. journey.

I thank my remaining two committee members, **Professors Francis Arnold** and **Doug Rees** for their constructive comments on my research over the years and their

genuine enthusiasm for my post-doctoral work.

Mo Renta and **Olga Batygin**, assistants to Jackie and Dianne, respectively, provided outstanding support throughout my thesis. Mo and Olga were kind counselors and sympathetic ears. They also kept their laboratories and, in Mo's case, the CCE Division, running smoothly. I enjoyed all of Mo's stories about Idgy, Bammy, and Grayson.

I extend my gratitude to the many **Newman** and **Barton Lab members** past and present for their personal and academic mentorship, and for being both good friends and great co-workers. I particularly thank **Dr. Lars Dietrich**, who taught me the basics of molecular biology and microbiology. Lars is a brilliant researcher and I know he will become an outstanding mentor to his young laboratory at Columbia University. Also deserving of special recognition are **Dr. Amie Boal**, who initiated my thesis project, and **Dr. Alon Gorodetsky**, who helped me realize that I could, after all, conquer my most loathed experimental technique: electrochemistry. **Dr. Brian Zeglis** offered many words of encouragement and on occasion, tough love, that I eventually tried to take to heart. He also took me camping many times. **Dr. Paul Lee** continually amused and encouraged me, and he also proved a great swing-dance partner and artistic director for cake-baking projects. **Eric Olmon** and **Wendy Mercer** were incredibly supportive proofreaders, providers of pep-talks, and fellow members of the incoming class of September, 2005. **Wendy,** in particular, was a remarkably sympathetic friend and colleague. She commiserated with me through all the trials and tribulations of transferring labs, navigating an interdisciplinary research project, coping with laboratory politics, and battling a scientific world that sometimes seems aligned against you. It helps that she also likes cute

bunnies and let me babysit hers on occasion. Wendy defends her thesis a few months after I defend mine, and I wish her the best of luck and fortune when she begins her post-doc at Cornell.

Just down the hall from the Barton Lab, the Okumura Group students were also great friends. Thanks to **Dr. Laurence Yeung** and **Kana Takematsu** for listening to my many stories of laboratory misfortunes. It helps that their lab's break room almost always has cookies.

As a first-year, I joined other women students for weekly tea nights to chat about our experiences. Thanks to **Dr. Katy Muzikar, Brinton Seashore-Ludlow, Sylvie Wintenberger, Dr. Jen Stockdill**, and **Dr. Jenny Roizen** for all the tea and cookies. During the early years of graduate school, I also joined a **Potluck Group** that met on Sundays. They allowed me to start each week with lots of free food and social interaction. They also organized many outings downtown, ski trips, and other adventures in Los Angeles and California. Many of these people have "grown-up" in graduate school with me, and I wish them luck as they complete their Ph.D.s and move onto future endeavors. Special congrats go to **Eve Stenson** and **Brian Standley** and **Adrienne Erickcek** and **Nick Law**, two potluck couples who recently married, and **Janelle Bray, Dan Grin, Jessie Rosenberg**, and **Laurence Yeung**, who graduated last June. I am sure that **Kana Takematsu, Matt Kelley**, **Wendy Mercer**, and **Neil Halelamien** will soon follow. Members of this potluck group convinced me to join the **Caltech Ballroom Dance Club**, which proved an unexpected source of entertainment and social activities. Thanks for putting up with my terrible DJ-ing skills, my staunch anti-sparkle-ism, and my relentless assertions that Lindy Hop is much a better dance than Ceroc or West Coast Swing (because

it is!!!).

Later in graduate school, I joined **Prufrock Group**, a cooking group that also became a community of treasured friends. They provided many free meals and the opportunity to transfer my scientific frustrations into culinary endeavors. There's a reason I'm good at chopping potatoes! More importantly, though, most members of Prufrock were close to my year in graduate school and, as we grew up together, they also provided me with encouragement and perspective as I navigated its difficult moments. Many thanks to **Zuli Kurji, Theresa Emery, Matt Kelley, Heather McCaig**, and **Michael Mendenhall** for keeping my stomach full and your ears open. In addition to maintaining the Prufrock club's backyard garden, **Heather McCaig** was an outstanding roommate for the better part of five years. Heather is a fellow tea fanatic, an outstanding cook of often unusual cuisines, and a willing participant in late-night social gatherings. I am still amazed that she and I have managed to remain such good friends after living in the same apartment for so long. She has remained upbeat and undeniably quirky through all the difficulties of her Ph.D. project, and I admire her spirit. I am also constantly amazed when I watch her dance ballet. I wish Heather all the best in her post-Caltech endeavors.

In addition to providing a wonderful support network of friends and fellow scientists-in-training, Caltech maintains many student support services that have greatly enriched my Ph.D. experience. I cannot endorse the **Caltech Y** enough. The staff there is always positive and friendly, and they maintain constant smiles and calm demeanors amidst what must be a very demanding workload for such a small office. In addition to overseeing volunteer activities and cultural trips in which I have participated, the Caltech Y **Outdoor Committee** has dramatically changed my years in Pasadena. I came to Southern California

expecting to find Mickey Mouse, some palm trees, and lots of smog. The Caltech Y's many day hikes, camping trips, and backpacking trips helped me realize that California is a much more diverse, complex, and gorgeous place to live. I did not realize that so much natural beauty and dramatically varying landscapes were located so close to such populous urban centers. The Caltech Y taught me to love the great outdoors, which prompted my participation in the **Caltech Hiking Society**, and my desire to travel all over California, and even into Arizona's Grand Canyon, in search of wilderness adventure. These adventures inspired me to elect a post-doc that involves research in Yellowstone National Park.

Another student support office that was pivotal for my progress was the Dean of Graduate Student's' Office. The two desks in the front of this office are occupied by **Icy Ma** and **Natalie Gilmore**, who pleasantly saw me through the piles of paperwork that marked almost every milestone on the path towards a Ph.D. The wonderful **Dr. Felicia Hunt** occupies one of the back offices. I cannot say enough about Felicia. Her job requires constant oversight and organizing of campus events, recruitment weekends, orientation, and graduate student quality of life concerns, although she always makes the time to meet one-on-one with those students who need assistance. She is an academic advisor, emotional counselor, and professional advocate all rolled into one. She fights for all students, particularly women, to have the opportunity to complete their Ph.D. and pursue a career in science. Felicia helped me overcome my most severe obstacles in graduate school and celebrate the tiniest accomplishments. She is still, carefully, but proudly, watching me approach the finish line.

Before her promotion to assistant dean, Felicia Hunt oversaw the Caltech Women's Center. Felicia and her successor, **Portia Harris**, arranged programming, camaraderie

building, social networking, and assistance for women graduate students, who comprise a surprisingly small percentage of Caltech's student population. All of the workshops, presentations, and conversations were tremendously inspiring. Among the encouraging women I met at the Women's Center were **Dr. Heather Hunt**, a recent Caltech graduate who had a difficult thesis project but is currently on the cusp of accepting an academic position, and **Dr. Roberta Hansman**, an oceanographer who encouraged my enthusiasm for environmental science. Roberta helped me search for post-doctoral positions and even drove me to one of my interviews. I wish Roberta best of luck as she completes her post-doc and moves on in the scientific world.

Speaking of environmental scientists, I thank the **Caltech Roof Lab** (a.k.a. the Flagan and Seinfeld research groups) for becoming such close friends in grad school and welcoming me into their community. You guys were my "lab away from lab" and provided a social network when I needed time away from my own research group. I was invited into the Roof Lab family by **Dr. Arthur Chan**, who has been an incredible source of strength and support throughout my Ph.D. I appreciate Arthur taking me away from campus so often when I needed the distraction, and for continuing to check in on me from his current post in Berkeley. He is also going to get me a golden retriever someday. Arthur, his roommate **Andrew Metcalf**, and labmate **Christine Loza**, cooked for me frequently.

The Chemistry Department Staff, past and present, proved crucial for my ability to complete research projects. Thanks to **Dian Buchness, Laura Howe, Agnes Tong, Steve Gould, Tom Dunn, Anne Penney, Chris Smith, Joe Drew**, and **Ron Koen**.

I thank **Geoff** and **Karen Blake** for letting me be partially adopted by their

gorgeous golden retriever, **Ranger**, during the last few months of my thesis. All the walks, frolics, and play sessions were such joyful diversions from the tedium of writing. On that note, many of Ranger's kindred spirits also helped me endure the difficulties of graduate school. My gratitude goes to **Leia, Thao, Cody, Pepper, Salt, Tiger, Eva, Jackie, Tricksy, Twiggy, Toby, Oliver, Tobi, Xerion, Coco, Puffy, Plato, Socrates**, **Buttercup**, and their respective people.

I would like to conclude these acknowledgments with a personal note to my family. First, let me thank my ever-nurturing grandmother, **Consuelo Soto**. Her kindness and gentle voice always make me smile. She amuses me and cares for me, and it is great to know that I will always have a place to stay when I visit Austin, TX. One of my grandmother's other great fans is my sister, **Claire Romano**. Claire is the family artist, and while our personalities differ greatly, she has never failed to call (eventually...) to celebrate a birthday or when she hears of an accomplishment. She also keeps me posted on all the events of Central Texas, a region that I miss dearly. I appreciate her encouragement of her nerdy older sister during the Ph.D. struggle. An admirer of both my grandmother and my sister is **Boerne**, the family's shepherd-hound mix. It's weird that I admire Boerne so much, given that his life ambitions rarely extend beyond napping or acquiring that one extra piece of cheese. However, Boerne was rescued from mistreatment, and he has recovered from those conditions and learned to love my family with so much positive zeal. I hope that when I have to endure difficult circumstances, I will also be able to remain hopeful and someday land in a world surrounded by treats, toys, and adoring companions.

Graduate school and all its struggles have actually strengthened my relationship with my parents, **Martha** and **Michael Romano**, and their continued support has meant the

world to me. I thank them for understanding that yes, I'm still alive, even though I do not always call on time. I thank them for trying to learn about everything from DNA to microbiology in order to understand what I do. I thank them for letting me haul them up California mountaintops on long hikes, wheezing all the way, because they wanted to experience what I enjoy about living in Pasadena. I thank them for sending me to the Los Angeles Philharmonic on my birthday because they know that I want to experience both fine arts and science during my Ph.D. studies. I thank them for reminding me that they are always proud of how much I have achieved, no matter how insignificant I perceive my accomplishments to be. I also thank them for letting me explore this unknown and mysterious career path—graduate school—and for pledging their unfailing devotion to me through whatever lies ahead.

Abstract

The work performed herein describes three proteins: Uracil DNA glycosylase (UDG) from *Archaeoglobus fulgidus*, MutY, and Endonuclease III (EndoIII) from *Escherichia coli*. They are DNA repair glycosylases that contain [4Fe-4S] clusters. While the catalytic mechanisms of many BER enzymes have been studied in detail, questions remain about how these enzymes search the vast amount of cellular DNA to find their substrates, and why some require a [4Fe-4S] cluster. The iron-sulfur cluster is not necessary for catalysis, and it only displays a physiologically relevant midpoint potential when bound to DNA. We have proposed that UDG, MutY, and EndoIII use their [4Fe-4S] clusters to participate in DNA-mediated charge transport (CT), and that these proteins mediate long-range electrochemical signaling in order to detect DNA damage.

This scheme for DNA damage detection assumes that CT occurs efficiently between the DNA helix and the [4Fe-4S] cluster of the bound protein. In order for efficient CT to occur, a pathway of amino acids must be present that facilitates CT between the DNA and the iron-sulfur cluster. For each of the enzymes mentioned, this pathway was explored through mutagenesis. In UDG, MutY, and EndoIII, several amino acids thought to be important for CT were mutated and the resulting proteins were characterized biochemically. Their CT capabilities were analyzed by cyclic voltammetry on DNA-modified electrodes. In these experiments, the mutants' signal intensities were quantified and compared to those of wild-type enzyme. An attenuated signal relative to wild-type protein may indicate that the mutant is deficient in CT and that the targeted amino acid is part of the protein-DNA CT pathway in the native enzyme. Many mutants were also

screened by enzymatic assays and circular dichroism spectroscopy to further characterize their DNA-binding properties and structural stability.

The *A. fulgidus* UDG mutants examined, C17H, C85S, and C101S, all contained mutations in the cysteine residues that ligate the [4Fe-4S] cluster. These mutants were designed to determine how the iron-sulfur cluster coordination environment affects protein-DNA CT. The mutants exhibited varying signal strengths relative to WT UDG on DNA-modified electrodes. C85S produced a weaker signal, indicating a CT deficiency. The signal intensity from C101S was within error of that of WT, and the signal from C17H was larger than that of WT, possibly indicating that this mutant is less structurally stable than WT UDG.

In *E. coli* MutY, position Y82 aligns with Y165 in MUTYH, a residue in which mutations have been found in many colorectal cancer patients. To better understand the correlation between protein-DNA CT and colorectal cancer, the MutY mutants Y82C and Y82L were prepared and characterized. Y82C exhibited a CT deficiency relative to WT MutY, whereas Y82L did not. These data indicate that Y82 forms part of the CT pathway in native *E. coli* MutY, but that other long-chain amino acids, such as leucine, can also mediate CT efficiently at this position.

Several different mutants of *E. coli* EndoIII were examined. First, the Y82 position was targeted, since the aligning MUTYH residue has been found mutated in colorectal cancer patients and because this residue is located near the protein-DNA interface. Five mutations were made at or near the Y82 position, and their cyclic voltammetry signals demonstrated that aromatic amino acids best mediate CT at this position. Other residues towards the interior of the protein, Y75, Y55, and F30 were also mutated to alanines.

These mutants exhibited CT deficiencies, implicating the residues as part of a potential CT pathway. Residues W178 and Y185, located near the [4Fe-4S] cluster of EndoIII, were also mutated to alanines. The resulting mutants produced larger signals than that of WT EndoIII. These mutants were later shown by circular dichroism spectroscopy to be less stable structurally than WT EndoIII. All of the mutants mentioned exhibited enzymatic properties similar to those of WT, suggesting that they are able to bind DNA and excise damage nucleobases as well as the native enzyme. Several of these mutants were also used in a mutagenesis-based experiment to assay how EndoIII variants help MutY search for DNA lesions, although data from these experiments showed no significant differences in mutation rate between strains expressing different EndoIII variants.

In total, the mutagenesis studies performed here helped determine the characteristics of BER enzymes that enable them to mediate DNA-protein CT. All these enzymes must contain a stable, well-protected metallocluster that charge can access through a series of CT-facilitating amino acids. In discovering several residues important for protein-DNA CT in UDG, MutY, and EndoIII, we have strengthened support for the hypothesis that these enzymes facilitate DNA-mediated CT *in vivo*. These enzymes may in fact be part of a much larger array of redox-active DNA-binding proteins that communicate electrochemically to help each other detect and repair DNA lesions inside the cell.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Chapter 1

Chapter 2

Chapter 3

Appendix II

Appendix III

Chapter 1

[4Fe-4S] Cluster Base Excision Repair Glycosylases

Introduction

Cells constantly undergo an onslaught of potentially mutagenic processes that include exposure to ultraviolet light, reactive oxygen species, and intracellular alkylating agents [1–3]. If cells cannot detect and repair the DNA damage that these processes create, then the genomic mutations that ensue could have a profound effect on the cells' ability to faithfully replicate and survive [4]. Consequently, most organisms have evolved a stunning array of enzymes that detect and repair damaged DNA [3], including the *E. coli* glycosylases endonuclease III (EndoIII) and MutY and *Archaeaglobus fulgidus* Uracil DNA glycosylase (UDG) (Figure 1.1). These enzymes are part of the base excision repair (BER) pathway, meaning that they repair damaged DNA by cleaving the *N*-glycosidic bond of a damaged nucleotide and excising it from the helix [5–8]. Endonuclease III repairs damaged pyrimidines [9, 10], MutY removes adenine from mismatches between adenine and 7,8-dihydro-8-oxo-2'-deoxyguanosine (8–oxo-G) [8, 11, 12], and UDG excises uracil present in DNA [13]. The molecular mechanisms by which these proteins excise their substrates are fairly well-understood, but less attention has been devoted to understanding how BER glycosylases detect DNA damage in the first place [2].

We have proposed that BER enzymes may detect DNA damage through DNA-mediated charge transfer [14, 15]. The base pair π-stack of duplex DNA is capable of mediating charge transport (CT) through distances of over 200 Å

[16–18]. Importantly, alterations in the π-stack, such as those produced by DNA lesions, can disrupt the flow of CT through the helix [19–22]. BER enzymes may sense these CT disruptions, and are thus alerted to the presence of a DNA lesion. In this model for lesion detection, the [4Fe-4S] clusters of BER enzymes are crucial to their *in vivo* function because they are the redox-active moieties that mediate CT interactions with the DNA helix. Several lines of evidence suggest that the [4Fe-4S] clusters of BER enzymes function as CT mediators. The clusters do not catalyze the base excision reaction [5], and they are not readily oxidized or reduced within a physiologically relevant range of potentials [9]. This apparent lack of activity has led some authors to suggest that the [4Fe-4S] clusters have a structural role [5, 23, 24]. However, there is some biochemical evidence that MutY can fold properly without the [4Fe-4S] cluster [25]. Furthermore, physiologically relevant midpoint potentials can be measured when the proteins are bound to DNA [14, 26], arguing that the iron-sulfur clusters confer redox capabilities rather than structural stability. These data have led to the development of the following model for DNA damage detection: 1) The [4Fe-4S] cluster of a BER enzyme typically resides in the 2^+ state [9]. Oxidative stress activates the enzyme to the 3^+ state, causing it to and bind DNA more tightly [27]. This oxidative stress may be funneled to the protein through DNA-mediated CT after having formed at guanine residues [28]. 2) The newly oxidized protein loses an electron that gets transmitted through the DNA helix and ultimately reduces a second, distally located DNA-bound repair protein. 3) Reduction reduces the DNA binding affinity of this second enzyme, which then

dissociates from the DNA and re-binds to a different portion of the genome [27]. If part of the DNA helix is damaged, then this damage will impair CT in step 2 so that the BER enzymes will remain bound to DNA and continue to search for the damaged nucleotide.

We have tested this model for DNA damage detection through DNA-mediated CT though a variety of experiments. These include *in vivo* assays of cooperativity between BER enzymes [29], AFM assays of whether BER enzymes redistribute onto mismatch-containing DNA strands [29], and experiments that detected an increase in DNA binding affinity upon oxidation of EndoIII [27]. However, one important component of this model remains to be examined: the mechanism by which charge is transported between the DNA helix and the metallocluster of the bound repair enzyme. In other charge-transfer active proteins, CT between redox active moieties often occurs through a pathway of amino acids that can act as "stepping stones" though which electrons or holes pass as they traverse the protein [30–33]. These CT-active residues can be discovered through mutagenesis experiments in which CT candidate residues are mutated, and the resulting protein sample is assayed for a deficiency in CT activity. Here, assays of electrochemical activity were performed on site-directed mutants of three different redox-active BER glycosylases: Uracil DNA Glycosylase (UDG), MutY, and Endonuclease III (EndoIII). These individual proteins are described in detail below.

Uracil DNA Glycosylase

Uracil DNA Glycosylases (UDG) remove uracil and related compounds from DNA [2], (Figure 1.2). Uracil can be misincorporated into DNA by synthesis enzymes [34]. However, more commonly, it is formed by hydrolytic deamination of cytosine [2]. If not repaired, G:U base pairs will result in G:C → A:T transversion mutations [35]. Indeed, *E. coli* cells in which UDG has been inactivated show an increase in G:C → A:T transversion rates [35]. UDG enzymes are found in all three domains of life, although differences in substrate, structure, and active site residues have caused them to be divided into six different families [36]. Family I, II, and III UDG enzymes are found in eukaryotes and some bacteria [36]. UDG proteins from families IV, V, and VI have, thus far, only been isolated from extremophilic organisms. They are distinguished from other UDG families in that they contain a [4Fe-4S] cluster. Given its distance from the putative substrate binding site, this [4Fe-4S] cluster is unlikely to be catalytic in UDG. We have proposed that extremophilic UDG enzymes use their [4Fe-4S] clusters to participate in electrochemical long-range signaling for DNA damage detection. The ability to detect lesions electrochemically may prove particularly useful for extremophilic organisms because rates of cytosine deamination increase with temperature [37]. Many family IV, V, and VI UDGs have been isolated from thermophilic prokaryotes [38–40]. In order to better understand the role of the [4Fe-4S] clusters in thermophilic UDGs, C17H, C85S, and C101S mutants of *Archaeoglobus fulgidus* UDG were prepared. The mutations target residues that ligate the [4Fe-4S]

cluster. *A. fulgidus* is a thermophilic archaeon, and the mutants that are examined here were tested to determine how changes in the iron-sulfur cluster coordination environment affect the charge-transfer properties of the protein.

MutY

MutY homologues are found in all three domains of life and are part of a repair system that helps cells respond to oxidative stress. When excess reactive oxygen species form *in vivo*, these species can damage DNA, particularly at guanine residues since guanine has the lowest oxidation potential of all the nucleobases [41, 42]. When guanine is oxidized, it forms 7,8-dihydro-8-oxoguanine (8-oxo-G, Figure 1.2) [43]. Excess intracellular 8-oxo-G is removed by the enzyme MutT [11]. If it gets incorporated into DNA, it is excised by MutM [11, 12]. However, if 8-oxo-G remains misincorporated, then subsequent rounds of DNA replication will mistakenly pair an adenine molecule with it [44]. The next round of DNA replication will then place a thymine across from this adenine, resulting in a G:C → T:A transversion. MutY is the "final defense" against these transversions, as it removes adenine mispaired with 8-oxo-G [11, 44, 45]. The enzymes that repair 8-oxo-G-based lesions were first discovered by genetic mapping of cells with a strong G:C → T:A mutator phenotype. The resulting experiments identified the *mutY* locus as being among those responsible [45].

In human cells, MutY also plays an important role in maintaining genomic integrity. The human homologue of MutY, MUTYH, has recently drawn the

attention of colorectal cancer researchers. Patients with mutations in MUTYH are predisposed towards developing colorectal cancer [47]. These patients tend to acquire mutations in the Adenomatous Polyposis Coli (APC) gene, which governs the proliferation of colonic cells [47, 48]. If APC gene mutations accumulate, and MUTYH is unable to detect and repair them, then colorectal tissue may become cancerous. Several different cancer-relevant MUTYH mutations have been discovered [47], and two of their *E. coli* counterparts Y82C and Y82L were prepared and examined for electrochemical activity herein.

Endonuclease III

Endonuclease III (EndoIII) acts upon a variety of substrates, excising several oxidative damage products of cytosine and thymine from the DNA backbone [2] (figure 2). Cells lacking EndoIII exhibit only a weak mutator phenotype, suggesting that many of its substrates may also be repaired by other enzymes. EndoIII contains a [4Fe-4S] cluster which, we argue, allows it to communicate electrochemically with other DNA-bound proteins such as MutY [14] and SoxR [49, 50] in order to detect DNA lesions. For the experiments discussed in this report, eleven mutants of EndoIII were prepared targeting residues suspected to form part of the CT pathway between the DNA and the [4Fe-4S] cluster in EndoIII. These mutants were then characterized biochemically to determine whether their CT capabilities were decreased relative to those of WT EndoIII. Because EndoIII

bears structural and amino acid sequence homology to MUTYH, several of the EndoIII mutants examined are also relevant to colorectal cancer research.

In total, the research performed in this thesis uses three proteins as model systems to better understand the biological relevance of DNA-mediated charge transfer. For each of these three proteins, UDG, MutY, and EndoIII, several site-directed mutants were prepared with four questions in mind: 1) How do perturbations near the [4Fe-4S] cluster affect the CT properties of redox active DNA repair enzymes? 2) Which types of amino acid residues best mediate protein-to-DNA CT? 3) What possible amino acid pathways may exist through which charge flows? 4) What are the biological consequences of having an impaired CT pathway? These questions were addressed though a variety of biochemical techniques including enzymatic activity assays, cyclic voltammetry on DNA-modified electrodes, circular dichroism spectroscopy, and *in vivo* assays of cooperativity between select BER proteins. In total, these experiments present how biological systems could exploit DNA CT *in vivo* in order to detect and repair damaged DNA.

Figure 1.1: Crystal structures of the base excision repair proteins to be investigated. Each contains a [4Fe-4S] cluster. EndoIII (*E. coli*), MutY (*B. stearothermophilus*), and UDG (*T. thermophilus*) structures were adapted from references [5, 23, 51, 52], respectively, and formatted in PyMol [53].

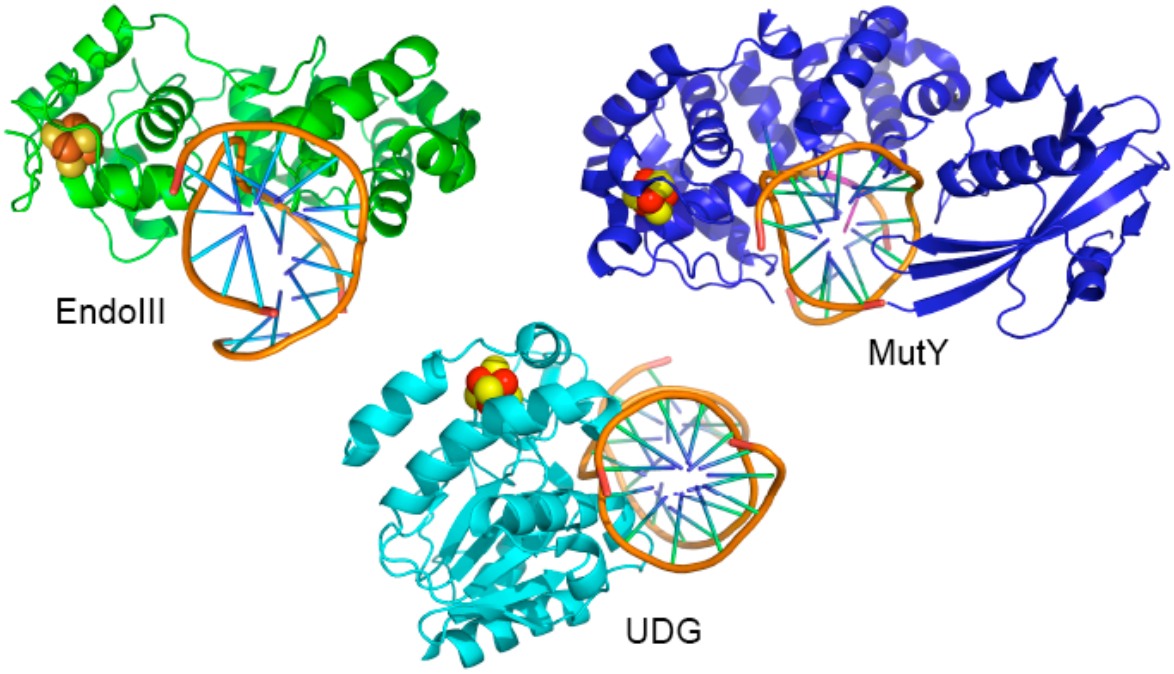
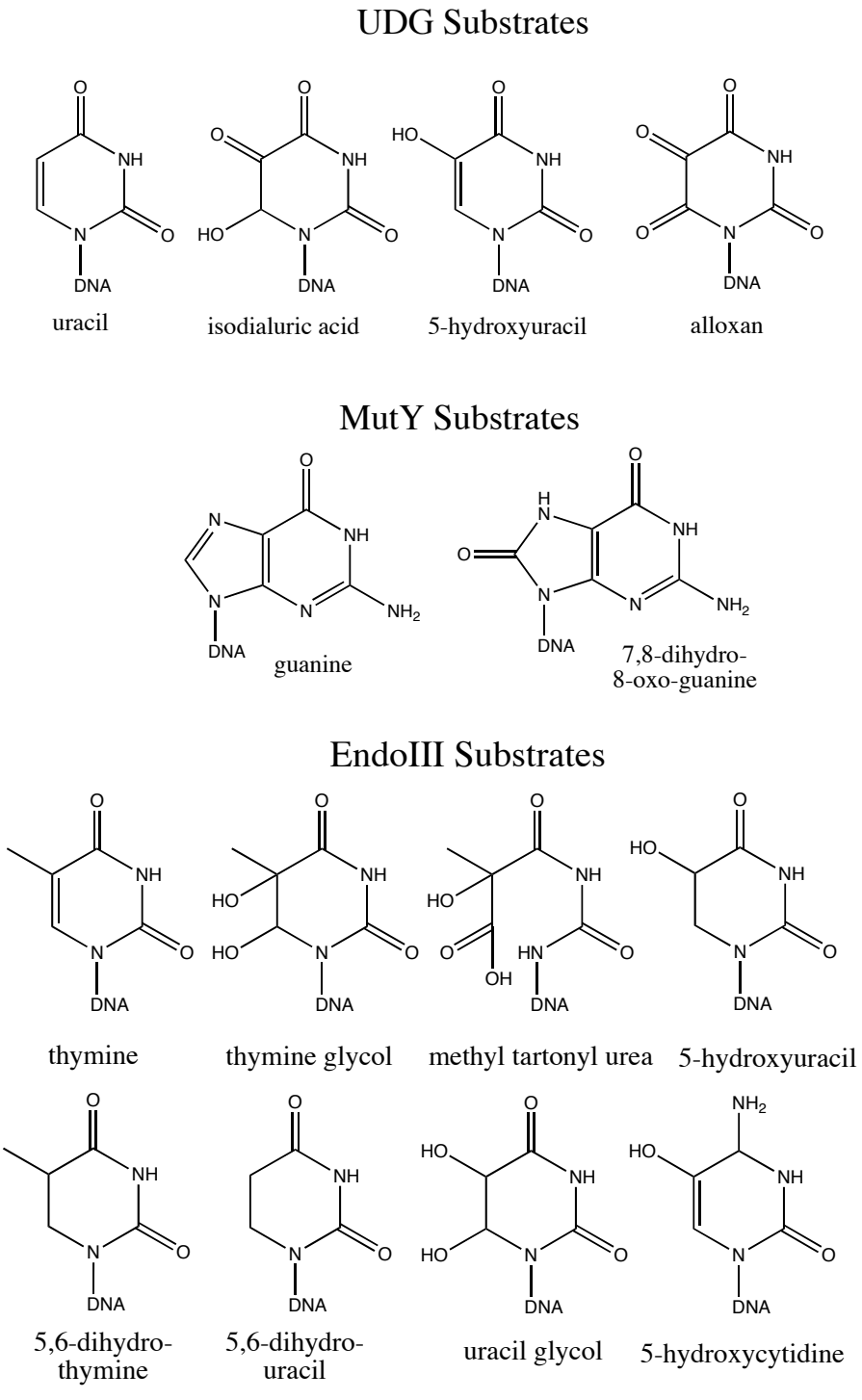


Figure 1.2: Substrates of BER enzymes. UDG excises uracil and related compounds from DNA, MutY excises adenine mispaired with 7,8-dihydro-8-oxoguanine, and EndoIII targets oxidized pyrimidines.



References:

1. Friedberg, E.C., *DNA damage and repair.* Nature, 2003. **421**: 709–715.
2. David, S.S., and S.D. Williams, *Chemistry of Glycosylases and Endonucleases Involved in Base Excision Repair.* Chemical Reviews, 1998. **98**: 1221–1261.
3. Friedberg, E.C., G.C. Walker, and W. Siede, *DNA Repair and Mutagenesis*. 1995, Washington, D.C.: ASM Press.
4. Lindahl, T., *Instability and decay of the primary structure of DNA*. Nature, 1993. **362**: 709–715.
5. Thayer, M.M., et al., *Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure.* The EMBO Journal, 1995. **14**(16): 4108–4120.
6. Manuel, R.C., et al., *Reaction Intermediates in the Catalytic Mechanism of Escherichia coli MutY DNA Glycosylase.* Journal of Biological Chemistry, 2004. **279**: 46930–46939.
7. Demple, B., and S. Linn, *DNA N- glycosylases and UV repair.* Nature, 1980. **287**: 203–208.
8. Au, K.G., et al., *Escherichia coli mutY gene encodes an adenine glycosylase active on G-A mispairs.* Proc. Natl. Acad. Sci. USA, 1989. **86**: 8877–8881.
9. Cunningham, R.P., et al., *Endonuclease III Is an Iron-Sulfur Protein.* Biochemistry, 1989. **28**: 4450–4455.
10. Breimer, L.H., and T. Lindahl, *DNA Glycosylase Activities for Thymine Residues Damaged by Ring Saturation, Fragmentation, or Ring Contraction Are Functions of Endonuclease III in Escherichia coli.* Journal of Biological Chemistry, 1984. **259**(9): 5543–5548.
11. Michaels, M.L., and J.H. Miller, *The GO System Protects Organisms from the Mutagenic Effect of the Spontaneous Lesion 8-Hydroxyguanine (7,8-Dihydro-8-Oxoguanine).* Journal of Bacteriology, 1992. **174**(20): 6321–6325.
12. Michaels, M.L., et al., *Evidence that MutY and MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA.* Proc. Natl. Acad. Sci. USA, 1992. **89**: 7022–7025.
13. Lindahl, T., *An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues.* Proc. Natl. Acad. Sci. USA, 1974. **71**(9): 3649–53.
14. Boal, A.K., et al., *DNA-Bound Redox Activity of DNA Repair Glycosylases Containing [4Fe- 4S] Clusters.* Biochemistry, 2005. **44**: 8397–8407.
15. Yavin, E., et al., *Electron trap for DNA-bound repair enzymes: A strategy for DNA-mediated signaling.* Proc. Natl. Acad. Sci. USA, 2006. **103**(10): 3610–3614.

16. Hall, D.B., R.E. Holmlin, and J.K. Barton, *Oxidative DNA damage through long-range electron transfer.* Nature, 1996. **382**: 731–735.
17. Nunez, M.E., D.B. Hall, and J.K. Barton, *Long-range oxidative damange to DNA: effects of distance and sequence.* Chem. Biol., 1999. **6**: 85–97.
18. Kelley, S.O., et al., *Single-base mismatch detection based on charge transduction through DNA.* Nucleic Acids Research, 1999. **27**(24): 4830–4837.
19. Kelley, S.O., and J.K. Barton, *Electron Transfer Between Bases in Double Helical DNA.* Science, 1999. **283**: 375–381.
20. Kelley, S.O., et al., *Photoinduced Electron Transfer in Ethidium-Modified DNA Duplexes: Dependence on Distance and Base Stacking.* J. Am. Chem. Soci., 1997. **119**(41): p. 9861–9870.
21. Boal, A.K., and J.K. Barton, *Electrochemical Detection of Lesions in DNA.* Bioconjugate Chem., 2005. **16**: 312–321.
22. Boon, E.M., et al., *Mutation detection by electrocatalysis at DNA-modified electrodes.* Nature Biotechnology, 2000. **18**: 1096–1100.
23. Fromme, J.C. and G.L. Verdine, *Structure of a trapped endonuclease III-DNA covalent intermediate.* The EMBO Journal, 2003. **22**(13): 3461–3471.
24. Lukianova, O.A., and S.S. David, *A role for iron-sulfur clusters in DNA repair.* Current Opinion in Chemical Biology, 2005. **9**:145–151.
25. Porello, S.L., M.J. Cannon, and S.S. David, *A Substrate Recognition Role for the [4Fe-4S]2+ Cluster of the DNA Repair Glycosylase MutY.* Biochemistry, 1998. **37**: 6465–6475.
26. Boon, E.M., et al., *DNA-mediated charge transport for DNA repair.* Proc. Natl. Acad. Sci. USA, 2003. **100**:12543–12547.
27. Gorodetsky, A.A., A.K. Boal, and J.K. Barton, *Direct Electrochemistry of Endonuclease III in the Presence and Absence of DNA.* J. Am. Chem. Soc., 2006. **128**:12082–12083.
28. Yavin, E., et al., *Protein-DNA charge transport: Redox activation of a DNA repair protein by guanine radical.* Proc. Natl. Acad. Sci. USA, 2005. **102**(10): p. 3546-3551.
29. Boal, A.K., et al., *Redox signaling between DNA Repair Proteins for efficient lesion detection.* Proc. Natl. Acad. Sci. USA, 2009. **106**(36): 15237–15242.
30. Stubbe, J., et al., *Radical Initiation in the Class I Ribonucleotide Reductase: Long-Range Proton-Coupled Electron Transfer?* Chemical Reviews, 2003. **103**(6):2167–2202.
31. Cordes, M., et al., *Influence of amino acid side chains on long-distance electron transfer in peptides: Electron hopping via "Stepping Stones".* Angewandte Chemie—International Edition, 2008. **47**(18):3461–3463.
32. Giese, B., M. Graber, and M. Cordes, *Electron transfer in peptides and proteins.* Curr. Opin. Chem. Biol., 2008. **12**(6):755–9.

33. Giese, B., et al., *Electron Relay Race in Peptides.* Journal of Organic Chemistry, 2009. **74**(10):3621–3625.
34. Mosbaugh, D.W., and S.E. Bennett, *Uracil-excision DNA repair.* Prog. Nucleic Acid Res. Mol. Biol., 1994. **48**:315–70.
35. Duncan, B.K. and B. Weiss, *Specific Mutator Effects of Ung (Uracil-DNA Glycosylase) Mutations in Escherichia-Coli.* Journal of Bacteriology, 1982. **151**(2): 750–755.
36. Pearl, L.H., *Structure and function in the uracil-DNA glycosylase superfamily.* Mutat. Res., 2000. **460**(3–4):165-81.
37. Lindahl, T. and B. Nyberg, *Heat-Induced Deamination of Cytosine Residues in Deoxyribonucleic-Acid.* Biochemistry, 1974. **13**(16): 3405–3410.
38. Sakai, T., et al., *Mutagenesis of uracil-DNA glycosylase deficient mutants of the extremely thermophilic eubacterium Thermus thermophilus.* DNA Repair, 2008. **7**(4): 663–669.
39. Hoseki, J., et al., *Crystal structure of a family 4 Uracil-DNA glycosylase from Thermus thermophilus HB8.* Journal of Molecular Biology, 2003. **333**(3): 515–526.
40. Kosaka, H., et al., *Crystal structure of family 5 uracil-DNA glycosylase bound to DNA.* Journal of Molecular Biology, 2007. **373**(4): 839–850.
41. Saito, I., et al., *Photoinduced DNA Cleavage Via Electron-Transfer—Demonstration That Guanine Residues Located 5' to Guanine Are the Most Electron-Donating Sites.* Journal of the American Chemical Society, 1995. **117**(23): 6406–6407.
42. Steenken, S., and S.V. Jovanovic, *How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution.* Journal of the American Chemical Society, 1997. **119**(3): 617–618.
43. Neeley, W.L., and J.M. Essigmann, *Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products.* Chemical Research in Toxicology, 2006. **19**(4): 491–505.
44. Michaels, M.L., et al., *A Repair System for 8-Oxo-7,8-Dihydrodeoxyguanine.* Biochemistry, 1992. **31**(45):10964–10968.
45. Nghiem, Y., et al., *The Muty Gene—a Mutator Locus in Escherichia-Coli That Generates G.C->T.A Transversions.* Proceedings of the National Academy of Sciences of the United States of America, 1988. **85**(8): 2709–2713.
46. Cupples, C.G., and J.H. Miller, *A set of lacZ Mutations in Escherichia coli that Allow Rapid Detection of Each of the Six Base Substitutions.* Proc. Natl. Acad. Sci. USA, 1989. **86**:5435–5439.
47. Cheadle, J.P., and J.R. Sampson, *MUTYH-associated polyposis—From defect in base excision repair to clinical genetic testing.* DNA Repair, 2007. **6**: 274–279.

48. Fearnhead, N.S., M.P. Britton, and W.F. Bodmer, *The ABC of APC.* Human Molecular Genetics, 2001. **10**(7): 721–733.
49. Lee, P.E., B. Demple, and J.K. Barton, *DNA-mediated redox signaling for transcriptional activation of SoxR.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(32):13164–13168.
50. Gorodetsky, A.A., et al., *DNA binding shifts the redox potential of the transcription factor SoxR.* Proc. Natl. Acad. Sci. USA, 2008. **105**(10): 3684–3689.
51. Fromme, J.C., et al., *Structural basis for removal of adenine mispaired with 8- oxoguanine by MutY adenine DNA glycosylase.* Nature, 2004. **427**(652-656): p. 652.
52. Kosaka, H., et al., *Crystal structure of uracil-DNA glycosylase in complex with AP:C containing DNA*. RCSB Protein Data Bank 2DP6, 2007.
53. *The PyMOL Molecular Graphics System, Version 1.2r3pre*, Schrödinger, LLC.

Chapter 2

[4Fe-4S] Cluster Ligand Mutants of *Archaeoglobus Fulgidus* Uracil DNA Glycosylase

UDG protein samples were prepared by L. Engstrom and X. Xu from the laboratory of Professor S. David (University of California, Davis).

Abstract

Archaeoglobus fulgidus is a hyperthermophilic archaeon that expresses a [4Fe-4S] containing uracil DNA glycosylase (UDG) enzyme. UDG proteins excise uracil that has been misincorporated into DNA. This uracil is often formed by deamination of cytosine, a process whose rate is greatly accelerated at high temperatures. In order to cope with this excess DNA damage, thermophilic UDGs may contain unique features that enable them to better detect and repair their substrate. We have proposed that *A. fulgidus* UDG uses its [4Fe-4S] cluster to detect DNA lesions through DNA-mediated CT. To better understand the charge-transfer characteristics of *A. fulgidus* UDG, several variants were created that contain mutations in the cysteine residues that ligate the iron-sulfur cluster. The C17H, C85S, and C101S mutants were examined by cyclic voltammetry on DNA-modified electrodes to determine whether they were CT deficient relative to wild-type UDG. C17H produced a signal that was larger than that of WT UDG, C101S produced a signal within error of that of WT, and C85S produced a weaker signal than that of WT. These results indicate that changes in the iron-sulfur cluster coordination environment of UDG can drastically perturb the DNA-bound CT properties of this enzyme.

Introduction

Extremophilic species have long been of interest to life sciences research

because of their ability to survive in a variety of harsh environments. Among the many extremophiles that currently receive increased research attention is the hyperthermophilic archaeon *Archaeoglobus fulgidus*, which can survive in geothermal springs and high-temperature oil fields [1, 2]. Adapting to such high temperatures probably requires a range of physiological changes relative to organisms that live under less extreme conditions. These changes may include modified DNA repair pathways to combat excess mutagenesis induced by high temperatures. For example, cytosine deamination occurs at greater frequency at high temperatures [3], and this process converts cytosine to uracil in DNA. Although thermophiles likely endure higher rates of cytosine deamination in their natural environment, those whose mutation rates have been examined do not exhibit higher rates of mutation than non-thermophiles [4].

Researchers who study this phenomenon have pointed focused on Uracil DNA glycosylase (UDG), the enzyme that removes misincorporated uracil from DNA. Many thermophilic UDGs, (which belong to Families IV, V, and VI of the UDGs [5–9]) contain a [4Fe-4S] cluster [5–7, 10, 11], (Figure 2.1). The presence of this iron-sulfur cluster in *A. fulgidus* UDG makes it comparable to other [4Fe-4S] cluster base excision repair (BER) enzymes such as MutY and EndoIII, which our laboratory has also studied [12]. The [4Fe-4S] clusters in these enzymes are redox-active at physiologically relevant potentials only when bound to DNA [12]. The data led us to hypothesize that thermophilic UDGs scan the genome electrochemically in search of DNA damage.

In our model for DNA damage detection, a bound enzyme transmits an electron to another protein bound to a distal location on the DNA. Upon reduction, this "distal" protein loses some of its DNA affinity, dissociates, and re-binds to a different region of the genome. However, if a lesion is present in the DNA sequence between the two bound proteins, CT will not proceed efficiently, both proteins remain oxidized and DNA-bound, and then localize to find and repair the damage between them. This model is supported by the fact that DNA-mediated charge-transfer (CT) can occur efficiently over long distances [13, 14], but is attenuated by mismatches and damage sites in the DNA helix [14, 15]. We have also shown that the affinity of certain BER enzymes for DNA is stronger in their oxidized 3^+ form than in their reduced 2^+ form [16]. Because CT would provide a means of DNA damage detection that is much faster than a processive search [17], perhaps the presence of a [4Fe-4S] cluster enables UDGs from thermophilic organisms to detect DNA damage more efficiently and thus cope with excess, temperature-induced DNA damage.

This model of DNA damage detection assumes that CT proceeds efficiently from the DNA helix to the [4Fe-4S] cluster of a bound enzyme such as UDG. Accordingly, UDG must contain amino acids that ligate the metallocluster, maintain it in a CT-active and stable conformation, and facilitate CT between the DNA and the protein. In order to study which amino acids mediate CT in UDG, several site-directed mutants were prepared and characterized electrochemically.

The *A. fulgidus* UDG mutants examined here, C17H, C85S, and C101S, all

contain mutants at one of the four cysteine residues that ligate the iron-sulfur cluster. In family IV UDG's, the iron-sulfur cluster is ligated by a C-X_2-C-X_n-C-$X_{14\text{-}17}$-C sequence where n=70–100, and the cysteine locations are well conserved [11, 18], (Figure 2.2). These UDG mutants were examined on DNA-modified electrodes, and the signal strength of different mutants' cyclic voltammogram signals were quantified relative to WT protein. An attenuated electrochemical signal indicates a reduced ability to mediate CT. Through these experiments, several mutants of *A. fulgidus* UDG were detected that have CT abilities different than those of WT protein.

Materials and Methods

Oligonucleotide synthesis. The sequences used for quantitative electrochemistry experiments were SH-(C_6H_{12})-(RR)- 5'-GA GAT ATA AAG CAC GCA-3' and complement, SH-(C_6H_{12})- 5'-AGT ACA GTC ATC GCG-3' and complement, and SH-(C_6H_{12})- 5'-AGT ACA GTC ATC GCG-3' paired with RR- 3'- TCA TGT CAG TAG CGC- 5'. "RR" stands for Redmond red, a redox active probe used to quantify the amount of DNA on the electrode surface [19]. All DNA sequences were prepared on an AB 3400 DNA Synthesizer from ABI using standard phosphoramidite chemistry. Phosphoramidites were purchased from Glen Research. Oligonucleotides were purified by HPLC on a C18 reverse phase column and characterized by MALDI-TOF mass spectrometry.

DNA Annealing. For electrochemistry experiments, complementary strands were mixed in a 1:1 ratio in DNA Buffer (10 mM NaPi, 50 mM NaCl, pH = 7.5), heated to 90ºC for several minutes, and then allowed to cool slowly to room temperature over 1–4 hours.

Preparation of DNA-Modified Electrodes. After hybridization to their complement, oligonucleotides duplexes were deposited onto an Au surface in phosphate buffer (50 mM NaCl, 5 mM sodium phosphate, pH 7.0) and incubated for 24–36 hours. Au substrates were purchased from Molecular Imaging. The electrode surface was then washed, and further passivated by incubation with 100 mM mercaptohexanol in phosphate buffer for 30–45 minutes. The electrode surface was washed again in phosphate buffer, and then protein buffer (see below).

Cyclic Voltammetry. Preparation of the DNA-modified Au electrode and subsequent cyclic voltammetry experiments were performed as described above and previously [12, 15, 17, 20–22]. The DNA–modified Au electrode served as the working electrode, a Pt wire served as the auxiliary electrode, and the reference was either an Ag/AgCl electrode modified with an agarose tip or a 66-EE009 Ag/AgCl reference electrode (ESA Biosciences). Unless otherwise noted, all scans were taken at a rate of 50 mV/s on a CH Instruments 760 potentiostat. Each experiment used 50 μL of 50–200 μM of protein in UDG Buffer: (50 mM NaPi pH =7, 50 mM NaCl, 0.5 mM EDTA, 25% glycerol). The protein signal intensity was normalized to DNA concentration using the intensity of the Redmond red signal. For each

comparison, the protein samples were measured consecutively on the same electrode surface.

Protein Purification. UDG was prepared as documented previously [12]. The concentration was quantified using ε_{390}= 20,000 $M^{-1}cm^{-1}$.

Results

Quantitative electrochemistry of BER Enzymes on DNA-modified electrodes. WT UDG, C17H, C85S, and C101S were examined on electrodes prepared with Redmond-red modified DNA. Redmond red is a redox-active probe whose cyclic voltammetry signal intensity can be used to quantify the amount of DNA present on the electrode surface [19]. The Redmond red was positioned so that it could contact the Au surface directly without requiring a DNA-mediated signal. In cyclic voltammograms taken on these electrodes, two peaks are visible, one from Redmond red and one from the protein sample (Figure 2.3). In these experiments, the signal from C17H is consistently larger than that of WT UDG, the signal strength from C101S is within error of that of WT UDG, and the signal from C85S is consistently weaker than that of WT UDG. The signal strength ratios of these variants were 2.2, 0.5, and 1.0, respectively (Table 2.1).

Discussion

All the mutants examined contain substitutions in one of the cysteine residues that ligates the iron-sulfur cluster. Histidine-substituted mutants were

prepared because this residue is a naturally occurring iron-sulfur cluster ligand in some proteins [23]. Substitution of histidine for cysteine did not severely reduce the strength of the mutants' cyclic voltammetry signals relative to wild-type enzyme. As a matter of fact, C17H produced a stronger signal on DNA-modified electrodes than WT UDG, possibly because histidine's aromatic character may render it particularly effective at mediating CT [24]. Serine is not known to occur as a cluster ligand in nature, but perhaps its hydroxyl side chain could substitute for the thiol found in cysteine. C101S, for example, produced a signal strength within error of that of WT UDG. C85S produced an incredibly weak signal relative to wild-type enzyme, although the discrepancy in signal attenuation between the two serine mutants could result from their position within the UDG structure. In the crystal structure of *T. thermophilus* UDG, another thermophilic organism, the residue equivalent to C101 is oriented towards the interior of the protein, making it the least solvent-exposed of the cysteines [11]. Consequently, perturbations at this residue may only minimally affect the stability of the [4Fe-4S] cluster. It is also possible that solvent could ligate the [4Fe-4S] cluster in some of the mutants and/or that minor structural rearrangements could occur that allow other residues to ligate the cluster instead of the native cysteines [25].

These data have identified types of amino acids that can effectively substitute for cysteine in ligating the [4Fe-4S] cluster of UDG and still permit efficient CT. Of the three mutants examined here, the histidine mutant was most likely to be functional and the serine mutants were somewhat functional.

Importantly, though, all of these mutants exhibited a midpoint potential within error of that of WT UDG, which was reported as 95 mV versus NHE [11]. These measurements demonstrate that the [4Fe-4S] cluster is intact in these mutants, and that they are in the range characteristic of high potential iron proteins (HiPIPs) that transition between the 2^+ and 3^+ oxidation states [26]. This potential is also physiologically relevant, suggesting that UDG could participate in electron transfer reactions *in vivo.* These experiments also show that changes in the coordination sphere of the iron-sulfur cluster of UDG can dramatically impact its electrochemical activity. Further work will be necessary to determine whether such mutations undermine the ability of UDG to detect and repair DNA damage inside a cell.

Figure 2.1: Structure of DNA-bound Family V UDG from *T. thermophilus* HB8. A uracil substrate is shown in purple. Adapted from reference [7] and formatted in PyMol [27] PDB File: 2DEM.

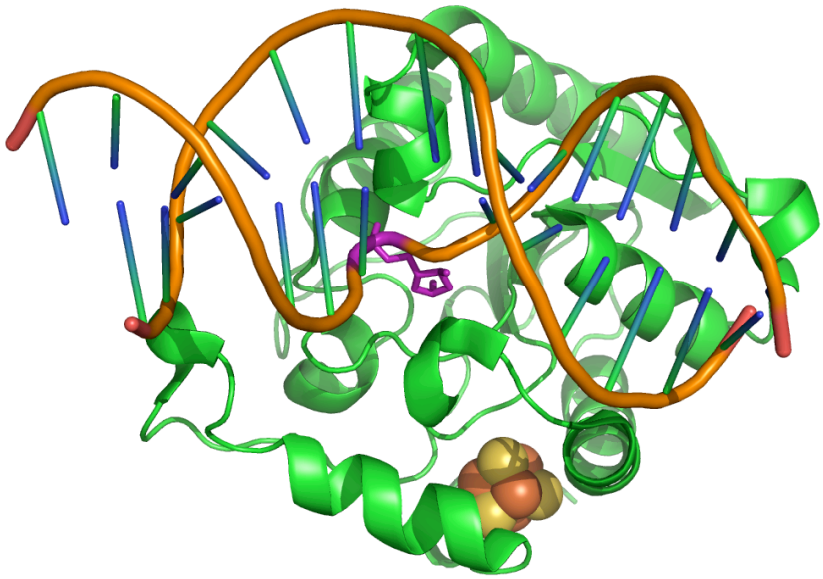


Figure 2.2: Sequence alignment of Family IV UDG proteins. All sequences are from thermophilic organisms. Alignment was made using ClustalW. Conserved cysteine residues are emphasized in orange, other highly conserved residues are highlighted in blue.

```
TThermophilus   -----MTLELLQAQAQNCTACRLMEGRTRVVFGEGNPDAKLMIVGEGPGEEEDKTGRPFV 55
Tmaritima       MYTREELMEIVSERVKKCTACPLHLNRTNVVVGEGNLDTRIVFVGEGPGEEEDKTGRPFV 60
Afulgidus       ----MESLDDIVREIMSCRKCDLHKTKTNYVPGVGNEKAEIVFVGEAPGRDEDLKGEPFV 56
                              *  *

TThermophilus   GKAGQLLNRILEAAGIPREEVYITNIVKCRPPQNRAPLPDEAKICTDKWLLKQIELIAPQ 115
Tmaritima       GRAGMLLTELLRESGIRREDVYICNVVKCRPPNNRTPTPEEQAACG-HFLLAQIEIINPD 119
Afulgidus       GAAGKLLTEMLASIGLRREDVYITNVLKCRPPNNRDPTPEEVEKCG-DYLVRQLEAIRPN 115
                                        *               *

TThermophilus   IIVPLGAVAAEFFL----GEKVSITKVRGKWYE---WHG-IKVFPMFHPAYLLRNPSRAP 167
Tmaritima       VIVALGATALSFFVD---GKKVSITKVRGNPID---WLGGKKVIPTFHPSYLLRNRS--- 170
Afulgidus       VIVCLGRFAAQFIFNLFDLEFTTISRVKGKVYEVERWGKKVKVIAIYHPAAVLYRPQ--- 172

TThermophilus   GSPKHLTWLDIQEVKRALDALPPKERRPVKAVSQEPLF 205
Tmaritima       ---NELRRIVLEDIEKAKSFIK-KEG------------ 192
Afulgidus       -----LREEYESDFKKIGELCGKKQPTLFDYL------ 199
```

Figure 2.3: Cyclic voltammograms of UDG mutants on DNA-modified electrodes. Representative cyclic voltammograms are shown for C17H (orange), C85S (green), and C101S (red). In each image, the scan of WT UDG is also shown (blue), along with a buffer-only control scan (gray).

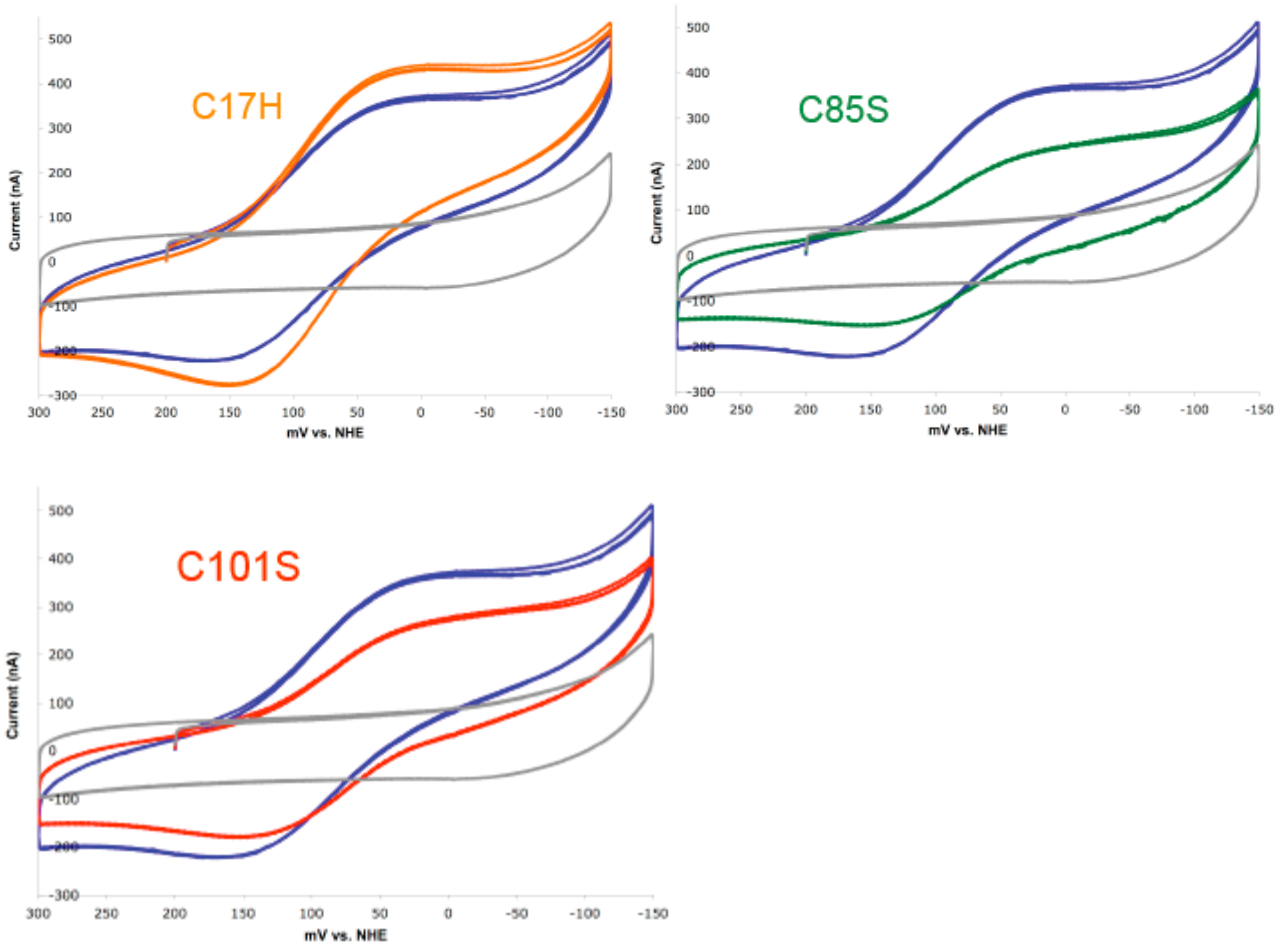


Table 2.1: Quantitative summary of electrochemical data for UDG mutants. Three scans of each mutant were taken. For each sample, WT UDG was measured on the same electrode. The signal strength ratio (area under the cyclic voltammogram curve) was tabulated for each trial.

Trial	Mutant to WT Signal Strength Ratio (RR Corrected)		
	C17H	C85S	C101S
1	3.2	0.5	1.0
2	1.1	0.4	0.9
Average	**2.2**	**0.5**	**1.0**

References:

1. Stetter, K.O., et al., *Hyperthermophilic Archaea Are Thriving in Deep North-Sea and Alaskan Oil-Reservoirs.* Nature, 1993. **365**(6448): 743–745.
2. Klenk, H.P., et al., *The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon Archaeoglobus fulgidus.* Nature, 1997. **390**(6658): 364–370.
3. Lindahl, T., and B. Nyberg, *Heat-Induced Deamination of Cytosine Residues in Deoxyribonucleic-Acid.* Biochemistry, 1974. **13**(16): 3405–3410.
4. Jacobs, K.L., and D.W. Grogan, *Rates of spontaneous mutation in an archaeon from geothermal environments.* Journal of Bacteriology, 1997. **179**(10): 3298–3303.
5. Sakai, T., et al., *Mutagenesis of uracil-DNA glycosylase deficient mutants of the extremely thermophilic eubacterium Thermus thermophilus.* DNA Repair, 2008. **7**(4): 663–669.
6. Hoseki, J., et al., *Crystal structure of a family 4 Uracil-DNA glycosylase from Thermus thermophilus HB8.* Journal of Molecular Biology, 2003. **333**(3): 515–526.
7. Kosaka, H., et al., *Crystal structure of family 5 uracil-DNA glycosylase bound to DNA.* Journal of Molecular Biology, 2007. **373**(4): 839–850.
8. Chung, J.H., et al., *A novel uracil-DNA glycosylase family related to the helix-hairpin-helix DNA glycosylase superfamily.* Nucleic Acids Research, 2003. **31**(8): 2045–2055.
9. Sandigursky, M., and W.A. Franklin, *Thermostable uracil-DNA glycosylase from Thermotoga maritima, a member of a novel class of DNA repair enzymes.* Current Biology, 1999. **9**(10): 531–534.
10. Parikh, S.S., et al., *Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA.* EMBO Journal, 1998. **17**(17): 5214–5226.
11. Hinks, J.A., et al., *An iron-sulfur cluster in the Family 4 uracil-DNA glycosylases.* Journal of Biological Chemistry, 2002. **277**(19):16936–16940.
12. Boal, A.K., et al., *DNA-Bound Redox Activity of DNA Repair Glycosylases Containing [4Fe- 4S] Clusters.* Biochemistry, 2005. **44**:8397–8407.
13. Slinker, J.D., et al., *DNA Charge Transport over 34 nm.* Nature Chemistry, 2011. **3**: 230–235.
14. Nuñez, M.E., D.B. Hall, and J.K. Barton, *Long- range oxidative damange to DNA: effects of distance and sequence.* Chem. Biol., 1999. **6**: 85–97.
15. Boal, A.K., and J.K. Barton, *Electrochemical Detection of Lesions in DNA.* Bioconjugate Chem., 2005. **16**: 312–321.
16. Gorodetsky, A.A., A.K. Boal, and J.K. Barton, *Direct Electrochemistry of Endonuclease III in the Presence and Absence of DNA.* J. Am. Chem. Soc., 2006. **128**:12082–12083.

17. Boal, A.K., et al., *Redox signaling between DNA Repair Proteins for efficient lesion detection.* Proc. Natl. Acad. Sci. USA, 2009. **106**(36): 15237–15242.
18. Lukianova, O.A., and S.S. David, *A role for iron-sulfur clusters in DNA repair.* Current Opinion in Chemical Biology, 2005. **9**(2):145–151.
19. Buzzeo, M.C., and J.K. Barton, *Redmond Red as a Redox Probe for the DNA-Mediated Detection of Abasic Sites.* Bioconjugate Chemistry, 2008. **19**(11): 2110–2112.
20. Boon, E.M., et al., *DNA-mediated charge transport for DNA repair.* Proc. Natl. Acad. Sci. USA, 2003. **100**: 12543–12547.
21. Kelley, S.O., et al., *Single-base mismatch detection based on charge transduction through DNA.* Nucleic Acids Research, 1999. **27**(24): 4830–4837.
22. Kelley, S.O., et al., *Long-Range Electron Transfer through DNA Films.* Angew. Chem. Int. Ed., 1999. **38**(7): 941–945.
23. Volbeda, A., et al., *Crystal structure of the nickel-iron hydrogenase from Desulfovibrio gigas.* Nature, 1995. **373**(6515): 580–587.
24. Stubbe, J., et al., *Radical Initiation in the Class I Ribonucleotide Reductase: Long-Range Proton-Coupled Electron Transfer?* Chemical Reviews, 2003. **103**(6): 2167–2202.
25. Boal, A.K., *DNA-mediated Charge Transport in DNA Repair*, Thesis in Chemistry. 2008, California Institute of Technology: Pasadena, CA.
26. Cowan, J.A. and S.M. Lui, *Structure-function correlations in high-potential iron proteins.* Advances in Inorganic Chemistry, Vol 45, 1998. **45**: p. 313-350.
27. *The PyMOL Molecular Graphics System, Version 1.2r3pre*, Schrödinger, LLC.

Chapter 3

Charge-Transfer Properties of Two Y82 Position Mutations of *Escherichia Coli* MutY

EPR experiments were performed by E. Yavin and E.D.A. Stemp. MutY protein samples were prepared by S. Kundu and X. Xu from the laboratory of Professor S. David (University of California, Davis).

Abstract

MutY from *Escherichia coli* is a [4Fe-4S] cluster containing DNA repair glycosylase that excises adenine mispaired with 8-oxo-7,8-dihydrodeoxyguanine. The human equivalent of MutY, MUTYH, has attracted the attention of cancer researchers because mutations in MUTYH predispose patients towards developing colorectal cancer. One mutation that occurs frequently in cancer patients is Y165C. Y165 in MUTYH aligns with Y82 in *E. coli* MutY, making Y82 position mutants intriguing candidates for experiments that examine the protein's ability to detect and repair DNA damage. Here, we characterized the electrochemical properties of MutY variants Y82C and Y82L. Y82C exhibits a weaker cyclic voltammetry signal on DNA-modified electrodes than wild-type MutY, and it is less able to regenerate a nitroxyl radical signal through DNA-mediated CT, as detected by EPR spectroscopy. By contrast, Y82L performed within error of WT MutY in both sets of experiments. These data demonstrate that Y82 may mediate protein-to-DNA CT *in vivo* in MutY and support the model for DNA-mediated signaling between DNA repair enzymes as a means of DNA damage detection.

Introduction

MutY is a [4Fe-4S] cluster-containing glycosylase that excises adenine mispaired with 8-oxo-7,8-dihydrodeoxyguanine [1, 2]. It is a member of the base excision repair (BER) family of DNA glycosylases [3] and shares structural homology to the helix-hairpin-helix family of DNA-binding proteins [4]. MutY

and related proteins have received much attention recently because mutations in the human version of MutY, MUTYH, are found in many colorectal cancer patients and have been associated with a predisposition towards developing this disease [5–10]. One of the mutations that occurs most frequently in cancer patients is Y165C [7, 8]. Y165 in human MUTYH aligns with Y82 in *E. coli* MutY, making Y82C a relevant mutant for research into how MUTYH Y165C leads to colorectal cancer progression. Y82L was chosen because *E. coli* EndoIII, which is also structurally similar to *E. coli* MutY, contains a leucine residue at the position equivalent to Y82 in MutY. Enzymatic assays of Y82C and Y82L have already been performed and indicate that Y82C exhibits a higher K_D and lower enzymatic turnover rate than WT MutY [9]. Y82L, by contrast, performed within error of WT MutY in both of these assays [9]. Yet, one feature of these mutants that has yet to be examined is their charge-transfer capability.

MutY contains a [4Fe-4S] cluster that is redox-active when bound to DNA [11]. The midpoint potential is within a physiologically relevant range, suggesting that the DNA-bound form of this enzyme could participate in electron-transfer reactions *in vivo* [11]. In our model for DNA damage detection, a bound enzyme could transmit an electron to another protein bound to a distal location on the DNA. Upon reduction, this "distal" protein would lose some of its DNA affinity, dissociate, and re-bind to a different region of the genome. However, if a lesion is present in the DNA sequence between the two bound proteins, then CT will not proceed efficiently, both proteins will remain oxidized and DNA-bound, and they

will localize to the site of damage between them. This model is supported by the fact that DNA-mediated charge-transfer (CT) can occur efficiently over long distances, but is attenuated by mismatches and damage sites in the DNA helix [12, 13]. We have also shown that the affinity of certain BER enzymes for DNA is stronger in their oxidized 3^+ form than in their reduced 2^+ form [14].

For this model of DNA damage detection using DNA-mediated CT to function, CT must proceed efficiently from the DNA helix to the [4Fe-4S] cluster of a bound enzyme. We wondered whether Y82 could mediate DNA-to-protein CT in MutY. In addition to its biomedical relevance, Y82 is a good candidate for CT studies because the equivalent residue in *Bacillus stearothermophilus* MutY has been shown to intercalate into the DNA [15] (Figure 3.1). If the same occurs in *E. coli*, this intercalation would allow residue Y82 to act as an important mediator for CT from the DNA into the MutY protein. Herein, two sets of experiments were performed to test the CT capabilities of MutY Y82C and Y82L.

The first experiment was an EPR-based study that examined how effectively the protein mutants could regenerate a free-radical signal. DNA was prepared that contains a uridine base modified with a stable nitroxyl radical [16]. Following chemical oxidation of the nitroxyl radical to the corresponding N-oxo-ammonium ion using $IrCl_6^{2-}$, addition of MutY or UDG results in the regeneration of the nitroxyl radical, as followed by electron paramagnetic spectroscopy (EPR). This signal regeneration is consistent with electron trapping from the reduced (2+) to the oxidized (3+) form of the cluster. The second set of experiments was

electrochemically based and examined the mutants on a DNA-modified electrode. The signal strength of different mutants' cyclic voltammogram signals were quantified relative to WT protein. An attenuated electrochemical signal indicates a reduced ability to mediate CT. Through EPR and cyclic voltammetry, Y82 of *E. coli* MutY was examined as a facilitator of CT between the DNA and the [4Fe-4S] cluster of MutY.

Materials and Methods

Oligonucleotide synthesis. The sequences used for quantitative electrochemistry experiments were SH-(C_6H_{12})-(RR)- 5'-GA GAT ATA AAG CAC GCA-3' and complement, SH-(C_6H_{12})- 5'-AGT ACA GTC ATC GCG-3' and complement, and SH-(C_6H_{12})- 5'-AGT ACA GTC ATC GCG-3' paired with RR- 3'- TCA TGT CAG TAG CGC- 5'. "RR" stands for Redmond red, a redox active probe used to quantify the amount of DNA on the electrode surface [17]. Synthesis of the spin-labeled deoxyuridine and its incorporation into a 10-mer oligonucleotide (10-mer: 3' –GATG**U***CAGCA-5', where U* = spin-labeled uridine through acetylene linker) was previously reported [16]. Complementary DNA sequences include a 26-mer oligo 3'- CATAGCCGCAATCGCCGACTAGAGCC-5' annealed to a 36-mer oligo 5'- GTATCGGCGTTAGCGGCTGATCTCGGCTACAGTCGT-3'. All DNA sequences were prepared on an AB 3400 DNA Synthesizer from ABI using standard phosphoramidite chemistry. Phosphoramidites were purchased from Glen

Research. Oligonucleotides were purified by HPLC on a C18 reverse phase column and characterized by MALDI- TOF mass spectrometry.

DNA Annealing. For electrochemistry experiments, complementary strands were mixed in a 1:1 ratio in DNA Buffer (10 mM NaPi, 50 mM NaCl, pH = 7.5), heated to 90ºC for several minutes, and then allowed to cool slowly to room temperature over 1–4 hours. For the EPR experiments, DNA annealing was carried out initially by heating to 90ºC a solution containing the 26:36-mer strands in DNA Buffer (10 mM NaPi, 50 mM NaCl, pH = 7.5) at a 1:1 ratio, and then slowly cooling this solution to room temperature over at least 90 minutes. Next, the 10-mer spin labeled DNA (0.9:1 ratio to the 26:36-mer duplex) was added to the preassembled 26:36-mer duplex by heating the solution to 45ºC for two minutes and allowing it to cool to ambient temperature. Duplex formation with the 10-mer spin labeled DNA was verified by EPR spectroscopy.

Preparation of DNA-Modified Electrodes. After hybridization to their complement, oligonucleotides duplexes were deposited onto an Au surface in phosphate buffer (50 mM NaCl, 5 mM sodium phosphate, pH 7.0) and incubated for 24–36 hours. Au substrates were purchased from Molecular Imaging. The electrode surface was then washed, and further passivated by incubation with 100 mM mercaptohexanol in phosphate buffer for 30–45 minutes. The electrode surface was washed again in phosphate buffer, and in protein buffer (see below).

Cyclic Voltammetry. Preparation of the DNA-modified Au electrode and subsequent cyclic voltammetry experiments were performed as described

previously [11, 12, 18–21]. The DNA-modified Au electrode served as the working electrode, a Pt wire served as the auxiliary electrode, and the reference was either an Ag/AgCl electrode modified with an agarose tip or a 66-EE009 Ag/AgCl reference electrode (ESA Biosciences). Unless otherwise noted, all scans were taken at a rate of 50 mV/s on a CH Instruments 760 potentiostat. Each experiment used 50 μL of 50–200 μM of protein in MutY buffer (20 mM sodium phosphate pH=7.5, 100 mM NaCl, 5% glycerol, 1 mM EDTA). For MutY measurements, the protein signal intensity was normalized to DNA concentration using the intensity of the Redmond red signal. For each comparison, the protein samples were measured consecutively on the same electrode surface.

Protein Purification. MutY was prepared as documented previously [22]. The concentration was quantified using ε_{410}= 17,000 $M^{-1}cm^{-1}$.

EPR Spectroscopy. X-band EPR spectra were obtained on an EMX spectrometer (Bruker, Billerica, MA) equipped with a rectangular cavity working in the TE_{102} mode. A quartz flat cell (100 μL) was used in all ambient–temperature experiments. A frequency counter built into the microwave bridge provided accurate frequency values. DNA samples consisted of 26–36 preassembled duplexes (8 μM) and 10-mer spin-labeled DNA (7.2 μM) in 75 μL of buffer (10 mM NaPi, 29 mM NaCl, 2.6 mM $MgCl_2$, pH = 6.2). Chemical oxidation of the spin-labeled probe was then accomplished by addition of 2 μL of $K_2[IrCl_6]$ (2.5 mM in 10 mM NaPi, 50 mM NaCl, pH = 7.5), as determined by a significant attenuation in the EPR signal. After this oxidation, protein was added to the

sample, and the spectrum was immediately recorded (typically after two minutes). EPR parameters were as follows: microwave power = 20 mW, receiver gain = $1x10^4$, and modulation amplitude = 4 G.

EcoRI Control Experiment. A stock solution of EcoRI (40 units/ 1 μL) was dialyzed in protein buffer (100 mM NaCl, 20 mM NaPi, pH=7.5, 1 mM EDTA, 5% glycerol) for four hours at 4°C. Protein activity post-dialysis was tested by incubating dialyzed EcoRI with genomic DNA (extracted from *HeLa* cells) at 37°C for 30 minutes, and then visualizing the digest on an agarose gel relative to cell extracts treated with un-dialyzed EcoRI. Both samples showed similar levels of activity as indicated by comparable DNA smears (data not shown). EcoRI concentration was then determined using a fluorescence-based detection kit (Nano-Orange/ Molecular Probes). An estimated 200 μM stock solution was prepared for the EPR assay.

Results

EPR-based detection of electrons generated from DNA-bound [4Fe-4S] cluster enzymes. The spectra were acquired at room temperature and are shown before and after the addition of the chemical oxidant, $IrCl_6^{2-,}$ and after addition of protein. Upon addition of Ir(IV), the EPR signal decreases substantially, indicating that the spin label was efficiently oxidized from the nitroxyl radical to the EPR silent *N*-oxo-ammonium ion.

Two MutY mutants, Y82C and Y82L, were also examined. Y82C initially

exhibits low signal regeneration relative to WT MutY (Figures 3.2a and 3.2b), although the EPR signal gradually increases to WT level over the course of 28 minutes (Figure 3.3). Y82L, however, regenerated the EPR signal to the same level as WT MutY (Figure 3.4). As a control, these experiments were also performed with EcoRI to verify that the observed EPR signal regeneration originated from the [4Fe-4S] cluster. No apparent signal regeneration was observed after EcoRI was added to the oxidized spin-labeled sample (Figure 3.5).

Quantitative electrochemistry of BER enzymes on DNA-modified electrodes. Quantitative electrochemistry experiments were used to examine the MutY variants Y82C and Y82L. Consistent with EPR results, the cyclic voltammetry signal produced by Y82C was much weaker than that of WT MutY (Figure 3.6a), whereas the signal produced by Y82L was within error of that produced by wild-type MutY (Figure 3.6b). Both mutants produce midpoint potentials within error of WT MutY (Table 3.2).

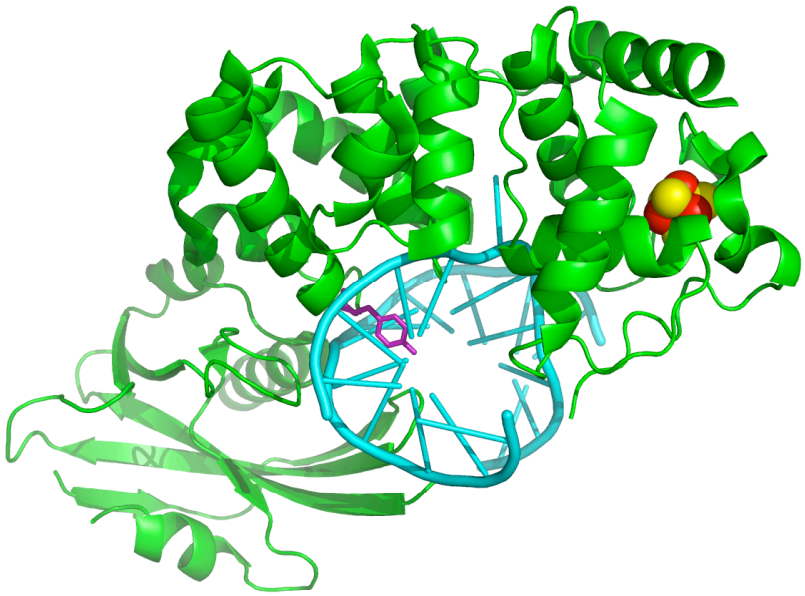
Discussion

The MutY mutants examined were both chosen based on relevance to cancer research, sequence conservation, and the ability to intercalate into the DNA π-stack. Y82C displayed both an attenuated electrochemical signal and a reduced ability to regenerate a nitroxyl radical EPR signal. This result is consistent with the hypothesis that MUTYH Y165C mutants are less able to electrochemically detect and repair DNA damage, thus causing patients with this mutation to develop

colorectal cancer. Y82L produced an electrochemical signal and EPR signal strength within error of that of WT MutY, indicating that the leucine substitution does not dramatically impede DNA-to-protein CT in MutY.

These data on Y82 position mutants establish Y82 as a CT-relevant residue in MutY. They also indicate that bulky residues that can intercalate into the DNA helix, such as tyrosine and leucine, are best able to facilitate CT in this enzyme, whereas shorter side chains, such as the thiol of cysteine, are less capable of mediating CT. It should be noted that although Y82C exhibits a higher K_D than WT MutY, we expect that it bound the DNA just as well as WT during these experiments because the amount of protein used far exceeded the amount of DNA. Future work on this system could include research on other site-directed mutants to detect other segments of the CT pathway between DNA and the iron-sulfur cluster of MutY. If such studies are undertaken, special attention could be given to residues with long side chains that could intercalate into the DNA. Aromatic amino acids are also regarded as good mediators of protein CT [23–26]. The discovery of additional pathway residues would increase support for the hypothesis that MutY detects DNA damage through long-range DNA mediated signaling.

Figure 3.1: Crystal structure of MutY from *Geobacillus stearothermophilus* bound to DNA. Tyr88 (equivalent to Tyr82 in *E. coli* MutY) is highlighted in purple. Mutations at this position in the human equivalent of MutY are associated with a predisposition to colorectal cancer. The sulfur and iron atoms of the [4Fe-4S] cluster are annotated in yellow and red, respectively. Adapted from [15] and formatted in PyMOL (PBD: 1RRQ).



Figures 3.2a and 3.2b: EPR spectra before and after addition of WT-MutY (left, 5.8 μM) **and MutY-Y82C** (right, 5.8 μM). The DNA duplex (6.7 μM) was initially treated with $IrCl_6$ 2 μL (60 μM). Spectra shown are before oxidation (black), after oxidation before protein addition (red), and then four minutes after protein addition (green).

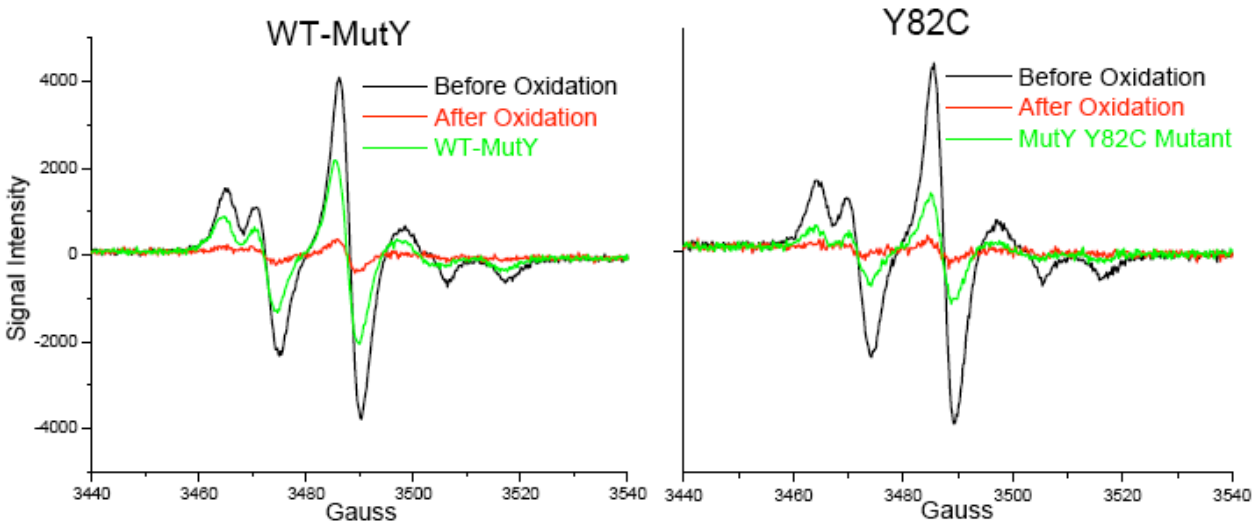


Figure 3.3: EPR spectra before and after addition of MutY-Y82C (5.7 μM) .The DNA duplex (6.7 μM) was initially treated with $IrCl_6$ 2 μL (60 μM). Spectra shown are before oxidation (black), after oxidation before protein addition (red), and then throughout a time course of 28 minutes after protein addition (green to blue).

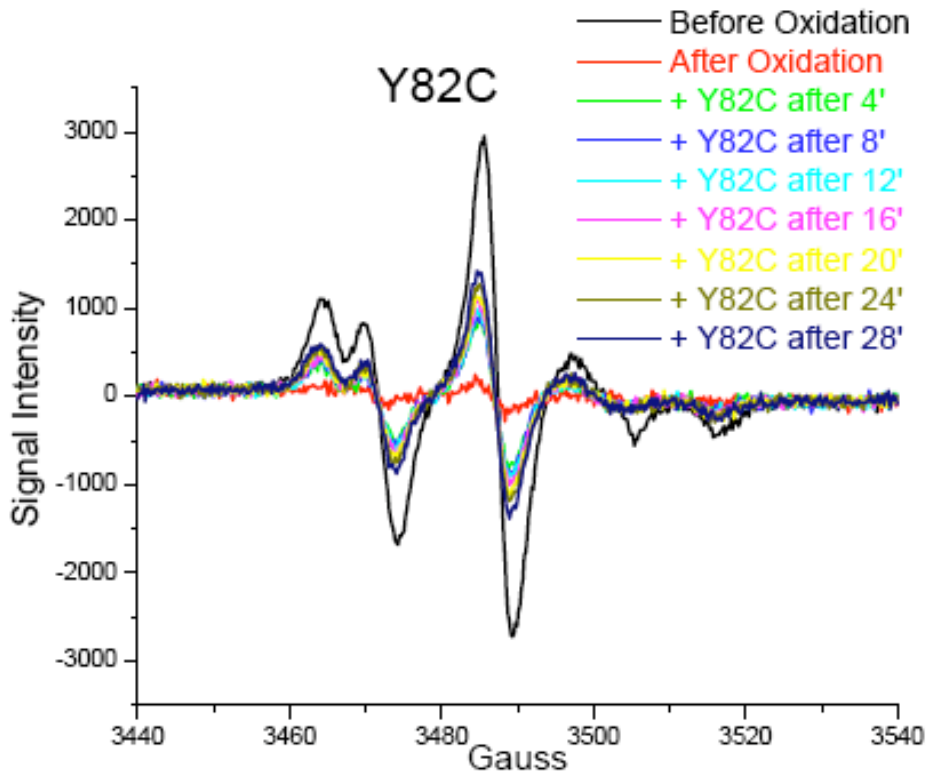


Figure 3.4: EPR spectra before and after addition of MutY-Y82L (5.7 μM). The DNA duplex (6.7 μM) was initially treated with $IrCl_6$ 2 μL (60 μM). Spectra shown are before oxidation (black), after oxidation before protein addition (red), and then after protein addition (green).

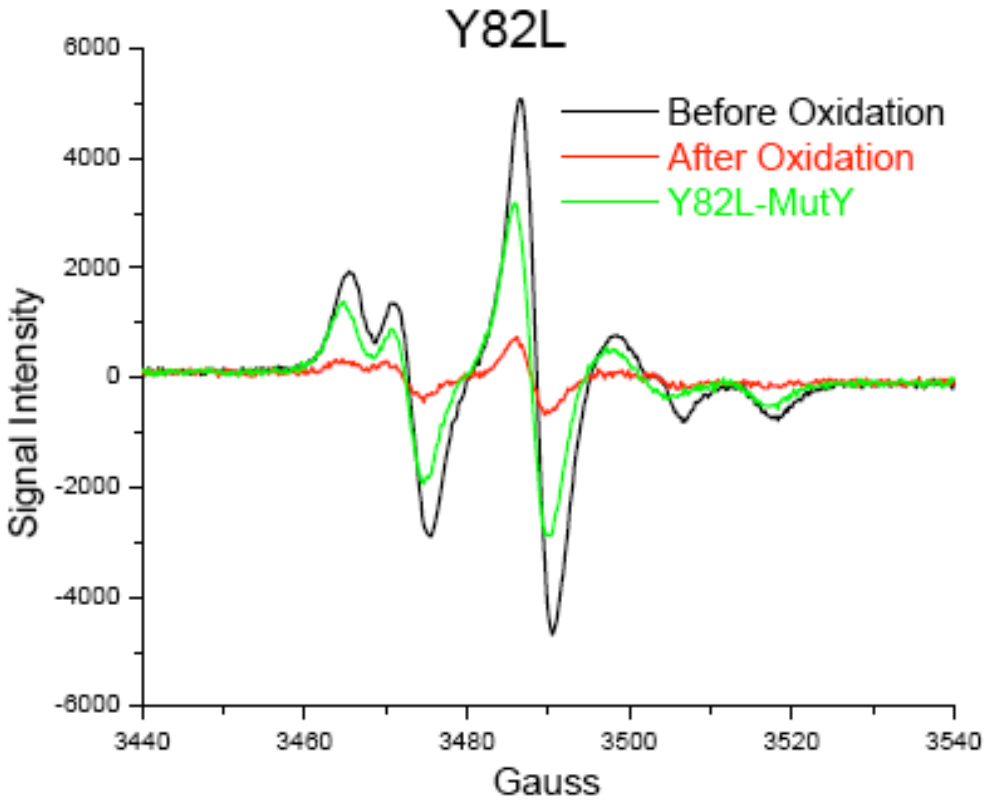


Figure 3.5: EPR spectra before and after addition of EcoRI (13.8 μM). The DNA duplex (6.7 μM) was initially treated with $IrCl_6$ 2 μL (60 μM). Spectra shown are before oxidation (black), after oxidation before protein addition (red), and then after protein addition (green).

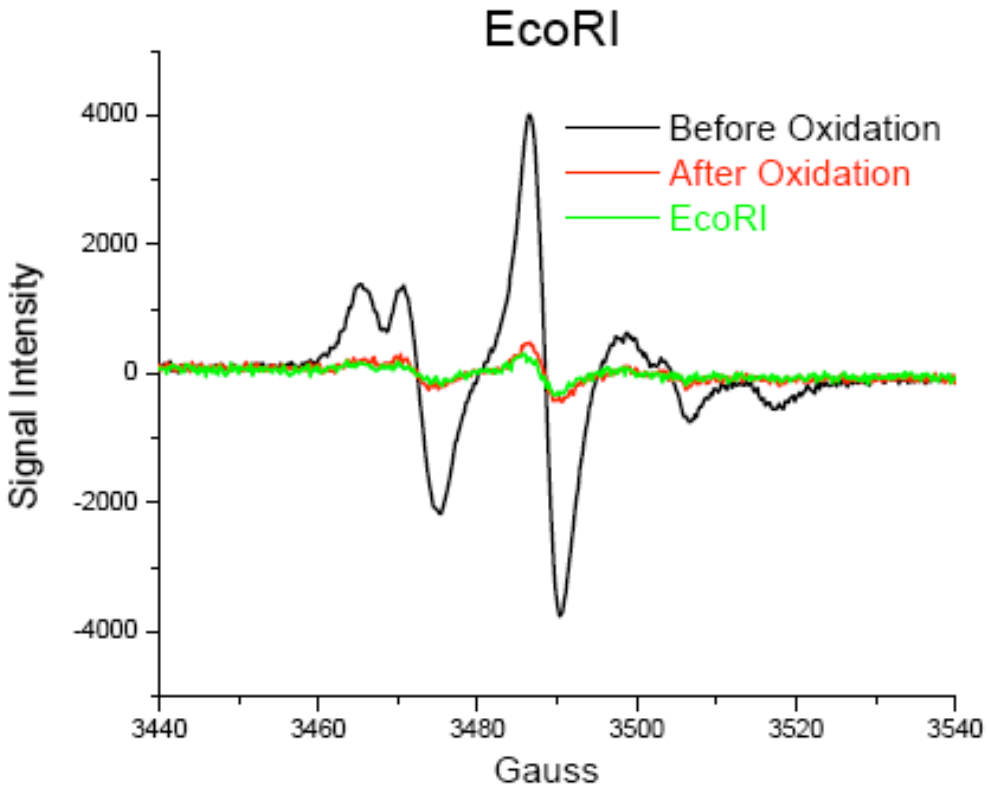


Table 3.1: Summary of the MutY data in terms of signal re-generation. The values are corrected for the non-zero value obtained after treating DNA samples with Ir(IV).

SAMPLE #	1	2	3	AVERAGE
WT-MutY	45[a]	45	47	46±1
Y82C (4')	27	27	23	26±2
Y82C (30')	39	42	42	41±2
Y82L	44	69	48	54±11

[a] Signal regeneration is based on the ratio of main EPR signal at g = 2.0 ($\Delta_{\text{peak to peak}}$) between the EPR signal observed after protein addition to that of untreated DNA.

Figure 3.6: Electrochemistry of MutY variants. The cyclic voltammograms of the MutY variants Y82C (pale blue) and Y82L (dark green) are shown relative to those of WT MutY (burgundy). The scan in buffer is also pictured (gray). Y82C is charge-transfer deficient relative to WT MutY and Y82L produces a signal within error of that of WT.

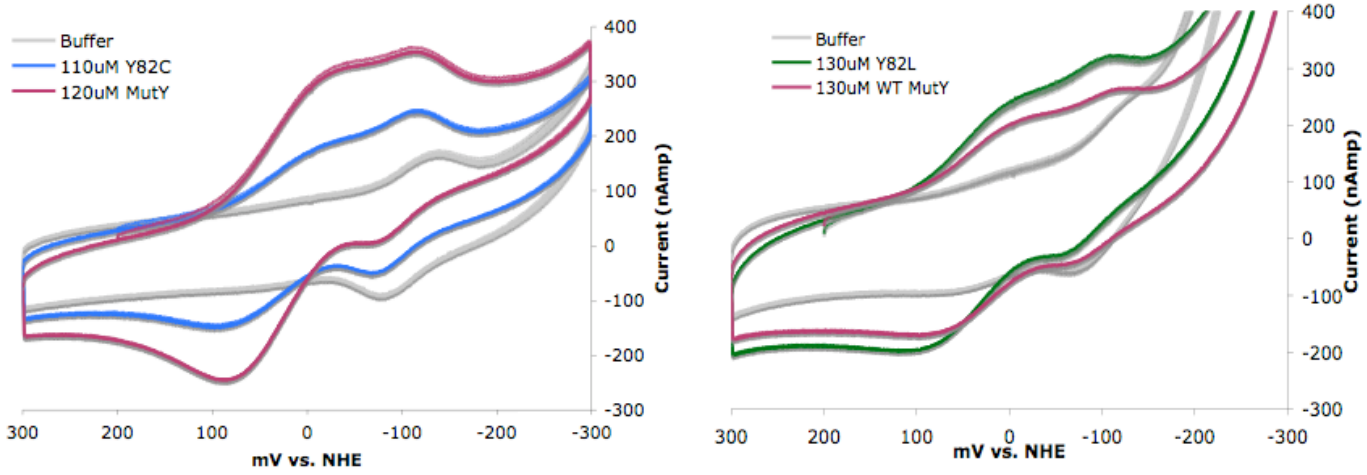


Table 3.2: Summary of quantitative electrochemical measurements of MutY variants

Variant	Midpoint Potential (mV vs. NHE)	Signal strength relative to WT (RR corrected)
Y82C	88 ± 8	0.25 ± 0.20
Y82L	88 ± 15	0.94 ± 0.17

References:

1. Nghiem, Y., et al., *The Muty Gene—a Mutator Locus in Escherichia-Coli That Generates G.C→T.A Transversions.* Proceedings of the National Academy of Sciences of the United States of America, 1988. **85**(8): 2709–2713.
2. Michaels, M.L., et al., *A Repair System for 8-Oxo-7,8-Dihydrodeoxyguanine.* Biochemistry, 1992. **31**(45):10964–10968.
3. David, S.S., and S.D. Williams, *Chemistry of Glycosylases and Endonucleases Involved in Base Excision Repair.* Chemical Reviews, 1998. **98**: 1221–1261.
4. Doherty, A.J., L.C. Serpell, and C.P. Ponting, *The helix-hairpin-helix DNA-binding motif: A structural basis for non-sequence-specific recognition of DNA.* Nucleic Acids Research, 1996. **24**(13): 2488–2497.
5. Bai, H.B., et al., *Functional characterization of human MutY homolog (hMYH) missense mutation (R231L) that is linked with hMYH-associated polyposis.* Cancer Letters, 2007. **250**(1):74–81.
6. Bai, H.B., et al., *Functional characterization of two human MutY homolog (hMYH) missense mutations (R227W and V232F) that lie within the putative hMSH6 binding domain and are associated with hMYH polyposis.* Nucleic Acids Research, 2005. **33**(2):597–604.
7. Cheadle, J.P. and J.R. Sampson, *MUTYH-associated polyposis—From defect in base excision repair to clinical genetic testing.* DNA Repair, 2007. **6**: 274–279.
8. Chmiel, N.H., A.L. Livingston, and S.S. David, *Insight into the Functional Consequences of Inherited Variants of hMYH Adenine Glycosylase Associated with Colorectal Cancer: Complementation Assays with hMYH Variants and Pre-stead-state Kinetics of the Corresponding Mutated E. Coli Enzymes.* J. Mol. Biol., 2003. **327**:431–443.
9. Livingston, A.L., et al., *Insight into the Roles of Tyrosine 82 and Glycine 253 in Escherichia coli Adenine Glycoslyase MutY.* Biochemistry, 2005. **44**:14179–14190.
10. Al-Tassan, N., et al., *Inherited variants of MYH associated with somatic G:C→T:A mutations in colorectral tumors.* Nature Genetics, 2002. **30**:227-232.
11. Boal, A.K., et al., *DNA-Bound Redox Activity of DNA Repair Glycosylases Containing [4Fe- 4S] Clusters.* Biochemistry, 2005. **44**:8397–8407.
12. Boal, A.K. and J.K. Barton, *Electrochemical Detection of Lesions in DNA.* Bioconjugate Chem., 2005. **16**: p. 312-321.
13. Nunez, M.E., D.B. Hall, and J.K. Barton, *Long-range oxidative damange to DNA: effects of distance and sequence.* Chem. Biol., 1999. **6**:85–97.

14. Gorodetsky, A.A., A.K. Boal, and J.K. Barton, *Direct Electrochemistry of Endonuclease III in the Presence and Absence of DNA.* J. Am. Chem. Soc., 2006. **128**: 12082–12083.
15. Fromme, J.C., et al., *Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase.* Nature, 2004. **427**: 652-656.
16. Yavin, E., et al., *Electron trap for DNA-bound repair enzymes: A strategy for DNA-mediated signaling.* Proc. Natl. Acad. Sci. USA, 2006. **103**(10): 3610–3614.
17. Buzzeo, M.C., and J.K. Barton, *Redmond Red as a Redox Probe for the DNA-Mediated Detection of Abasic Sites.* Bioconjugate Chemistry, 2008. **19**(11): 2110–2112.
18. Boal, A.K., et al., *Redox signaling between DNA Repair Proteins for efficient lesion detection.* Proc. Natl. Acad. Sci. USA, 2009. **106**(36): 15237–15242.
19. Boon, E.M., et al., *DNA- mediated charge transport for DNA repair.* Proc. Natl. Acad. Sci. USA, 2003. **100**: 12543–12547.
20. Kelley, S.O., et al., *Single-base mismatch detection based on charge transduction through DNA.* Nucleic Acids Research, 1999. **27**(24): 4830–4837.
21. Kelley, S.O., et al., *Long-Range Electron Transfer through DNA Films.* Angew. Chem. Int. Ed., 1999. **38**(7): 941–945.
22. Chmiel, N.H., et al., *Efficient recognition of substrates and substrate analogs by the adenine glycosylase MutY requires the C-terminal doman.* Nucleic Acids Research, 2001. **29**:553–564.
23. Shih, C., et al., *Tryptophan-accelerated electron flow through proteins.* Science, 2008. **320**(5884): 1760–1762.
24. Stubbe, J., et al., *Radical Initiation in the Class I Ribonucleotide Reductase: Long-Range Proton-Coupled Electron Transfer?* Chemical Reviews, 2003. **103**(6):2167–2202.
25. Giese, B., et al., *Electron Relay Race in Peptides.* Journal of Organic Chemistry, 2009. **74**(10):3621–3625.
26. Giese, B., M. Graber, and M. Cordes, *Electron transfer in peptides and proteins.* Curr. Opin. Chem. Biol., 2008. **12**(6): p. 755-759.

Chapter 4

Charge-Transfer Mutants of the DNA-Bound Glycosylase, Endonuclease III

H.B. Gray and D.K. Newman offered helpful advice. T.J. Ge provided technical assistance and E.D. Olmon helped prepare figures. The plasmid used for overexpression of EndoIII variants was generously donated by the laboratory of Professor S. Mayo (Caltech).

Abstract:

Endonuclease III (EndoIII) is a base excision repair (BER) glycosylase that targets oxidatively damaged pyrimidines and contains a [4Fe-4S] cluster. We have proposed that the [4Fe-4S] cluster is used in DNA-mediated signaling for DNA repair. Here, several mutants of *E. coli* EndoIII were prepared to probe efficiency of DNA/protein charge transport (CT). Cyclic voltammetry experiments on DNA-modified electrodes show that aromatic residues F30, Y55, Y75, and Y82 help mediate CT between DNA and the [4Fe-4S] cluster in EndoIII. At position 82, an interface between protein and DNA, aromatic residues best facilitate CT, while a serine substitution inhibits protein/DNA CT. Based on circular dichroism studies to measure protein stability, mutations at residues W178 and Y185 are found to destabilize the protein; these residues may function to protect the [4Fe-4S] cluster. These results suggest a pathway for protein/DNA CT and support the model for DNA-mediated signaling between DNA repair proteins as a means of DNA damage detection.

Introduction

First isolated from *Escherichia coli* and described in 1989 [1], Endonuclease III (EndoIII) is an enzyme that excises oxidatively damaged pyrimidines from DNA. EndoIII and MutY, a closely related enzyme that removes adenine mispaired with 8-oxo-7,8-dihydrodeoxyguanine [2, 3], belong to the base excision repair (BER) family of DNA repair proteins, whose members

repair DNA by excising damaged nucleobases from the DNA backbone. EndoIII and MutY homologues are found in many species, ranging from bacteria to humans, and their structural biology and enzymatic properties have been studied in great detail by both crystallography [4–9] and mutagenesis-based enzymology [8, 10–18]. MutY has also drawn attention from the biomedical community because patients with mutations in its human homologue, MUTYH, are predisposed towards developing colorectal cancer [19–23].

Despite these extensive studies, two features of the BER family of enzymes remain intriguing to consider. First, what is the role of the [4Fe-4S] cluster? The [4Fe-4S] cluster does not catalyze base excision in either enzyme, nor is it necessary for MutY to fold properly [11], although it is necessary for MutY to bind DNA. Secondly, how are these proteins able to scan the vast amount of DNA present in a cell in order to detect their substrate lesions? Both enzymes are present in low copy number *in vivo* (~ 30 copies per cell of MutY and ~ 500 copies per cell of EndoIII) [24], making a scan of the genome an unacceptably slow process through a processive mechanism [25].

In response to these questions, our laboratory has proposed a model in which MutY and EndoIII cooperatively scan the genome for DNA lesions using their [4Fe-4S] clusters to participate in DNA-mediated redox signaling [25–28]. DNA mediates CT very efficiently due to π-stacking, or the overlap of aromatic nucleotides whose many delocalized π electrons permit fast electron flow [29–31]. Proteins also mediate CT through a variety of mechanisms [32–37].

These CT mechanisms could permit charge to travel from the DNA into a bound protein, and then through the protein to a redox-active cofactor, such as a [4Fe-4S] cluster.

Both DNA and proteins act as CT mediators in our model for DNA damage detection by DNA-mediated CT. In this model, one protein bound to DNA could transmit an electron through the DNA to a distally bound protein, thus reducing the recipient protein. Reduction decreases the affinity of the protein for DNA [38], so this second protein would dissociate upon reduction and bind to a different location of the genome. However, damaged bases attenuate CT through DNA [39, 40]. Consequently, if DNA between the two proteins is damaged, then charge-transfer is impaired, causing both proteins to remain oxidized and DNA-bound, eventually localizing to the site of damage so that they can repair the lesion. In support of this model, we have found that MutY and EndoIII display physiologically relevant midpoint potentials only when bound to DNA [27, 38]. These protein redox signals are attenuated on electrodes modified with DNA containing an abasic site [27], establishing that the reaction is DNA-mediated. DNA CT can also occur over long distances [29, 41], suggesting that proteins could use it to scan large regions of a genome in order to detect lesions and repair DNA on a reasonable timescale.

One question that remains is how charge is transported from the DNA, through the protein, to the [4Fe-4S] cluster of MutY or EndoIII. Specifically, which amino acids in these proteins facilitate CT between the DNA helix and the

protein metal cluster? One EndoIII mutant, Y82A, has already been shown to be deficient in CT activity, implicating Y82 as a crucial residue in the CT pathway of EndoIII [25]. The current study builds on this result by characterizing several novel mutants of EndoIII to discover other residues that mediate protein-to-DNA CT. Several categories of mutants were prepared, including five additional mutants at or near the Y82 position, to better assess the types of amino acid side chains that are CT proficient in EndoIII. Mutations were also made targeting other aromatic residues in the protein because aromatic amino acids have been shown to mediate CT in other systems [33, 34, 36, 37]. A third category of mutations were made near the [4Fe-4S] cluster, assuming that amino acids close to this redox-active moiety would likely lie along the CT pathway. All of the mutants were characterized by cyclic voltammetry on DNA-modified electrodes, glycosylase assays, and circular dichroism (CD) spectroscopy. These experiments have identified several amino acids that facilitate CT between the DNA helix and the [4Fe-4S] cluster.

Materials and Methods

Preparation of DNA and Protein Samples for Electrochemistry Experiments. Redmond Red (RR) is a redox active probe used to quantify the amount of DNA on the electrode surface [42]. The sequences used for electrochemistry experiments were SH-(C_6H_{12})-(RR)-5'-GA GAT ATA AAG CAC GCA-3' and complement, and SH-(C_6H_{12})-5'-AGT ACA GTC ATC GCG-

3’ annealed to 5’-CGC GAT GAC TGT ACT-3’-RR. Redmond-Red was connected to the oligonucleotide via a pyrrolidinyloxy linker. Redmond Red-labeled DNA was prepared according to established methods [41]. Briefly, all DNA sequences were prepared using standard phosphoramidite chemistry on an Applied Biosystems 3400 DNA synthesizer. Phosphoramidites were purchased from Glen Research. For thiolated strands, the 5′ end was modified with the Thiol Modifier C6 S-S phosphoramidite using standard protocols from Glen Research, Inc. DNA modified with Redmond Red on the 3′ terminus was prepared on Epoch Redmond Red CPG columns from Glen Research with ultramild phosphoramidites and reagents. DNA modified with Redmond Red attached to the thiol was prepared with Epoch Redmond Red phosphoramidite from Glen Research with ultramild phosphoramidites and reagents.

Deprotection, purification, chemical modification of DNA, annealing, and preparation of the DNA-modified Au electrode were performed as described previously [43–46]. Au substrates were purchased from Agilent Technologies. Each experiment used 50 μL of 50–100 μM EndoIII in protein buffer (20 mM NaH_2PO_4 pH 7.5, 100 mM NaCl, 5% glycerol, 1 mM EDTA). The protein concentration was quantified using ε_{410}= 17,000 $M^{-1}cm^{-1}$ [47].

Cyclic Voltammetry. All cyclic voltammetry experiments were performed as described previously [25, 27]. A Au on mica (Molecular Imaging) electrode was assembled and incubated with thiol-modified DNA duplex for 24–36 hours. The electrode was then backfilled with 1 mM mercaptohexanol,

rinsed in DNA buffer (50 mM NaCl, 5 mM NaH_2PO_4 pH 7.5), and then rinsed in protein buffer before protein samples were added to the surface. The DNA-modified Au electrode served as the working electrode, a Pt wire served as the auxiliary electrode, and the reference was either a Ag/AgCl electrode modified with an agarose tip or a 66-EE009 Ag/AgCl reference electrode (ESA Biosciences). All scans were taken at a rate of 50 mV/s on a CH Instruments 760 potentiostat. To determine the relative CT efficiencies of EndoIII mutants, the signal intensity of these proteins was normalized to DNA concentration using the intensity of the Redmond Red signal. For each comparison, the protein samples were measured consecutively on the same electrode surface.

Preparation of EndoIII Over-Expression Construct. The *nth* gene that encodes EndoIII was cloned using the Failsafe Enzyme with Buffer G (Epicentre Biotechnologies) and plasmid pBBR1MCS-4 [48] which contains the *nth* gene as a template. (This construct was originally prepared by cloning the *nth* gene from *Escherichia coli* chromosomal DNA). Primers used were 5'- CGC CCGCG GTGGT ATG AAT AAA GCA AAA GCG CTG– 3' and 5'- CGC GGATCC TCA GAT GTC AAC TTT CTC TTT- 3'. PCR products were purified on a 1% agarose gel and then excised using the QIAquick gel extraction kit (Qiagen). These fragments were ligated into the BamHI and SacII sites of the pET11-ubiquitin-His vector, which expresses hexahistidine and ubiquitin tags at the N-terminus of the insertion site. This vector was derived from the pET11 vector (Novagen), but the ubiquitin gene and hexahistidine tag were engineered into the

vector by the laboratory of Professor Stephen Mayo, from whom this vector was a kind donation. The ligation reaction was used to transform both DH5α *E. coli* (Invitrogen) for sub-cloning, and BL21star(DE3)pLysS *E. coli* (Invitrogen) for over-expression. Transformants resistant to ampicillin and chloramphenicol were selected and the plasmids were isolated using the QIAPrep MiniPrep kit (Qiagen). After the size of the insert was verified by restriction digestion, plasmids from positive transformants were submitted for sequencing (Laragen). Freezer stocks of transformants containing the correct EndoIII sequence were prepared and stored at -80°C.

Purification of EndoIII and Mutants. Freezer stocks of BL21star(DE3)pLysS containing pET11-ubiquitin-His with the *nth* gene were used to inoculate 10 mL of LB media containing 100 μg/mL ampicillin and 30 μg/mL chloramphenicol. Each 10 mL culture was incubated at 37°C overnight and then used to inoculate 1 L of LB/ampicillin/chloramphenicol. Each 1 L flask was grown to OD_{600} = 0.6 at 37°C and then isopropyl β-D-1-thiogalactopyranoside was added to a total concentration of 0.3 mM. Cells were incubated at 30°C for 3.5 hours and harvested by centrifugation (5000 rpm, 10 minutes). Pellets were stored at -80°C. For lysis, pellets were dissolved in 25 mL lysis buffer (50 mM Tris-HCl, pH 8, 5% glycerol, 250 mM NaCl, 5 mM DTT, 1 mM phenylmethylsulfonyl fluoride) and lysed by sonication. Cell lysate was fractionated by centrifugation (7000 rpm, 10 minutes). The supernatant was filtered, and then loaded onto a 5 mL HisTrap HP column (GE Healthcare) pre-

equilibrated with binding buffer (20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole, 1 mM DTT, pH 7.4) at a flow rate of ~ 1 mL/min. The column was washed with binding buffer until the UV baseline was stable. The fusion protein was eluted using a gradient of 0–100% elution buffer (20 mM sodium phosphate, 0.5 M NaCl, 500 mM imidazole, 1 mM DTT, pH 7.4) over 10–20 column volumes on an AKTA FPLC (GE Healthcare). Fractions containing protein (determined by their yellow color) were pooled and loaded onto a Superose 12 column (GE Healthcare) that had been pre-equilibrated with Superose 12 Buffer (50 mM NaH_2PO_4, 0.15 M NaCl, pH 7.0). The tagged EndoIII, His_6-ubiquitin-EndoIII, eluted after ~ 12 mL. Fractions containing EndoIII were pooled and dialyzed overnight into protein storage buffer (20 mM sodium phosphate pH 7.5, 0.5 mM EDTA, 100 mM NaCl, 20% glycerol). Aliquots of 50–70 μL EndoIII were prepared, frozen in dry ice, and stored at -80°C. Protein purity was assessed by SDS-PAGE, and concentration was measured using ε_{410}= 17,000 $M^{-1}cm^{-1}$ [47].

Site-Directed Mutagenesis. Mutations were encoded on the pET-11 based EndoIII overexpression plasmid using the Quikchange Site-Directed Mutagenesis Kit (Stratagene). Primers used are listed in the supporting material, where uppercase letters indicate the encoded mutation. DNA primers were purchased from Integrated DNA Technologies. All mutagenized plasmids were sequenced (Laragen) to verify accurate mutagenesis before being used for protein over-expression.

Instrumentation. All protein samples were examined using a Beckman DU 7400 spectrophotometer. Protein concentration was measured using ε_{410} = 17,000 $M^{-1}cm^{-1}$ [47].

Circular Dichroism. Samples were measured at a concentration of ~ 5 μM at ambient temperature on a Model 62A DS Circular Dichroism Spectrometer (AVIV, Lakewood, NJ). For thermal denaturation experiments, the temperature was varied from 20°C to 60°C. Each trace shown is the average of at least three independent experiments.

DNA Glycosylase Assay. This protocol was adapted from methods described previously [10]. DNA strands synthesized were 5'- TGT CAA TAG CAA GXG GAG AAG TCA ATC GTG AGT CT- 3' where X = 5-hydroxyuracil and the complement with G opposite X. DNA was prepared using standard phosphoramidite chemistry using reagents purchased from Glen Research. Prior to annealing, single-strand DNA was purified using reversed-phased HPLC and the substrate-containing strand was 5'- radiolabeled using ^{32}P-ATP [49] with polynucleotide kinase purchased from Roche.

Glycosylase activity was determined by monitoring nick formation in the hydroxyuracil-containing duplex using denaturing gel electrophoresis [12]. For this assay, 100 nM radiolabeled duplex was incubated with EndoIII or a variant at a range of concentrations (10 nM, 100 nM or 1 μM) in reaction buffer (10 mM Tris pH = 7.6, 1 mM EDTA, 50 mM NaCl) for 15 minutes at 37°C. The reactions were quenched upon addition of NaOH to a final concentration of 100 nM.

Samples were dried, counted by scintillation, and diluted with loading buffer (80% formamide, 10 mM NaOH, 0.025% xylene cyanol, 0.025% bromophenol blue in Tris-Borate-EDTA buffer) to normalize the radioactivity. Samples were then heated at 90°C for five minutes prior to loading and then separated by denaturing PAGE. Glycosylase activity was determined by quantifying the amount of 14-mer product visualized in the gel relative to the total amount of DNA present.

Results

Electrochemistry on DNA-modified electrodes. The CT capabilities of several mutants of EndoIII were investigated on a DNA-modified electrode surface passivated with mercaptohexanol (Figure 4.1a). This strategy has been used previously to measure the DNA-bound electrochemical properties of proteins containing [4Fe-4S] clusters [25, 27]. In this study, the thiolated DNA duplex was also modified with Redmond Red, a redox probe whose signal intensity can be used to measure the amount of DNA on a given electrode surface [42]. The Redmond Red was positioned so that it could contact the Au surface and be directly reduced without requiring DNA-mediated reduction.

For each EndoIII sample, the electrochemical signal grew in over 30–45 minutes. Scan rate dependence measurements showed a linear relationship between the peak current of the protein and the square root of scan rate, indicative of a diffusion-limited process [50]. These same experiments showed a linear

relationship between the peak current of Redmond Red and scan rate, indicative of a surface-bound species. Figures 4.1b and 4.1c show typical cyclic voltammograms, with two peaks evident, one from the [4Fe-4S] cluter of the DNA-bound protein and one from the Redmond Red probe. When compared on these surfaces, all mutants exhibited midpoint potentials of roughly 80 (±30) mV versus NHE, and all were within error of that of WT (Table 4.1). While the DNA-bound potential for these mutants was seen to be invariant, significant variations in redox signal intensity were observed (Table 4.1 and Figure 4.1d). The signal strength of each mutant was quantified and normalized based on the amount of DNA present on the electrode. To compare the signal strength of each mutant relative to that of the WT, these signals were further normalized by the Redmond-Red-corrected signal strength of WT EndoIII measured on the same surface.

Among the first samples to be examined in this fashion were mutants relevant to colorectal cancer research, those at positions L81 and Y82. Y82F and Y82W exhibit signal strengths within error of that of WT, suggesting that other aromatic residues at position 82 confer CT capabilities equivalent to those of the native tyrosine. By contrast, Y82S exhibits a very weak electrochemical signal relative to WT (Figure 1b, 1d). The signal intensity from Y82C is on average the same as that of WT, although highly variable, possibly due to the ability of cysteine to facilitate electron transfer in certain contexts [32, 51, 52], albeit through the formation of an unstable radical.

After the experiments with Y82 mutants established the importance of aromatic residues for DNA/protein CT, other aromatic amino acids in EndoIII were targeted for mutagenesis studies. F30A, Y55A, Y75A, and H140A were examined. The first three of these mutants all displayed a CT deficiency relative to WT EndoIII (Figure 4.1d). In the protein structure, Y75, Y55, and F30 are relatively close to Y82, and form a line along one side of the protein [5, 7, 9]. These residues may comprise a pathway of aromatic amino acids through which electrons can travel efficiently. Such aromatic "π-ways" have been found in other peptides [36, 37]. The H140A mutant, by contrast, proved CT-proficient in the experiments performed here so no conclusions can be made about its electrochemical properties. However, this residue was found to be enzymatically important (Table 4.2).

The final category of EndoIII mutants examined were W178A and Y185A, substitutions involving aromatic residues close to the [4Fe-4S] cluster. These mutants were expected to produce weak electrochemical signals because of their proximity to the metallocluster, but they instead produced signals that were large and highly variable relative to WT (Figures 4.1c, 4.1d). One explanation for this phenomenon is protein aggregation, as denatured samples can produce large and erratic signals on DNA-modified electrodes (data not shown). However, because the midpoint potentials of these mutants are within error of those produced by other EndoIII samples, it is unlikely that the [4Fe-4S] cluster degraded in W178A and Y185A, given that [3Fe-4S] clusters have different

potentials [53–55] than the [4Fe-4S] cluster. These results with W178A and Y185A prompted further experiments, described below, that suggest these residues may help protect the [4Fe-4S] cluster from exposure to solvent.

Enzymatic assays of mutants' glycosylase activity. The glycosylase activity of each mutant was measured according to established methods [10, 25]. Briefly, 35-mer strands of DNA were synthesized with 5-hydroxyuracil (5-OH-dU), an EndoIII substrate analogue, incorporated into the sequence. The sequence was 5'-radiolabeled with ^{32}P-ATP and annealed to its complement. Solutions of 100 nM DNA were incubated with 1 μM enzyme. Active enzymes will remove 5-OH-dU from the DNA backbone, leaving an abasic site whose phosphodiester bonds are cleaved with 1 M NaOH. This cleavage produces a 14-mer strand, which can be visualized by denaturing PAGE and autoradiography (Figure 4.2). The amount of 14-mer relative to the total quantity of DNA was used to assess enzymatic activity. Mutants that are glycolytically active are able to bind DNA as well as WT. Consequently, this experiment is an important complement to the electrochemical studies, as it verifies that any weak cyclic voltammetry signals produced by mutants stem from a CT deficiency and not an inability to bind DNA.

Of the eleven mutants examined, most were found to exhibit glycosylase activity within error of that of WT EndoIII. The exceptions were F30A and H140A. F30A was slightly impaired in its glycolytic activity, possibly because F30 is located near the substrate binding pocket of EndoIII [5, 7, 9], so mutations

at this position may impede DNA binding. H140A is deficient in glycosylase activity relative to WT (Table 2). In earlier crystallographic studies, the H140 residue was identified as possibly helping EndoIII bind to DNA by providing a positively charged histidine residue that could help coordinate the negatively charged DNA helix [7]. The enzymatic activity results presented here are consistent with that hypothesis.

Circular dichroism to examine the structural stability of EndoIII mutants. While most of the mutants examined gave electrochemical signals with intensities weaker than or within error of those of WT EndoIII, the mutants W178A and Y185A were exceptions to this trend. Consequently, CD experiments were performed to examine the secondary structure of W178A and Y185A. The CD spectra of WT EndoIII and several mutants were examined for both fully folded and thermally denatured protein (Figure 4.3). All folded samples exhibited similar spectral shapes.

In order to examine the stability of W178A and Y185A relative to WT EndoIII, thermal denaturation experiments were performed. In these experiments, the protein sample was heated gradually and its ellipticity was measured as a function of temperature. The ellipticity of all variants measured decreased with increasing temperature, but W178A and Y185A denatured at a lower temperature than WT EndoIII and a CT proficient mutant, Y82F (Figure 4.3). These data indicate that W178A and Y185A are less stable than WT EndoIII.

Discussion

Biochemical characterization of EndoIII Mutants. Eleven mutants of EndoIII were prepared and characterized biochemically in order to identify residues important for protein/DNA CT. They were selected based on several criteria including sequence conservation, aromatic character, and relevance to colorectal cancer research. Additionally, mutations were made in several different regions of the protein to ascertain CT characteristics of different domains. When assayed by cyclic voltammetry on DNA-modified electrodes, all of the mutants display midpoint potentials within error of that of WT EndoIII, indicating an intact [4Fe-4S] cluster, irrespective of the mutation. Thus our focus shifted to monitoring changes in signal intensity as a function of mutation, reflecting differences in coupling along the path for DNA/protein CT. Glycosylase activity was also measured to assay enzymatic activity and, importantly, confirm DNA-binding.

Variations at the protein-DNA interface. L81 and Y82 are both residues whose MUTYH equivalents exhibit mutations in cancer patients [21], and Y82 is well conserved. Additionally, L81 and Y82 are extremely close to the DNA interface in bound EndoIII. This proximity is important when the CT properties of DNA are considered. DNA is able to mediate CT by means of π-stacking, or the overlap of aromatic nucleotides whose many π-bonds provide an efficient conduit for electron flow [28–30, 56–60]. L81 and Y82 may be able to intercalate into the DNA helix [4]. L81 and Y82 would thus join the network of π-bonds in

DNA that facilitate electron flow. Consequently, L81 and Y82 were seen as important residues for mediating CT. Earlier research from our laboratory confirmed this hypothesis by establishing that Y82A is CT deficient relative to WT EndoIII [25]. To follow up on these experiments, additional mutants were made at L81 and Y82 to assess, in more detail, the characteristics of CT-proficient amino acids in EndoIII. L81C is CT-proficient, so no definite conclusions can be made about the electrochemical properties of L81. The mutant Y82S is electrochemically deficient relative to WT EndoIII, whereas Y82F and Y82W are CT proficient, indicating that aromatic amino acids are important mediators of CT at the Y82 position.

Consideration of pathways for DNA-mediated CT. Crystal structures of DNA-bound EndoIII place the [4Fe-4S] cluster and the DNA at roughly 14 Å apart at their closest [5, 7, 9]. At this distance, CT could proceed by either a single-step tunneling process or a multi-step tunneling process, also called hopping, in which amino acids act as "stepping stones" for the electron as it travels through the protein [33, 34, 61–63]. In hopping systems, many of the amino acids through which electrons hop are aromatic residues [34, 61]. With this information in mind, additional mutations were made at aromatic residues closer to the interior of EndoIII to find residues that could be part of a CT pathway. Emphasis was placed on aromatic residues near Y82, as this tyrosine has already been established as CT-active. F30, Y55, and Y75 all display weaker electrochemical signals than WT EndoIII, suggesting they participate in a CT

pathway. The distances covered are 9 Å from Y82 to Y75, 8 Å from Y75 to Y55, and 6 Å from Y55 to F30, which are reasonable distances to consider in the context of hopping [33]. Time-resolved spectroscopy will be necessary to distinguish these pathways. Tunneling efficiency is exponential with distance [62], so even a slight reduction in tunneling distance can increase CT rates significantly.

Beyond F30, the remainder of the CT pathway was not probed for several reasons. First, many residues in the region between F30 and the [4Fe-4S] cluster are part of the helix-hairpin-helix motif of EndoIII, a well-conserved motif in DNA-binding proteins [9, 64] that helps these proteins interact with their substrates. Mutations in this region could render the protein unstable and/or unable to bind DNA. Second, mutations in this region of the protein may not exhibit a detectable CT deficiency even if the targeted residues are able to facilitate CT, because multiple CT pathways could be present, as is often the case in redox-active proteins [63]. In ribonucleotide reductase, for example, electrons are known to find alternative pathways between two redox centers if a primary pathway is impeded by mutagenesis or incorporation of a non-natural amino acid [34, 35]. Similar phenomena could occur in EndoIII, especially given that all of the mutants examined display at least a partial signal, and that the degree of attenuation lessens as the mutations are made further from the DNA. The signal from Y82A is 50 % ± 13 % that of WT [25], whereas the signal from F30A is 88 % ± 3 % that of WT. These values suggest that protein/DNA CT may involve

Y82, but that it could circumvent F30. Third, mutations very close to the [4Fe-4S] cluster could compromise the stability of this cluster, as may have been the case with the mutants W178A and Y185A.

Characterization of EndoIII mutants with large electrochemical signals. After establishing that mutations in aromatic residues are more likely to impede CT in EndoIII, the mutants W178A and Y185A were expected to produce weak cyclic voltammetry signals. However, their CT activity is unexpectedly high, larger even than that of WT EndoIII. This increase in signal strength may be attributed to structural changes near the [4Fe-4S] cluster of the protein. The loss of the bulky, aromatic residues could increase the conformational flexibility of EndoIII, allowing water molecules to access the protein and accelerate CT to the [4Fe-4S] cluster [65, 66].

One question that could be raised by this result is why alanine is not present at positions 178 and 185 in the native protein since it permits more efficient CT than tryptophan and tyrosine at these positions. A possible answer to this question is presented by the circular dichroism data. W178A and Y185A are found to be less stable structurally than WT, and the resulting instability or flexibility could affect the nearby [4Fe-4S] cluster. Aromatic residues could play a protective role as well. If the [4Fe-4S] cluster decomposes, as could happen as a result of oxidative stress [12, 53], then this cluster could degrade to a [3Fe-4S] cluster. Given the proximity of W178 and Y185 to the [4Fe-4S] cluster of EndoIII, it is also possible that these aromatic residues shield the cluster from

solvent exposure under physiological conditions. However, these mutants still exhibit WT level enzymatic activity, indicating that the structural perturbations are probably localized.

Relevance to cancer research. Another noteworthy aspect of this study is the relevance of these EndoIII in the context of colorectal cancer and DNA repair. EndoIII bears structural similarity MUTYH. When MUTYH acquires certain mutations, it is no longer able to repair damage in the adenomatous polyposis coli (APC) gene, which regulates the proliferation of colonic cells [21, 67]. APC is thought to be particularly susceptible to MUTYH activity because it contains 216 "GAA" codons in which a G:C → T:A transversion would result in a stop codon. Genes associated with other cancers do not contain as many such sites [21]. Several MUTYH mutations have been detected in colorectal cancer patients, including those at positions Y114 and Y166 [21]. These residues align with F30 and Y82 in EndoIII, respectively. Since certain mutations at Y82 are deficient in CT activity but not enzymatic activity, perhaps the MUTYH mutants are similarly deficient in CT. According to our BER search model, these mutants would be less effective at detecting DNA damage, allowing mutations to accumulate in the APC gene that could lead to colorectal cancer.

EndoIII mutants and the model of DNA-mediated CT for DNA repair. Previous work in our laboratory has developed a model in which EndoIII and other redox-active DNA-bound proteins use DNA-mediated signaling to search the genome for lesions. Here, this model was further investigated by examining

specific amino acids in EndoIII that could mediate CT between the DNA and the [4Fe-4S] cluster. Four residues are found to be part of this CT pathway (F30, Y55, Y75, and Y82). Of the remaining residues studied, one is found to be important for enzymatic activity (H140) and two are found that contribute to protein stability (W178, Y185). In total, these mutants help demonstrate that EndoIII contains a well protected [4Fe-4S] cluster to which electrons travel from the DNA helix using a series of amino acids. In partially elucidating a pathway through which charge can travel between the DNA and the [4Fe-4S] cluster, these data further support the possibility that DNA-bound proteins communicate *in vivo* by means of DNA-mediated long-range signaling.

Table 4.1: DNA-bound electrochemistry of EndoIII mutants [a]

EndoIII	Midpoint potential (mV vs. NHE)[b]	Mutant / WT Ratio (RR corrected)
WT	78	
F30A	97	0.9 ± 0.03
Y55A	97	0.7 ± 0.1
Y75A	88	0.6 ± 0.2
L81C	84	1.1 ± 0.2
Y82C	92	0.8 ± 0.6
Y82F	91	1.2 ± 0.4
Y82S	91	0.4 ± 0.2
Y82W	92	1.4 ± 0.4
H140A[c]	71	1.0 ± 0.2
W178A	83	3.5 ± 2.3
Y185A	85	2.2 ± 1.2

[a] Experimental conditions are described in materials and methods. Each experiment used 50 μL of 50–100 μM EndoIII in protein buffer (20 mM NaH_2PO_4 pH 7.5, 100 mM NaCl, 5% glycerol, 1 mM EDTA).

[b] Measurements have an uncertainty of ± 30 mV.

[c] This sample was not measured using Redmond-Red-modified DNA, although extra trials were performed to verify the midpoint potential and signal strength ratio that are shown.

Table 4.2: Summary of glycosylase assay results with EndoIII mutants [a]

EndoIII	% Activity relative to WT at 1 μM concentration
F30A	93.7 ± 2.6
Y55A	96.7 ± 1.0
Y75A	99.8 ± 4.0
L81C	99.2 ± 0.6
Y82C	99.2 ± 0.5
Y82F	99.4 ± 1.1
Y82S	99.4 ± 1.6
Y82W	98.1 ± 3.6
H140A	39.1 ± 6.8
W178A	98.2 ± 1.7
Y185A	97.3 ± 0.6

[a] Experimental conditions were as described in materials and methods. Experiments were conducted using 1 μM protein and 100 nM annealed duplex in reaction buffer (10 mM Tris pH = 7.6, 1 mM EDTA, 50 mM NaCl) for 15 minutes at 37°C. The reactions were quenched upon addition of NaOH to a final concentration of 100 nM. Samples were dried, counted by scintillation, and diluted with loading buffer (80% formamide, 10 mM NaOH, 0.025% xylene cyanol, 0.025% bromophenol blue in Tris-Borate-EDTA buffer) to normalize the radioactivity. Samples were then heated at 90°C for five minutes prior to loading and then separated by denaturing PAGE. Glycosylase activity was determined by comparing the amount of 14-mer produced to the total amount of DNA. The glycosylase activity of each mutant as a percentage of WT is shown. Most mutants have activity within error of that of WT.

Table 4.3: Circular dichroism melting temperatures of select EndoIII variants [a]

Sample	T_m (C°)
WT	48 ± 1.8
Y82F	47 ± 0.2
W178A	42 ± 0.3
Y185A	43 ± 0.5

[a] Experimental conditions are described in materials and methods. To obtain melting temperatures, the spectra were fitted to a sigmoidal curve using Origin software, (OriginLab, Northampton, MA) and melting temperatures were obtained from these midpoints.

Table 4.4: EndoIII mutagenesis primers

Mutations are emphasized in bold, capital lettering.

Mutant	Primers
F30A	5'- ccgagcttaatttcagttcgcct**GCT**gaattgctgattgccgtactgc- 3' (forward) 5'- gcagtacggcaatcagcaattc**AGC**aggcgaactgaaattaagctcgg- 3' (reverse)
Y55A	5'- gcgacggcgaaactc**GCC**ccggtggcgaatacgcctgcagc -3'(forward) 5'- gctgcaggcgtattcgccaccgg**GGC**gagtttcgccgtcgc –3'(reverse)
Y75A	5'-gaaggggtgaaaacc**GCT**atcaaaacgattgggctttataacagc-3' (forward) 5'-gctgttataaagcccaatcgttttgat**AGC**ggttttcacccccttc-3' (reverse)
L81C	5'- gggtgaaaacctatatcaaaacgattggg**TGT**ataacagcaaagc- 3' (forward) 5'- gctttgctgttata**ACA**cccaatcgttttgatataggttttcaccc- 3' (reverse)
Y82C	5'-ggtgaaaacctatatcaaaacgattgggctt**TGT**aacagcaaagc-3' (forward) 5'-gctttgctgtt**ACA**aagcccaatcgttttgatataggttttcacc-3' (reverse)
Y82F	5'-ggtgaaaacctatatcaaaacgattgggctt**TTT**aacagcaaagc-3' (forward) 5'-gctttgctgtt**AAA**aagcccaatcgttttgatataggttttcacc-3' (reverse)
Y82S	5'- ggggtgaaaacctatatcaaaacgattgggctt**TCT**aacagcaaagc – 3' (forward) 5'- gctttgctgttagaaagcccaatcgttttgatataggttttcacccc-3' (reverse)
Y82W	5' - ggggtgaaaacctatatcaaaacgattgggctt**TGG**aacagcaaagc – 3' (forward) 5'-gctttgctgttccaaagcccaatcgttttgatataggttttcacccc-3' (reverse)
H140A	5'- ccgactattgctgtcgacacg**GCC**attttccgcgtttgtaatcg- 3' (forward) 5'- cgattacaaacgcggaaaat**GGC**cgtgtcgacagcaatagtcgg- 3' (reverse)
W178A	5'-gtcgactgccaccat**GCG**ttgatcctgcacgggcg-3' (forward) 5'- cgcccgtgcaggatcaa**CGC**atggtggcagtcgac- 3' (reverse)
Y185A	5'-cctgcacgggcgt**GCT**acctgcattgcccgcaagccccgc-3' (forward) 5'-gcggggcttgcgggcaatgcaggt**AGC**acgcccgtgcagg-3' (reverse)

Figure 4.1: Quantitative cyclic voltammetry of EndoIII variants on DNA-modified electrodes. A) Schematic of DNA-modified electrochemistry. A Au surface is treated with thiol-modified DNA and then backfilled with mercaptohexanol. The covalent Redmond Red redox probe, shown as a red oval, is used to quantify the amount of DNA on the surface. Each protein mutant was allowed to bind DNA on such a surface, and the [4Fe-4S] cluster was measured by cyclic voltammetry. The protein was then rinsed from surface and WT was measured on the same surface. B) Representative cyclic voltammogram of Y82S (brown) compared to WT EndoIII (blue). Y82S exhibits a weaker electrochemical signal. For reference, a scan taken in the absence of protein (buffer only) is also shown (gray). C) Representative cyclic voltammogram of W178A (orange) compared to WT EndoIII (blue) and a buffer-only scan (gray). D) The CT capability of each mutant was quantified based on the area under its redox peak. Each mutant's signal intensity was normalized to the intensity of the signal for WT measured on the same surface.

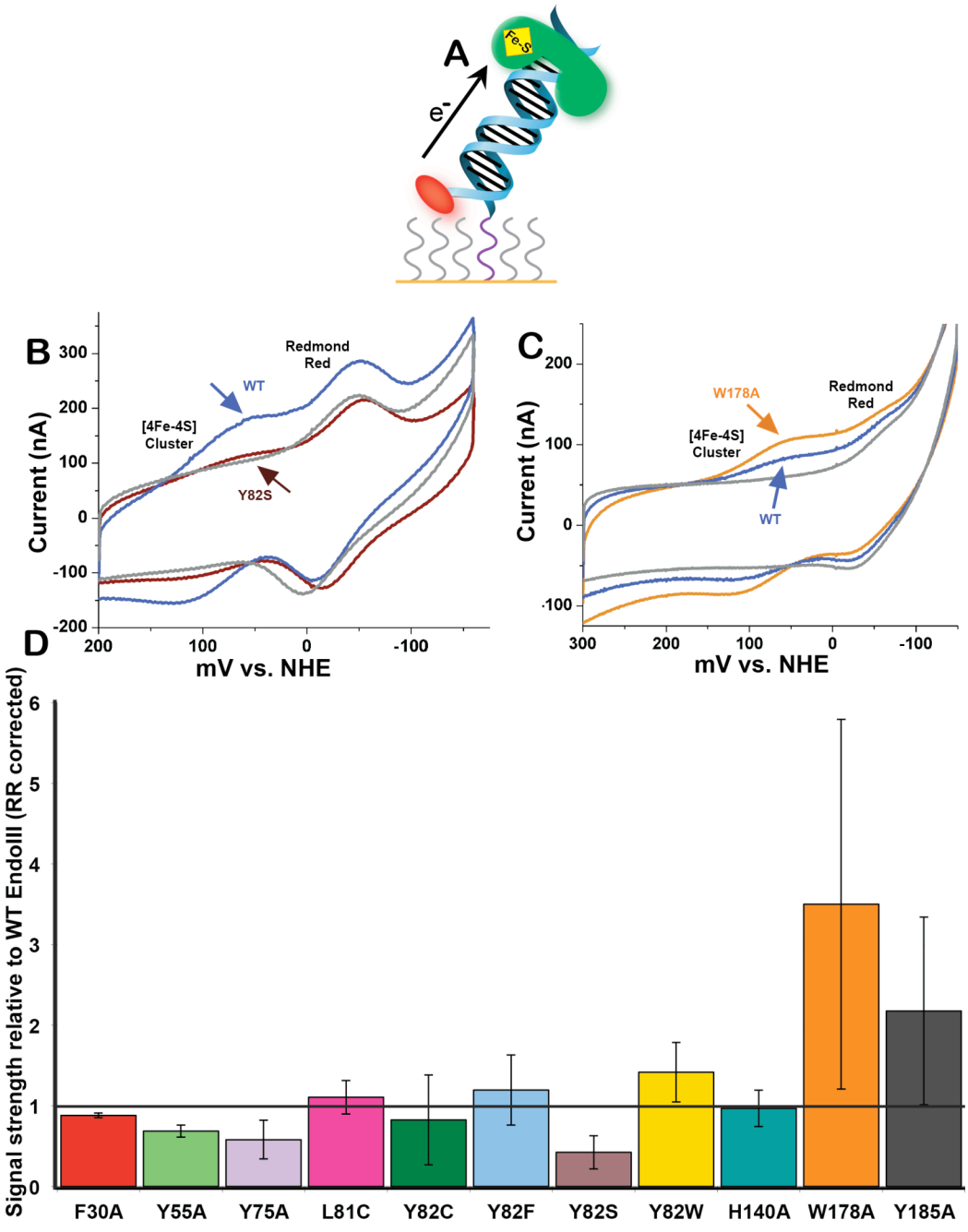

A
e-
Fe-S
B
Current (nA)
mV vs. NHE
Redmond Red
WT
[4Fe-4S] Cluster
Y82S
C
W178A
Redmond Red
[4Fe-4S] Cluster
WT
D
Signal strength relative to WT EndoIII (RR corrected)
F30A
Y55A
Y75A
L81C
Y82C
Y82F
Y82S
Y82W
H140A
W178A
Y185A

Figure 4.2: Glycosylase assay results of EndoIII variants. For this assay, a 1 μM, 100 nM, or 10 nM protein sample was incubated with 100 nM of 5'-radiolabeled 35-mer duplex DNA containing 5-hydroxy-uracil. Reactions were incubated for 15 minutes at 37°C and quenched with 1 M NaOH. Reactions were then examined by denaturing gel electrophoresis. Cleavage of the DNA results in a 14-mer. The representative gel below shows the results of WT EndoIII (blue) compared to Y75A (purple) and Y185A (black). Enzyme-free control reactions were loaded into the final two lanes. The reactions with 1 μM protein were used to compare the glycosylase activities of EndoIII variants (Table 4.2).

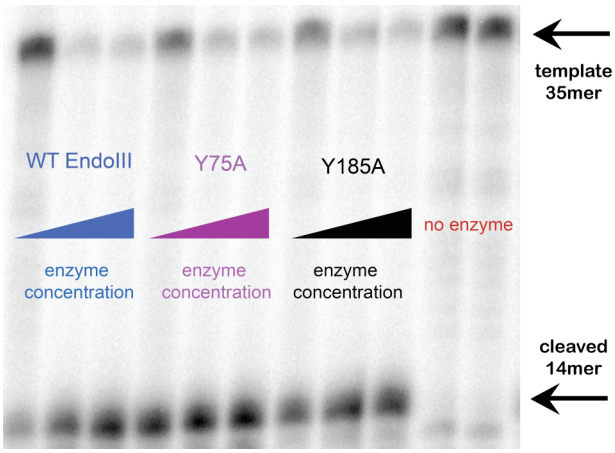


Figure 4.3: Circular dichroism spectra of EndoIII. A) Comparison of fully folded and denatured spectra of EndoIII. Samples were measured at room temperature at a concentration of ~ 5 μM. WT spectra are shown, although all mutants displayed a similar trend with denatured samples (gray) consistently displaying less ellipticity than fully folded samples (blue). The spectra shown are the average of three trials. B) Circular dichroism thermal denaturation of select EndoIII variants. Samples were measured at concentrations of ~ 5 μM. Each spectrum shown is the average of at least three independent experiments. The spectra show the fractional change in ellipticity for each variant measured, where the fully folded protein was given a value of 0 and the denatured form was given a value of 1, consistent with previous secondary structure studies [11]. The variants WT (dark blue) and Y82F (pale blue) display more structural stability than W178A (orange) and Y185A (black).

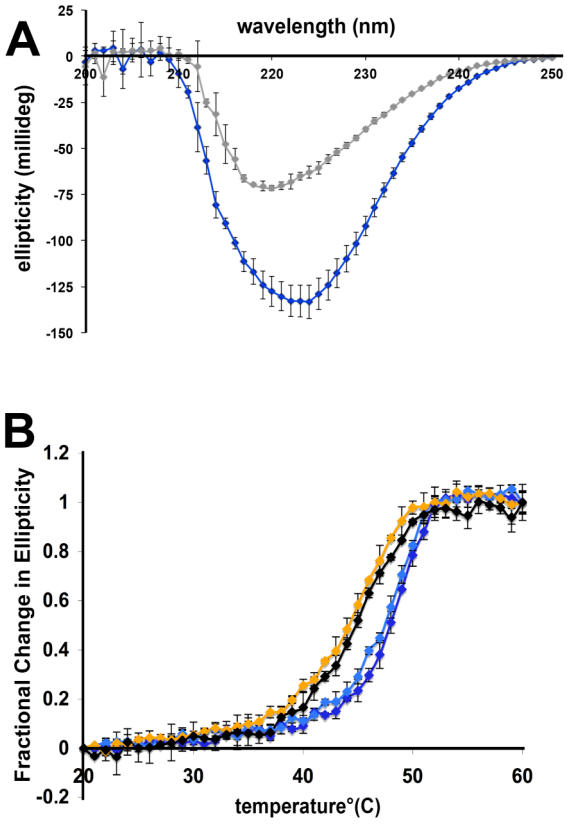
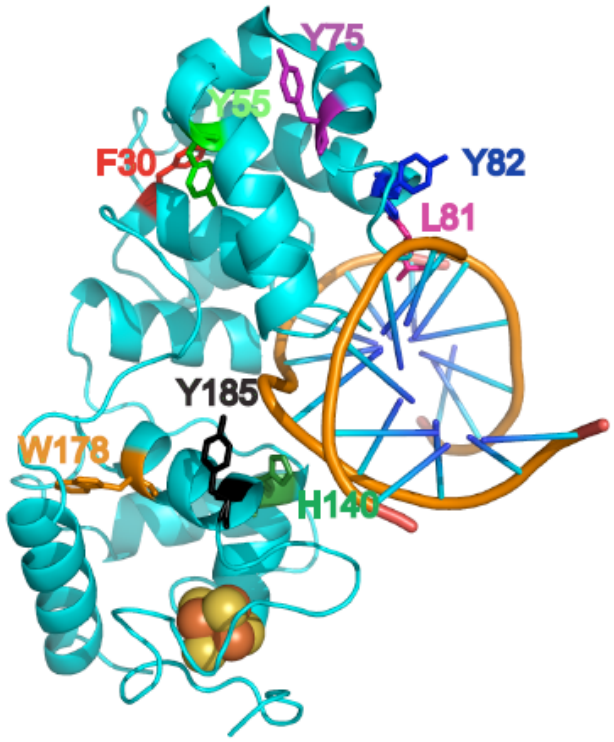


Figure 4.4: Crystal structure of DNA-bound EndoIII. Crystal structure was adapted from references [5, 9] and formatted using PyMol [69]. Residues targeted for mutagenesis studies are emphasized in color.



References:

1. Asahara, H., et al., *Purification and Characterization of Escherichia-Coli Endonuclease-III from the Cloned nth Gene*. Biochemistry, 1989. **28**(10): 4444–4449.
2. Nghiem, Y., et al., *The Muty Gene—a Mutator Locus in Escherichia-Coli That Generates G.C→T.A Transversions*. Proceedings of the National Academy of Sciences of the United States of America, 1988. **85**(8): 2709–2713.
3. Michaels, M.L., et al., *A Repair System for 8-Oxo-7,8-Dihydrodeoxyguanine*. Biochemistry, 1992. **31**(45): 10964–10968.
4. Fromme, J.C., et al., *Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase*. Nature, 2004. **427**:652–656.
5. Fromme, J.C., and G.L. Verdine, *Structure of a trapped endonuclease III-DNA covalent intermediate*. The EMBO Journal, 2003. **22**(13):3461–3471.
6. Guan, Y., et al., *MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily*. Nature Structural Biology, 1998. **5**:1058–1064.
7. Kuo, C.-F., et al., *Atomic Structure of the DNA Repair [4Fe-4S] Enzyme Endonuclease III*. Science, 1992. **258**(5081):434–440.
8. Manuel, R.C., et al., *Reaction Intermediates in the Catalytic Mechanism of Escherichia coli MutY DNA Glycosylase*. Journal of Biological Chemistry, 2004. **279**: 46930–46939.
9. Thayer, M.M., et al., *Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure*. The EMBO Journal, 1995. **14**(16): 4108–4120.
10. Watanabe, T., et al., *Engineering Functional Changes in Escherichia coli Endonuclease III Based on Phylogenetic and Structural Analyses*. Journal of Biological Chemistry, 2005. **280**(40): 34378–34384.
11. Porello, S.L., M.J. Cannon, and S.S. David, *A Substrate Recognition Role for the [4Fe-4S]2+ Cluster of the DNA Repair Glycosylase MutY*. Biochemistry, 1998. **37**: 6465–6475.
12. Messick, T.E., et al., *Noncysteinyl Coordination to the [4Fe-4S]$^{2+}$ Cluster of the DNA Repair Adenine Glycosylase MutY Introduced via Site-Directed Mutagenesis. Structural Characterization of an Unusual Histidinyl-Coordinated Cluster*. Biochemistry, 2002. **41**: 3931–3942.
13. Livingston, A.L., et al., *Insight into the Roles of Tyrosine 82 and Glycine 253 in Escherichia coli Adenine Glycoslyase MutY*. Biochemistry, 2005. **44**: 14179–14190.
14. Golinelli, M.-P., N.H. Chmiel, and S.S. David, *Site-Directed Mutagenesis of the Cysteine Ligands to the [4Fe- 4S] Cluster of Escherichia coli MutY*. Biochemistry, 1999. **38**: 6997–7007.

15. Chmiel, N.H., et al., *Efficient recognition of substrates and substrate analogs by the adenine glycosylase MutY requires the C-terminal domain*. Nucleic Acids Research, 2001. **29**: 553–564.
16. Bai, H.B., et al., *Functional characterization of two human MutY homolog (hMYH) missense mutations (R227W and V232F) that lie within the putative hMSH6 binding domain and are associated with hMYH polyposis*. Nucleic Acids Research, 2005. **33**(2): 597–604.
17. Bai, H.B., et al., *Functional characterization of human MutY homolog (hMYH) missense mutation (R231L) that is linked with hMYH-associated polyposis*. Cancer Letters, 2007. **250**(1): 74–81.
18. Chepanoske, C.L., et al., *A Residue in MutY Important for Catalysis Identified by Photocross-Linking and Mass Spectrometry*. Biochemistry, 2004. **43**: 651–662.
19. Chmiel, N.H., A.L. Livingston, and S.S. David, *Insight into the Functional Consequences of Inherited Variants of hMYH Adenine Glycosylase Associated with Colorectal Cancer: Complementation Assays with hMYH Variants and Pre-stead-state Kinetics of the Corresponding Mutated E. Coli Enzymes*. J. Mol. Biol., 2003. **327**: 431–443.
20. Al-Tassan, N., et al., *Inherited variants of MYH associated with somatic G:C → T:A mutations in colorectral tumors*. Nature Genetics, 2002. **30**: 227- 232.
21. Cheadle, J.P., and J.R. Sampson, *MUTYH-associated polyposis—From defect in base excision repair to clinical genetic testing*. DNA Repair, 2007. **6**: p. 274- 279.
22. Sampson, J.R., et al., *MutYH (MYH) and colorectal cancer*. Biochemical Society Transactions, 2005. **33**: 679–683.
23. Kundu, S., et al., *Adenine removal activity and bacterial complementation with the human MutY homologue (MUTYH) and Y165C, G382D, P391L and Q324R variants associated with colorectal cancer*. DNA Repair (Amst), 2009. **8**(12):1400–1410.
24. Demple, B., and L. Harrison, *Repair of Oxidative Damage to DNA—Enzymology and Biology*. Annual Review of Biochemistry, 1994. **63**: 915–948.
25. Boal, A.K., et al., *Redox signaling between DNA Repair Proteins for efficient lesion detection*. Proc. Natl. Acad. Sci. USA, 2009. **106**(36): 15237–15242.
26. Boal, A.K., *DNA-mediated Charge Transport in DNA Repair*, Thesis in Chemistry. 2008, California Institute of Technology: Pasadena, CA.
27. Boal, A.K., et al., *DNA-Bound Redox Activity of DNA Repair Glycosylases Containing [4Fe- 4S] Clusters*. Biochemistry, 2005. **44**: 8397–8407.
28. Genereux, J.C., A.K. Boal, and J.K. Barton, *DNA-Mediated Charge Transport in Redox Sensing and Signaling*. Journal of the American Chemical Society, 2010. **132**(3): 891–905.

29. Nunez, M.E., D.B. Hall, and J.K. Barton, *Long-range oxidative damange to DNA: effects of distance and sequence*. Chem. Biol., 1999. **6**: 85–97.
30. Genereux, J.C., and J.K. Barton, *Mechanisms for DNA Charge Transport*. Chemical Reviews, 2010. **110**(3):1642–1662.
31. Merino, E.J., A.K. Boal, and J.K. Barton, *Biological contexts for DNA charge transport chemistry*. Current Opinion in Chemical Biology, 2008. **12**(2): 229–237.
32. Giese, B., et al., *Electron Relay Race in Peptides*. Journal of Organic Chemistry, 2009. **74**(10): 3621–3625.
33. Shih, C., et al., *Tryptophan-accelerated electron flow through proteins*. Science, 2008. **320**(5884):1760–1762.
34. Stubbe, J., et al., *Radical Initiation in the Class I Ribonucleotide Reductase: Long- Range Proton- Coupled Electron Transfer?* Chemical Reviews, 2003. **103**(6): 2167–2202.
35. Chang, M.C., et al., *Site-specific replacement of a conserved tyrosine in ribonucleotide reductase with an aniline amino acid: a mechanistic probe for a redox-active tyrosine*. Journal of the American Chemical Society, 2004. **126**(51):16702–16703.
36. Giese, B., M. Graber, and M. Cordes, *Electron transfer in peptides and proteins*. Curr. Opin. Chem. Biol., 2008. **12**(6): 755-759.
37. Cordes, M., et al., *Influence of amino acid side chains on long-distance electron transfer in peptides: Electron hopping via "Stepping Stones"*. Angewandte Chemie—International Edition, 2008. **47**(18): 3461–3463.
38. Gorodetsky, A.A., A.K. Boal, and J.K. Barton, *Direct Electrochemistry of Endonuclease III in the Presence and Absence of DNA*. J. Am. Chem. Soc., 2006. **128**:12082–12083.
39. Boal, A.K., and J.K. Barton, *Electrochemical Detection of Lesions in DNA*. Bioconjugate Chem., 2005. **16**: 312–321.
40. Boon, E.M., et al., *Mutation detection by electrocatalysis at DNA-modified electrodes*. Nature Biotechnology, 2000. **18**: 1096–1100.
41. Slinker, J.D., et al., *DNA Charge Transport over 34 nm*. Nature Chemistry, 2010. **3**: 230–235.
42. Buzzeo, M.C., and J.K. Barton, *Redmond Red as a Redox Probe for the DNA-Mediated Detection of Abasic Sites*. Bioconjugate Chemistry, 2008. **19**(11): 2110–2112.
43. Boon, E.M., et al., *DNA-mediated charge transport for DNA repair*. Proc. Natl. Acad. Sci. USA, 2003. **100**: 12543–12547.
44. Kelley, S.O., et al., *Single-base mismatch detection based on charge transduction through DNA*. Nucleic Acids Research, 1999. **27**(24): 4830–4837.
45. Kelley, S.O., et al., *Long-Range Electron Transfer through DNA Films*. Angew. Chem. Int. Ed., 1999. **38**(7): 941–945.
46. Slinker, J.D., et al., *Multiplexed DNA-Modified Electrodes*. Journal of the American Chemical Society, 2010. **132**(8): 2769–2774.

47. Cunningham, R.P., et al., *Endonuclease III Is an Iron-Sulfur Protein*. Biochemistry, 1989. **28**: 4450–4455.
48. Kovach, M.E., et al., *Four new derivatives of the broad- host- range cloning vector pBBR1MCS carrying different antibiotic- resistance cassettes*. Gene, 1995. **166**: 175–176.
49. Sambrook, J., and D.W. Russell, *Molecular cloning: A Laboratory Manual*. 2001, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
50. Bard, A.J., and L.R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*. 2001, Hoboken, NJ: John Wiley & Sons, Inc.
51. Wang, M., et al., *Electron Transfer in Peptides with Cysteine and Methionine as Relay Amino Acids*. Angewandte Chemie—International Edition, 2009. **48**(23): 4232–4234.
52. Bhattacharjee, S., et al., *Electron transfer between a tyrosyl radical and a cysteine residue in hemoproteins: Spin trapping analysis*. Journal of the American Chemical Society, 2007. **129**(44): 13493–13501.
53. Camba, R., and F.A. Armstrong, *Investigations of the oxidative disassembly of Fe-S clusters in Clostridium pasteurianum 8Fe ferredoxin using pulsed-protein-film voltammetry*. Biochemistry, 2000. **39**(34): 10587–10598.
54. Beinert, H., *Iron-sulfur proteins: ancient structures, still full of surprises*. Journal of Biological Inorganic Chemistry, 2000. **5**(3): 409–409.
55. Beinert, H., *Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures*. Science, 1997. **277**: 653–658.
56. Giese, B., *Long-distance electron transfer through DNA*. Annual Review of Biochemistry, 2002. **71**: 51–70.
57. Kelley, S.O., and J.K. Barton, *Electron Transfer Between Bases in Double Helical DNA*. Science, 1999. **283**: 375–381.
58. Kelley, S.O., et al., *Photoinduced Electron Transfer in Ethidium-Modified DNA Duplexes: Dependence on Distance and Base Stacking*. J. Am. Chem. Soci., 1997. **119**(41): 9861–9870.
59. Holmlin, R.E., P.J. Dandliker, and J.K. Barton, *Charge transfer through the DNA base stack*. Angewandte Chemie—International Edition, 1997. **36**(24): 2715–2730.
60. Giese, B., *Electron transfer in DNA*. Current Opinion in Chemical Biology, 2002. **6**(5): 612–618.
61. Gray, H.B., and J.R. Winkler, *Long-range electron transfer*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(10): 3534–3539.
62. Gray, H.B., and J.R. Winkler, *Electron tunneling through proteins*. Q. Rev. Biophys., 2003. **36**(3): 341–72.
63. Onuchic, J.N., et al., *Pathway Analysis of Protein Electron-Transfer Reactions*. Annual Review of Biophysics and Biomolecular Structure, 1992. **21**: 349–377.

64. Doherty, A.J., L.C. Serpell, and C.P. Ponting, *The helix-hairpin-helix DNA-binding motif: A structural basis for non-sequence-specific recognition of DNA*. Nucleic Acids Research, 1996. **24**(13): p. 2488-2497.
65. Berghuis, A.M., et al., *Mutation of Tyrosine-67 to Phenylalanine in Cytochrome-C Significantly Alters the Local Heme Environment*. Journal of Molecular Biology, 1994. **235**(4): p. 1326-1341.
66. Casimiro, D.R., et al., *Electron-Transfer in Ruthenium-Modified Cytochromes-C—Sigma-Tunneling Pathways through Aromatic Residues*. Journal of Physical Chemistry, 1993. **97**(50): p. 13073-13077.
67. Fearnhead, N.S., M.P. Britton, and W.F. Bodmer, *The ABC of APC*. Human Molecular Genetics, 2001. **10**(7): p. 721-733.
68. *The PyMOL Molecular Graphics System, Version 1.2r3pre*, Schrödinger, LLC.

Chapter 5

Towards an *In Vivo* Assay to Examine the Effects of EndoIII Mutations on MutY Activity

Dr. A. Bose provided mentorship and technical assistance. Experiments were performed in the laboratories of Professor D.K. Newman (Caltech and MIT).

Abstract

Previously, our laboratory developed an assay to examine the effects of EndoIII mutations on the *in vivo* activity of MutY. This assay was designed to test our hypothesis that these two proteins communicate *in vivo* using DNA-mediated charge-transfer to search for sites of DNA damage. This assay showed that Y82A, an electrochemically deficient mutant of EndoIII, was less effective at helping MutY repair lesions than WT EndoIII. Later work in our laboratory discovered other mutants with varying charge-transfer properties, such as Y82S, Y75A, and W178A, and these samples were also considered good candidates for the *in vivo* assay. However, for these mutants, the assay protocol had to be modified. Methyl viologen was added to the bacteria cultures to increase oxidative stress, and bacterial strains were created in which the EndoIII mutants were expressed from the chromosome of an *E. coli* strain rather than on a plasmid. The experiments with methyl viologen showed that mutagenesis rates increase in the presence in of this oxidative-stress-inducing reagent. However, even in the presence of methyl viologen, no statistically significant differences were observed between strains expressing different variants of EndoIII. These results differ from those of previous experiments, but the discrepancy may result from differences in genetic background, growth phase at harvest, and media recipes. Importantly, the results of these experiments further elucidate the complex nature of cellular environments and their responses to mutagenesis and stress.

Introduction

Previously, our laboratory introduced a model for cooperative scanning for DNA damage using DNA-mediated charge-transfer (CT). According to this model, DNA-bound, redox-active proteins could communicate with each other by sending and receiving electrons and/or holes through the DNA helix. However, if the intervening DNA contains a lesion, then CT will not occur efficiently [1–3], signaling to the bound proteins that DNA damage is present. These proteins can then localize to the damaged site and execute repair. Our lab developed an *in vivo* assay to test this cooperative search model by examining how inactivation of *nth*, the gene that encodes Endonuclease III (EndoIII) affects the activity of MutY. EndoIII and MutY are both [4Fe-4S] cluster-containing DNA glycosylases whose *E. coli* forms have been shown to be redox-active when bound to DNA [1]. These proteins display similar midpoint potentials [1], indicating that they could transmit electrons to each other and potentially help each other search for their substrate DNA lesions.

To develop an assay to test whether MutY and EndoIII cooperate *in vivo*, the reporter strain *E. coli* CC104 was used. Developed by the laboratory of Professor Jeffrey Miller, CC104 contains an A→C mutation in the codon that encodes glutamic acid 461 of the *lacZ* gene, which encodes β-galactosidase. This mutation converts residue 461 to an alanine, rendering β-galactosidase inactive in wild-type CC104 strains [4]. However, oxidation processes within the cell can oxidize the guanine paired with the substituted cytosine, producing 7,8-dihydro-8-

oxoguanine (8-oxo-G) [5–7]. DNA replication machinery will mistakenly pair an adenine with 8-oxo-G, and then the next round of replication will place a thymine opposite this adenine. The end result of these processes is a G:C → T:A transversion at position 461 [5]. These "transverted" mutants of CC104 will then possess a functional copy of β-galactosidase and be able to grow on minimal medium supplemented with lactose as the sole carbon source.

However, the enzyme MutY removes adenine mispaired with 8-oxo-G, thus preventing CC104 from reverting to a *lac*$^+$ phenotype. Consequently, when CC104 colonies appear in low numbers on lactose-minimal media plates, the results indicate active MutY. High numbers of *lac*$^+$ revertant colonies suggest that MutY's activity has been compromised. Indeed, the number of *lac*$^+$ CC104 revertants increases 15-fold when *mutY* is genetically inactivated [8]. Our laboratory has also shown that the number of *lac*$^+$ revertants increases two fold when the *nth* gene is inactivated [8]. Given that MutY and EndoIII have not been shown to interact *in vivo*, this result suggests that EndoIII helps MutY detect and repair 8-oxo-G. This cooperation is not due to overlapping substrates, because a *mutY*$^-$*nth*$^-$ strain of CC104 produced as many *lac*$^+$ revertants, within error, as a *mutY*$^-$ strain [8].

This *lac*$^+$ reversion assay was also used to test how certain EndoIII mutations affected the ability of MutY to detect its substrates. A plasmid containing D138A, an electrochemically proficient but enzymatically deficient mutant, offered the same level of assistance to MutY as WT EndoIII [8]. However, Y82A, an enzymatically active but electrochemically deficient mutant was less able

to help MutY detect lesions [8]. These data support the model that EndoIII and MutY cooperatively search for DNA damage using DNA-mediated CT.

After examining these results, it was decided that this assay should be used to examine the effect of other EndoIII mutations on the activity of MutY. However, some changes to the protocol were recommended. First, the addition of oxidative stress-inducing reagents, such as methyl viologen, was encouraged because they could increase the level of internal reactive oxygen species (ROS), cause more guanine oxidation, and possibly increase the overall revertant count. High numbers of revertants could potentially enhance subtle differences between strains of CC104 expressing different *nth* mutants.

Second, the assay needed to be adjusted such that plasmids were not required. Previously, EndoIII mutants were expressed on plasmids and then transformed into CC104*nth*$^-$ cells. However, plasmids require the use of antibiotics, which could interact with the aforementioned oxidative stress-inducing chemicals. Strains expressing plasmids may also contain slight differences in plasmid copy number from cell to cell, making data difficult to interpret and increasing the possibility of errors and outliers during the assay. Consequently, strains were constructed in which *nth* variants were expressed on the chromosomes of CC104. The *nth* variants were cloned into a plasmid that expresses genes off of a rhamnose-inducible promoter [9], and then this plasmid was allowed to integrate into the λ-phage attachment site of the strain CC104Δ*nth*. After these strains were prepared, the *lac*$^+$ reversion assay was performed to examine the effect of methyl viologen

concentration on the number of *lac*$^+$ revertants of the strain CC104Δ*nth* $P_{rha\text{-}nth}$. Then, the assay was performed using strains of CC104 expressing chromosomal copies of the *nth* variants WT, Y82A, Y82S, Y75A, and W178A.

Materials and Methods

Bacterial Strains and Plasmids. The vectors pKOV, pAH120, pINT-ts, and strain *E. coli* CC104 were obtained from the Coli Genetic Stock Center of Yale University.

Cloning. All PCR reactions were carried out using either the Failsafe PCR System with Buffers D or G (Epicentre Biotechnologies), or KOD polymerase (Merck4 Biosciences). Plasmids were isolated using the QIAprep Spin Miniprep Kit (Qiagen) and DNA products were isolated from agarose gels using the Wizard Kit (Promega). All primers were purchased from Integrated DNA Technologies.

Creation of CC104Δ*nth*. The methods used here were adapted from those described previously [10]. Regions 1000 bp upstream and downstream of the *nth* gene were PCR amplified from chromosomal DNA using primers CARJKB172, delnthfusrev, delnthfusfor, and CARJKB173. The resulting PCR products were used as templates in a second PCR experiment, resulting in a product of ~ 2000 bp that omits the *nth* gene. This product was ligated into the BamHI and NotI sites of the vector pKOV. Electrocompetent CC104 cells were then transformed with the pKOV plasmid containing the Δ*nth* construct and recovered at 30°C. The transformation reactions were plated onto Luria-Bertani (LB) + chloramphenicol

medium and incubated at either 30°C or 42°C. After incubation, the ratio of colony number on the 42°C plate to colony number on the 30°C indicated high integration efficiency. Cells that had been incubated at 42°C were grown in liquid culture, serially diluted, and plated onto LB + sucrose medium. Sucrose-tolerant cells were selected and examined by PCR using primers CARJKB164, CARJKB165 to verify that the *nth* gene had been cleanly deleted. After this PCR reaction, the samples that appeared to be Δ*nth* were isolated, re-streaked, and analyzed by colony PCR again. This process was repeated at least three times to ensure that the samples were not contaminated with WT CC104. Finally, sequencing with primers CARJKB174, CARJKB175, CARJKB180 was used to verify that the isolated strains were truly Δ*nth*.

Creation of CC104Δ*nth* $P_{rha\text{-}nth}$ Strains. The methods described here are adapted from those previously published [9]. Variants of *nth* were PCR amplified from overexpression vectors using primers CARJKB164 and CARJKB165 and then ligated into the BamHI and NdeI sites of pAH120. Samples of pAH120 expressing different variants of *nth* were transformed into electrocompetent CC104Δ*nth* that were also expressing the plasmid pINT-ts. Cells were then incubated at 30ºC for one hour to allow chromosomal integration of pAH120, and then at 42ºC for 30 minutes, during which time pINT-ts should be cured. Cultures were spread onto Luria-Bertani (LB) + kanamycin media and incubated at 37ºC. This final incubation selects for cells that have integrated pAH120, since this plasmid confers kanamycin resistance. Colonies that emerged on these kanamycin selection plates

were screened using colony PCR and primers CARJKB164 and CARJKB165 to verify the presence of an *nth* variant on the chromosome. Further PCR screens were performed with primers given in reference [9] to verify that the pAH120 plasmid integrated only once.

Lac$^+$ Reversion Assay. Freezer stocks of the appropriate CC104 strains were streaked onto Luria-Bertani (LB) medium and incubated for 12 hours at 37°C. Ten different samples from each strain were inoculated into ten tubes of LB liquid medium supplemented with 1 mM rhamnose and grown for 13 hours at 37°C with shaking at 250 rpm. From each overnight culture, 200 μL of was transferred into 10 mL Minimal A Medium [11] supplemented with 0.4% glycerol and 1 mM rhamnose. Methyl viologen was added to these cultures to a final concentration of either 0 μM, 10 μM, 25 μM, or 50 μM. These cultures were grown at 37°C until OD_{600} = 0.4. The cultures were pelleted by centrifugation, rinsed in 5 mL Minimal A Medium with no carbon source, and plated onto Minimal A Medium supplemented with 0.2% lactose. The plates were incubated for 42 hours at 37°C, at which point *lac*$^+$ revertants were counted. Assays were performed at least two times for each sample, and ten replicates of each sample were used in each assay. The two highest and two lowest values were dismissed and the remaining six values were averaged. Errors are reported as standard deviations.

Results and Discussion

The effect of methyl viologen on the number of lac$^+$ revertants. One of the first assays performed with the newly created CC104Δ*nth* P*rha-nth* strain examined

whether the redox cycling reagent methyl viologen would affect the number of *lac*$^+$ revertants. The total number of revertants nearly doubled on increasing the methyl viologen concentration from 0 μM to 50 μM (Figure 5.1). This result was expected given that methyl viologen reacts with oxygen to generate superoxide [12], a reactive oxygen species (ROS). ROS can damage DNA, particularly at guanine residues since guanine is the most easily oxidized nucleobase [13–15].

Lac+ revertant counts as a consequence of nth *gene mutations*. After it was established that methyl viologen could increase the number of *lac*$^+$ revertants observed during the assay, methyl viologen was then incorporated into the protocol used to examine strains expressing different *nth* variants. The strains include CC104Δ*nth* and strains of CC104Δ*nth* P*rha-nth* expressing the *nth* variants WT, Y82A, Y82S, Y75A, and W178A. Even with the addition of methyl viologen, the results show no statistically significant differences between the number of *lac*$^+$ revertants produced by these different strains (Figure 5.2). It was surprising that these results were not consistent with results from plasmid-based assays performed previously [8]. However, the discrepancy may result from minor genetic differences in the parent CC104 strains used to produce these mutants. Other changes such as growth medium, growth phase at harvest, and nutrient source, could also affect the *in vivo* mutation rate, as discussed below. Also, expression levels of the *nth* gene under the different experimental conditions were not determined.

Variations in experiments. First, the number of *lac*$^+$ revertants produced by a given sample of CC104 seems to rely upon the genotype of that strain. Even subtle, non-detectable, genetic differences can lead to differences in internal mutagenesis rates. For example, our laboratory has worked with two different "wild-type" strains for CC104 *E. coli* during the past several years. One strain came from a collaborator's laboratory (Professor Jeffrey Miller, University of California Los Angeles) and one came from a strain library (Coli Genetic Stock Center, Yale University). Although both strains have the same genotype, they produced consistently different numbers of *lac*$^+$ revertants when measured side-by-side under identical experimental conditions (data not shown). This result speaks to the fact that even small, unnoticeable, changes in the genetic background of given strains can affect their mutagenesis rates. It was also observed that the sequence of the *lacZ* gene found in the strains in our laboratory was different than that published when Cupples and Miller reported the creation of strain CC104 [4]. However, it remains to be determined whether the *lacZ* sequence discrepancy accounts for the difference between previously published assay results [8] and the data presented here.

Second, cell cultures in the CC104Δ*nth* $P_{rha\text{-}nth}$ assays were harvested in mid-exponential phase to ensure that they were actively growing while plated. By contrast, cultures in the plasmid-based assay were harvested in late exponential phase [8]. Growth phase may affect mutagenesis rates because levels of expression of oxidative stress response genes can change in stationary phase [16], as can cells' resistance to peroxide [17]. However, metabolism and expression of certain genes

slow in stationary phase [17], possibly making it more difficult for colonies to grow on lactose-minimal medium.

Third, the assay using CC104Δ*nth* $P_{rha\text{-}nth}$ strains required different growth media than the plasmid-based assay. Rhamnose had to be added to the growth medium to induce expression of the *nth* variants [18, 19]. Expression from the rhamnose-inducible promoter is inhibited by glucose [19], which was used in the plasmid-based assay. Consequently, glycerol was used as the carbon source in the experiments presented here. This change in carbon source could be significant because carbon source has been shown to affect mutagenesis rates through mechanisms thought to involve carbon catabolite repression (reviewed in references [20, 21]). Mutagenesis reporter strains produce larger numbers of mutant colonies when grown on glycerol, a non-repressive growth substrate, than when grown on glucose, a highly repressive growth substrate [22, 23]. Glucose-related compounds with even more repressive effects, such as glucose-6-phosphate and methyl-α-D-glucopyranoside, exert even more powerful anti-mutagenic effects [22]. It is unclear exactly how carbon catabolite repression influences mutagenesis rates, but authors hypothesize that the SOS DNA repair pathway (reviewed in reference [24]) may be affected. Carbon catabolite repression depresses levels of cyclic AMP (cAMP) [20, 21], an important cellular signaling molecule. In turn, low cAMP levels reduce the expression of genes involved in SOS DNA repair, which is highly mutagenic [24]. Perhaps the anti-mutagenic effects of carbon catabolite repression are part of the reason why glucose transport is induced in response to oxidative

stress in *E. coli* through the action of the *SoxRS* genes [25]. The exact relationship between metabolism and mutation remains unclear, but the change in growth medium could have impacted the outcome of the lac^+ reversion assay.

Although the goal of the experiments presented here was to determine how mutations in EndoIII affect the activity of MutY, the data do not provide such an answer. However, the process of optimizing these experiments and interpreting the results offered yet another perspective on the profound complexity of *E. coli* cells. The phenotype observed previously by Boal, et al. [8] relied upon a careful combination of cell genotype, growth phase, and growth conditions. Future work examining with *in vivo* affects of *nth* mutations will need to consider these parameters.

Table 5.1: Primers used in *lac*$^+$ reversion experiments

Name	Sequence	Purpose
CARJKB172	5'-actagtgcggccgcgaactgcgcctttctgactggatc-3'	Primer for overlap extension to create Δ*nth* construct, and for verifying absence of *nth* on CC104 chromosome. It binds to a sequence 1000 bp upstream of *nth* gene.
CARJKB173	5'-ggcgcgccggatcccataaccaatcgccagcacaatagc-3'	Primer for overlap extension to create Δ*nth* construct, and for verifying absence of *nth* on CC104 chromosome. It binds to a sequence 1000 bp downstream of *nth* gene.
delnthfusfor	5'- gcattgccaaacggtgaaacagggaatgtctgatgaagaaaaggggtaacaccgattaccccattg-3'	Primer for overlap extension to create Δ*nth* construct. It binds to a sequence immediately following the *nth* gene.
delnthfusrev	5'-caatggggtaatcggtgttacccctttttcttcatcagacattccctgtttcaccgtttggcaatgc-3'	Primer for overlap extension to create Δ*nth* construct. It binds to a sequence immediately preceding the *nth* gene.
CARJKB174	5'-acattgtttgacggtgcaga-3'	Sequencing Δ*nth* strains
CARJKB175	5'-ggtcaatatttgttgcgtgtg-3'	Sequencing Δ*nth* strains
CARJKB180	5'-gcaatggcacattgtttgac-3'	Sequencing Δ*nth* strains
CARJKB164	5'-ggcgcgcccatatgaataaagcaaaacgcctggagatc-3'	Primer for cloning *nth* alleles into pAH120, NdeI site
CARJKB165	5'-ggcgcgccggatcctcagatgtcaactttctctttgtattc-3'	Primer for cloning *nth* alleles into pAH120, BamHI site
CARJKB176	5'-cgaattcaggcgctttttag-3'	Sequencing *nth* alleles after chromosomal insertion
CARJKB177	5'- tctgctggaaccactttcagt-3'	Sequencing *nth* alleles after chromosomal insertion
CARJKB178	5'-tgaaagtggttccagcagag-3'	Sequencing *nth* alleles after chromosomal insertion
CARJKB179	5'- gggagtgggacaaaattgaa-3'	Sequencing *nth* alleles after chromosomal insertion

Figure 5.1: Lac$^+$ revertants as a function of methyl viologen concentration. Methyl viologen was added to cultures of CC104Δ*nth* P*rha-nth* growing in liquid minimal medium at the concentrations shown. Cultures were grown to OD_{600}= 0.4, pelleted, rinsed once, and then plated on minimal A medium supplemented with 0.2% lactose. *Lac*$^+$ revertants were counted after 42 hours incubation at 37°C.

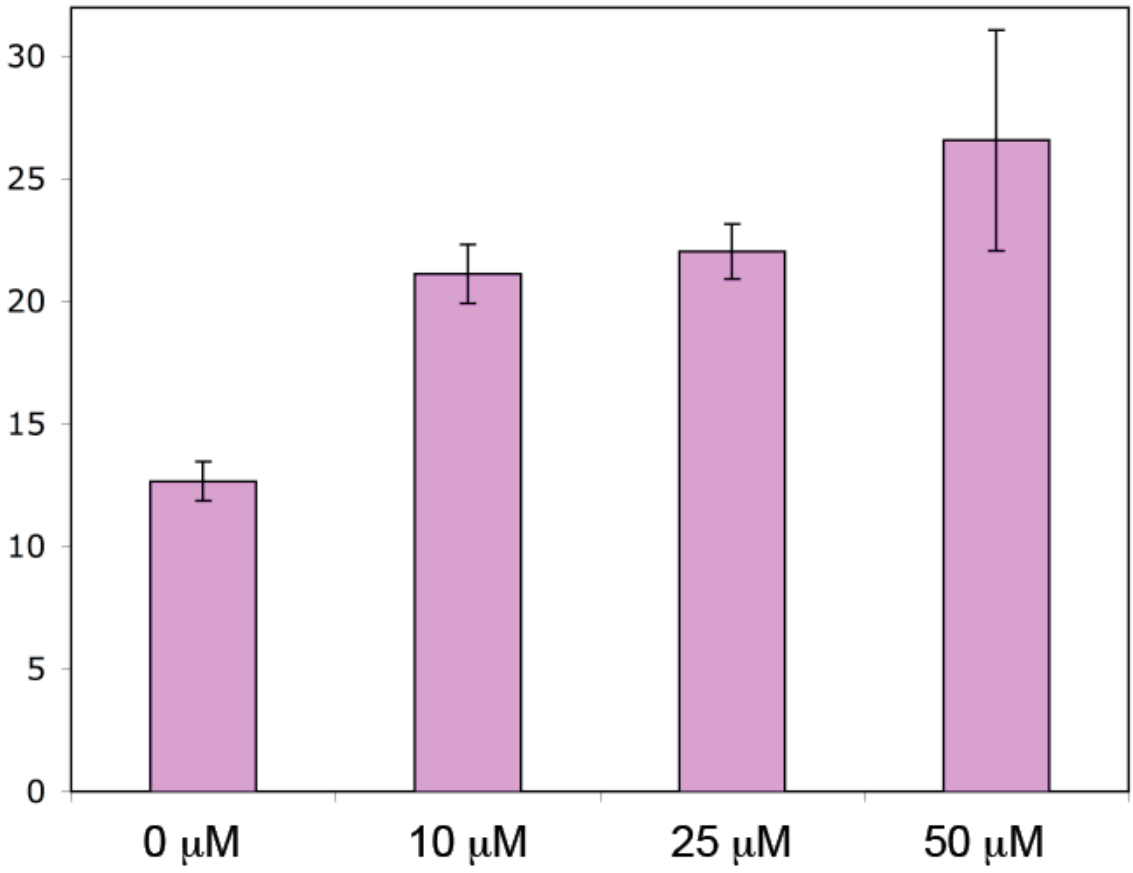


Table 5.2: Average *lac*+ revertants counted for different concentrations of methyl viologen. Results are averaged over four trials. Values are reported as the number of revertants per ~2x10^8 cells/mL.

Concentration of methyl viologen	0 μM	10 μM	25 μM	50 μM
No. of lac+ revertants	12.7 ± 0.8	21.1 ± 1.2	22.0 ± 1.1	26.6 ± 4.5

Figure 5.2: Lac$^+$ revertants produced by CC104Δ*nth* P*rha-nth* strains expressing different *nth* alleles. Strains were grown in Minimal A Medium supplemented with 0.4% glycerol, 1 mM rhamnose, and 50 μM methyl viologen. Cultures were grown to OD_{600}= 0.4, pelleted, rinsed once, and then plated on Minimal A Medium supplemented with 0.2% lactose. *Lac*$^+$ revertants were counted after 42 hours incubation at 37°C.

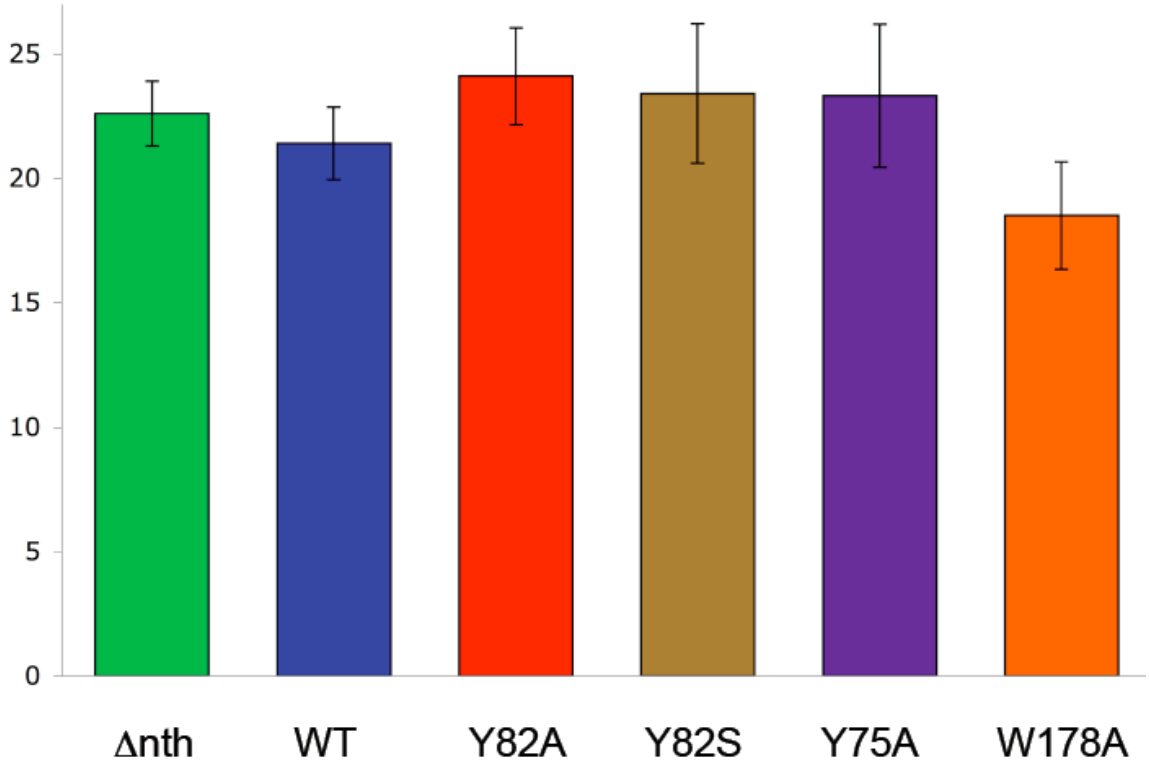


Table 5.3: Average *lac*$^+$ revertants counted for strains expressing different variants of *nth*. Values are reported as the number of revertants per ~ 2x10^8 cells/mL.

Strain	CC104Δ*nth*	CC104Δ*nth* P*rha-nth*	CC104Δ*nth* P*rha-Y82A*	CC104Δ*nth* P*rha-Y82S*	CC104Δ*nth* P*rha-Y75A*	CC104Δ*nth* P*rha-W178A*
No. of lac$^+$ revertants	22.6 ± 1.3	21.4 ± 1.5	24.1 ± 1.9	23.4 ± 2.8	23.3 ± 2.9	18.5 ± 2.2

References:

1. Boal, A.K., et al., *DNA-Bound Redox Activity of DNA Repair Glycosylases Containing [4Fe- 4S] Clusters.* Biochemistry, 2005. **44**: 8397–8407.
2. Boal, A.K., and J.K. Barton, *Electrochemical Detection of Lesions in DNA.* Bioconjugate Chem., 2005. **16**: 312–321.
3. Boon, E.M., et al., *Mutation detection by electrocatalysis at DNA-modified electrodes.* Nature Biotechnology, 2000. **18**: 1096–1100.
4. Cupples, C.G., and J.H. Miller, *A set of lacZ Mutations in Escherichia coli that Allow Rapid Detection of Each of the Six Base Substitutions.* Proc. Natl. Acad. Sci. USA, 1989. **86**: 5435–5439.
5. Neeley, W.L., and J.M. Essigmann, *Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products.* Chemical Research in Toxicology, 2006. **19**(4): 491–505.
6. Burrows, C.J., and J.G. Muller, *Oxidative nucleobase modifications leading to strand scission.* Chemical Reviews, 1998. **98**(3): 1109–1151.
7. David, S.S., V.L. O'Shea, and S. Kundu, *Base-excision repair of oxidative DNA damage.* Nature, 2007. **447**: 941–950.
8. Boal, A.K., et al., *Redox signaling between DNA Repair Proteins for efficient lesion detection.* Proc. Natl. Acad. Sci. USA, 2009. **106**(36): 15237–15242.
9. Haldimann, A., and B.L. Wanner, *Conditional-Replication, Integration, Excision, and Retrieval Plasmid-Host Systems for Gene Structure-Function Studies of Bacteria.* Journal of Bacteriology, 2001. **183**(21): 6384–6393.
10. Link, A.J., D. Phillips, and G.M. Church, *Methods for Generating Precise Deletions and Insertions in the Genome of Wild-Type Escherichia coli: Application to Open Reading Frame Characterization.* Journal of Bacteriology, 1997. **179**(20): 6228–6237.
11. Miller, J.H., *Experiments in Molecular Genetics*. 1972, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.
12. Bus, J.S., and J.E. Gibson, *Paraquat—Model for Oxidant-Initiated Toxicity.* Environmental Health Perspectives, 1984. **55**(APR):37–46.
13. Steenken, S., and S.V. Jovanovic, *How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution.* Journal of the American Chemical Society, 1997. **119**(3): 617–618.
14. Saito, I., et al., *Photoinduced DNA Cleavage Via Electron-Transfer—Demonstration That Guanine Residues Located 5′ to Guanine Are the Most Electron-Donating Sites.* Journal of the American Chemical Society, 1995. **117**(23): 6406–6407.
15. Genereux, J.C., A.K. Boal, and J.K. Barton, *DNA-Mediated Charge Transport in Redox Sensing and Signaling.* Journal of the American Chemical Society, 2010. **132**(3): 891–905.

16. Demple, B., *Regulation of Bacterial Oxidative Stress Genes.* Annual Review of Genetics, 1991. **25**: 315–337.
17. White, D., *The physiology and biochemistry of prokaryotes*. 2000, New York: Oxford University Press.
18. Cardona, S.T., and M.A. Valvano, *An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in Burkholderia cenocepacia.* Plasmid, 2005. **54**(3): 219–228.
19. Haldimann, A., L.L. Daniels, and B.L. Wanner, *Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the Escherichia coli phosphate regulon.* Journal of Bacteriology, 1998. **180**(5):1277–1286.
20. Gorke, B., and J. Stulke, *Carbon catabolite repression in bacteria: many ways to make the most out of nutrients.* Nat. Rev. Microbiol., 2008. **6**(8): 613–624.
21. Deutscher, J., *The mechanisms of carbon catabolite repression in bacteria.* Curr. Opin. Microbiol., 2008. **11**(2): 87–93.
22. Ambrose, M., and D.G. Macphee, *Catabolite repressors are potent antimutagens in Escherichia coli plate incorporation assays: experiments with glucose, glucose-6-phosphate and methyl-alpha-D-glucopyranoside.* Mutation Research—Fundamental and Molecular Mechanisms of Mutagenesis, 1998. **398**(1–2): 175–182.
23. Ambrose, M., and D.G. MacPhee, *Glucose and related catabolite repressors are powerful inhibitors of pKM101-enhanced UV mutagenesis in Escherichia coli.* Mutat. Res., 1998. **422**(1): p. 107-112.
24. Rattray, A.J., and J.N. Strathern, *Error-prone DNA polymerases: when making a mistake is the only way to get ahead.* Annu Rev Genet, 2003. **37**: p. 31-66.
25. Rungrassamee, W., X. Liu, and P.J. Pomposiello, *Activation of glucose transport under oxidative stress in Escherichia coli.* Archives of Microbiology, 2008. **190**(1): p. 41-49.

Chapter 6

Summary and Outlook

Summary: Pathways for Protein-DNA Charge Transfer

The experiments described in the preceding chapters examined the charge-transfer (CT) properties of several site-directed mutants of three base-excision repair enzymes: UDG, MutY, and EndoIII. These proteins are thought to participate in DNA-mediated long-range charge transfer in order to scan the genome for DNA damage. The mutants that were created and characterized were chosen in order to answer four questions about how amino-acid changes in the proteins affect protein-to-DNA CT: 1) How do perturbations in the [4Fe-4S] cluster binding domain affect the CT properties of DNA repair enzymes? 2) Which types of amino acid residues best mediate protein-to-DNA CT? 3) What possible amino acid pathways may exist for DNA-protein CT? 4) What are the biological consequences of having an impaired CT pathway?

The first of these questions, how the [4Fe-4S] cluster binding domain affects CT, was addressed by studying the UDG mutants C17H, C85S, C101S, and the EndoIII mutants W178A and Y185A. Altering the coordination environment could greatly decrease CT, as was the case with C85S UDG. However, C101S UDG produced an electrochemical signal that was within error of that of WT UDG. The UDG mutant C17H and the EndoIII mutants W178A and Y185A produced signals larger than those of the corresponding WT proteins. This signal increase was attributed to structural perturbations in the mutants' secondary structures that may have permitted water molecules to enter the protein and accelerate CT [1, 2].

The C17H mutation may have also placed a more CT-active residue closer to the cluster than that present in the native protein.

Mutagenesis studies of MutY Y82 position mutants and EndoIII Y82 position mutants investigated which types of residues best facilitate CT in their respective proteins. In MutY, long residues such as tyrosine and leucine that can intercalate into the DNA helix were the best CT mediators. In EndoIII, aromatic residues proved most CT-proficient. Several aromatic amino acids in EndoIII, including Y82, Y75, Y55, and F30, form what may be a partial pathway between the DNA and the [4Fe-4S] cluster.

The fourth question, the biological consequences of CT impediments has been addressed through colorectal cancer research. Mutations at residues Y114, Y165, and Y166 in human MUTYH have appeared in colorectal cancer patients [3, 4]. These residues align with F30, L81, and Y82, respectively, in *E. coli* EndoIII, and Y165 aligns with Y82 in *E. coli* MutY. EndoIII F30 and Y82, and MutY Y82 facilitate CT in their proteins. Therefore, the equivalent MUTYH residues may possibly mediate protein to DNA CT in human cells. Without this CT capability, MUTYH may be unable to detect and repair DNA damage, causing mutations to accumulate that would lead to colorectal cancer. CT deficiencies in biological systems could produce very deleterious consequences for the cell.

As a whole, the BER mutants examined here helped elucidate the properties that create a CT-active enzyme capable of participating in DNA-mediated long-range signaling. The protein must contain a well-protected, well-coordinated, redox

active cofactor (such as a [4Fe-4S] cluster). It must contain amino acids capable of intercalating into the DNA helix to allow charge to flow from the DNA into the protein, and it must contain a pathway of CT-active amino acids that can convey this charge from the DNA to the redox cofactor. Future experiments, such a time-resolved spectroscopy, will be necessary to determine how mutations in the CT pathway affect CT rates. Other ongoing *in vivo* experiments will determine how mutations in the CT pathways of these DNA repair proteins affect *in vivo* mutagenesis rates.

Outlook: Beyond DNA Repair

The majority of the experiments performed for this thesis investigated DNA-mediated charge-transfer (CT) for DNA-repair using *in vitro* characterization of proteins that may participate in this process. However, recent research is developing *in vivo* experiments to examine long-range electrochemical signaling as means of searching for DNA damage [5]. Such experiments include the "helper function" assay that investigated whether EndoIII (and variants) could help MutY detect and repair its substrates [5]. Another way to study the *in vivo* activity of EndoIII would be to investigate the role of the *rnf* operon in DNA repair. The *rnf* operon is a seven-gene operon of which the *nth* gene, which encodes EndoIII, is the final component [6] (Figure 6.1). This operon has not been fully characterized in *E. coli,* but could possibly regulate redox processes within the cell.

The notion that *rnf* genes control cellular redox processes is supported by several lines of evidence. First, the "*rnf*" operon is so named because its operon structure and gene sequences bear similarity to "rhodobacter nitrogen fixation" genes, or genes used by *Rhodobacter capsulatus* and *Azotobacter vinelandii* to reduce dinitrogen [7–9]. In *R. capsulatus, rnf* operon genes are thought to donate electrons to the nitrogenase complex [8, 9]. In *A. vinelandii*, *rnf* genes are thought to assist in iron-sulfur cluster assembly [7]. Similar gene products are also hypothesized to supply electrons for 2,4-dinitrophenol reduction in *R. capsulatus* [10], encode a ferredoxin:NAD^+ oxidoreductase for caffeate reduction in *Acetobacterium woodii* [11–13], and form a Na^+ translocating NADH:quinone oxidoreductase complex in *Vibrio alginolyticus* [14]. The fact that homologous genes participate in redox processes in so many other organisms makes it likely that they mediate redox chemistry in *E. coli.*

Other evidence that the *rnf* gene products facilitate redox processes from sequence predictions and preliminary biochemical characterization of the gene products. RnfB is hypothesized to contain several iron-sulfur clusters, RnfC has several domains that are similar to those found in other redox-active proteins, and RnfG is predicted to contain a FMN-binding domain[1]. RnfB and RnfC from *R. capsulatus* have been purified and were red-brown and brown, respectively [8, 14], colors often indicative of an iron-sulfur cluster. RnfB exhibited an absorbance and

[1] All predictions based on gene sequence information were either taken from the published literature or made using one of the following online resources: http://www.ncbi.nlm.nih.gov, http://www.uniprot.org, http://ca.expasy.org/.

EPR spectrum characteristic of [2Fe-2S] proteins and RnfC exhibited an absorbance spectrum characteristic of [4Fe-4S] proteins [14]. RnfC from *R. capsulatus* is also predicted to contain binding sites for NAD(H) and flavin mononucleotide (FMN), components of redox-active proteins [15]. However, neither RnfB nor RnfC could be characterized in great detail because the proteins were oxygen labile [14].

Few experiments have been performed on the *rnf* genes of *E. coli*, although one recent breakthrough demonstrated that these genes affect the activity of *E. coli* SoxR [16]. SoxR is a [2Fe-2S] transcription factor that regulates *E. coli* cells' response to oxidative stress [17]. When the iron-sulfur cluster of SoxR becomes oxidized, this protein then upregulates several genes that help the cell remove reactive oxygen species [17]. After the threat of oxidative stress is gone, *rnfC* helps re-reduce SoxR and stop the upregulation of oxidative stress response genes [16]. The mechanism of re-reduction remains unknown. Because SoxR is a DNA-binding protein that can participate in long-range DNA-mediated CT [18], it is possible that the *rnf* operon products that interact with it can also oxidize or reduce other redox-active DNA binding proteins such as MutY and EndoIII.

To better understand the cellular role of rnf proteins, experiments should be performed to address several questions: 1) Where do rnf proteins localize within the cell? 2) When are the *rnf* genes upregulated or downregulated? 3) Does the deletion of any *rnf* genes produce any detectable phenotypes? Regarding their localization, RnfA and RnfE have been shown to be membrane-bound [17, 19], and RnfD is also

predicted to localize to the membrane. RnfB and RnfC are predicted to contain both membrane and soluble domains (see footnote 1). It is possible that the rnf proteins form a membrane-bound complex through which reducing (and/or oxidizing) equivalents are transferred to other cellular components, possibly in an NAD^+/NADH mediated fashion (Figure 6.2). Fractionation and Western blotting experiments could confirm where rnf proteins localize within the *E. coli* cell.

Little is known about what factors alter expression of *rnf* genes. In *R.capsulatus*, their expression is decreased under iron limitation [8]. Addition of paraquat and hydrogen peroxide does not upregulate the expression of *nth* in *E. coli* [20], and so would unlikely upregulate the co-expressed *rnf* genes. Other conditions could be tested using qRT-PCR experiments. Because *rnf* gene products are so important for the activity of the SoxRS regulon, it would also be interesting to test whether they influence the activity of any other redox-active transcription factors in *E. coli*, such as OxyR [15], FNR [15], and MgrA [21]. Experimentally, a link between transcription factors and *rnf* products could be determined by a combination of genetics and molecular biology experiments. First, *rnf* genes could be genetically inactivated, particularly *rnfB* and *rnfC*, since these genes are predicated to contain the most redox-active domains. The activity of OxyR, FNR, and MgrA in these deletion strains could then be measured through qRT-PCR analysis of genes in their regulons or by β-galactosidase assays of *lacZ* fusions to any of these regulon genes. To further investigate whether *rnf* genes are involved in redox regulation, the ratio of a NAD^+ to NADH could also be quantified in *rnf*

deletion strains relative to wild-type *E. coli* strains. Pull-down assays could also be performed to determine which other *E. coli* proteins interact with *rnf*B and *rnfC.*

In total, the experiments performed on mutant BER proteins determined which amino acids mediate protein-to-DNA CT in the [4Fe-4S] glycosylase proteins UDG, MutY, and EndoIII. The experiments mentioned in this final chapter also aim to determine whether these proteins or any of their genomic neighbors play a broader *in vivo* role in regulating DNA repair, and/or other redox processes. EndoIII is part of an operon whose genes encode proteins that may be involved in uptake of oxidizing or reducing equivalents. The operon that encodes MutY also encodes YggX, which has been shown to help protect cells from oxidative stress [22–24]. Consequently, the BER enzymes and their operons, may play an important *in vivo* role beyond DNA repair. These enzymes and their genetic neighbors may be part of much broader system through which the cell maintains redox control and regulates processes that sustain biological activity.

Figure 6.1: *Rnf* operon. The *nth* gene, which encodes EndoIII, is the terminal gene of a seven-gene operon. The upstream genes of this operon play an undetermined role in *E. coli*. The *ydgK* gene is upstream of the *rnf* operon, but contains its own promoter. Figure adapted from reference [25].

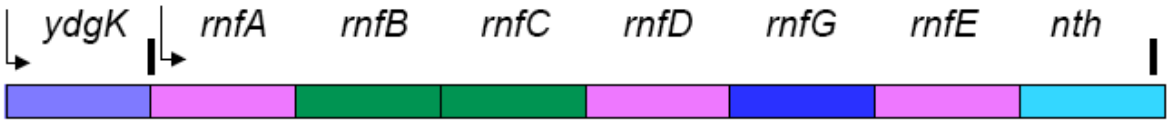
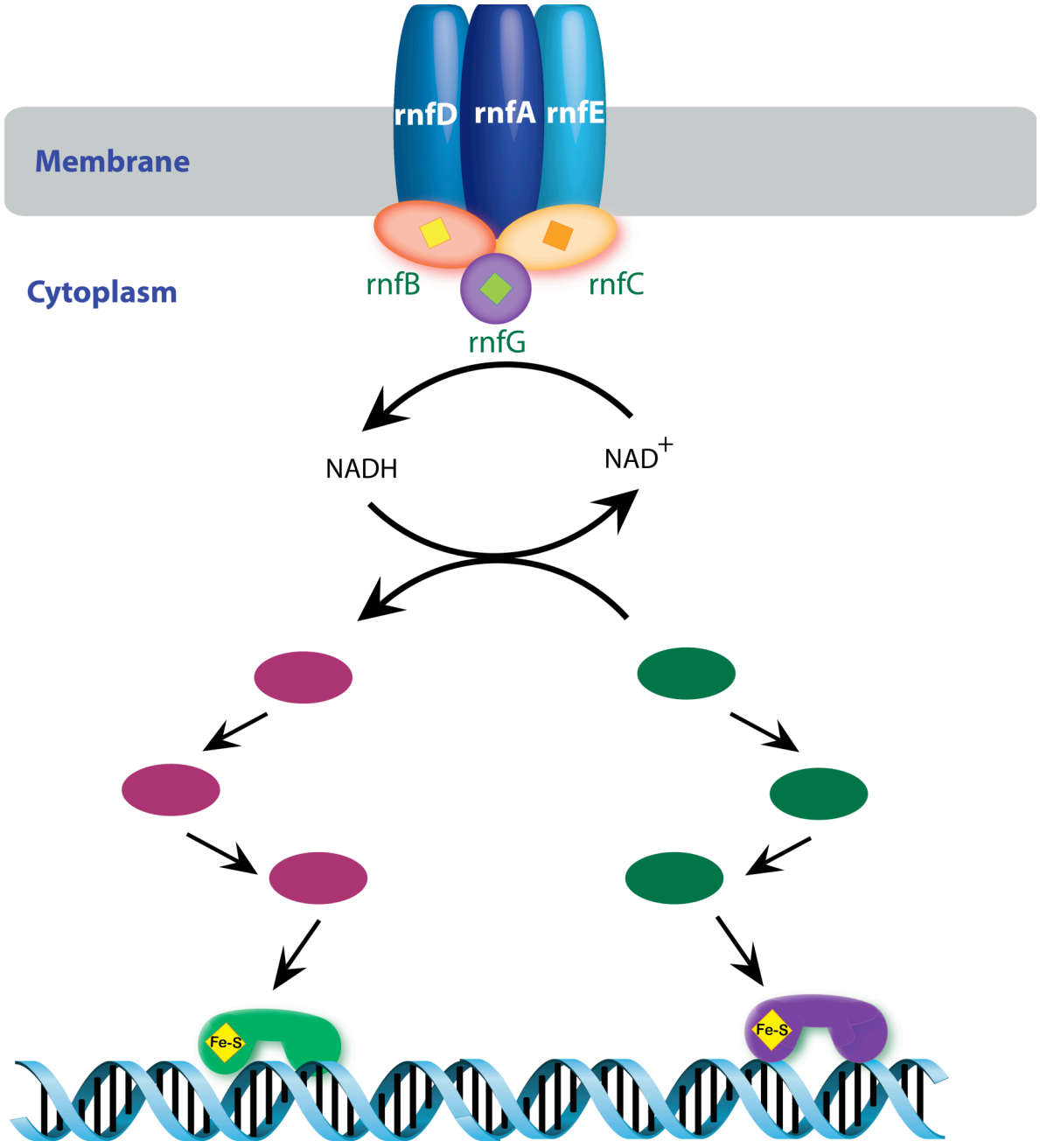


Figure 6.2: Possible cellular localization and function of rnf proteins. The products of the *rnfD, rnfA,* and *rnfE* genes are predicted to be membrane-bound. The products of the *rnfB* and *rnfC* genes are predicted to contain both membrane and soluble domains, while the *rnfG* gene product is predicted to be soluble. The latter three genes are predicted to encode binding sites for redox-active cofactors. Collectively, the rnf proteins may form a membrane-bound complex responsible for transporting reducing and/or oxidizing equivalents into the cell and passing them to other molecules *in vivo*. These reducing and/or oxidizing equivalents may be transferred through an NADH/NAD^+ mediated mechanism, since *rnf* genes encode binding sites for these molecules. The downstream recipients of these reducing and/or oxidizing equivalents may include redox-active DNA-binding proteins.


Membrane
rnfD
rnfA
rnfE
Cytoplasm
rnfB
rnfC
rnfG
NADH
NAD+
Fe-S
Fe-S

References:

1. Berghuis, A.M., et al., *Mutation of Tyrosine-67 to Phenylalanine in Cytochrome-C Significantly Alters the Local Heme Environment.* Journal of Molecular Biology, 1994. **235**(4): 1326–1341.
2. Casimiro, D.R., et al., *Electron-Transfer in Ruthenium-Modified Cytochromes-C—Sigma-Tunneling Pathways through Aromatic Residues.* Journal of Physical Chemistry, 1993. **97**(50): 13073–13077.
3. Cheadle, J.P., and J.R. Sampson, *MUTYH-associated polyposis—From defect in base excision repair to clinical genetic testing.* DNA Repair, 2007. **6**: 274–279.
4. Chmiel, N.H., A.L. Livingston, and S.S. David, *Insight into the Functional Consequences of Inherited Variants of hMYH Adenine Glycosylase Associated with Colorectal Cancer: Complementation Assays with hMYH Variants and Pre-stead-state Kinetics of the Corresponding Mutated E. Coli Enzymes.* J. Mol. Biol., 2003. **327**: 431–443.
5. Boal, A.K., et al., *Redox signaling between DNA Repair Proteins for efficient lesion detection.* Proc. Natl. Acad. Sci. USA, 2009. **106**(36): 15237-–15242.
6. Gifford, C.M., and S.S. Wallace, *The genes encoding endonuclease VIII and endonuclease III in Escherichia coli are transcribed as the terminal genes in operons.* Nucleic. Acids Res., 2000. **28**(3): 762–769.
7. Curatti, L., et al., *Genes required for rapid expression of nitrogenase activity in Azotobacter vinelandii.* Proc. Natl. Acad. Sci. USA, 2005. **102**(18): 6291–6296.
8. Jouanneau, Y., et al., *Overexpression in Escherichia coli of the rnf genes from Rhodobacter capsulatus—characterization of two membrane-bound iron-sulfur proteins.* Eur. J. Biochem., 1998. **251**(1-2): 54–64.
9. Schmehl, M., et al., *Identification of a new class of nitrogen fixation genes in Rhodobacter capsulatus: a putative membrane complex involved in electron transport to nitrogenase.* Mol. Gen. Genet., 1993. **241**(5–6): 602–615.
10. Saez, L.P., et al., *Role for draTG and rnf genes in reduction of 2,4-dinitrophenol by Rhodobacter capsulatus.* J. Bacteriol., 2001. **183**(5): 1780–1783.
11. Imkamp, F., et al., *Dissection of the caffeate respiratory chain in the acetogen Acetobacterium woodii: identification of an Rnf-type NADH dehydrogenase as a potential coupling site.* J. Bacteriol., 2007. **189**(22): 8145-8153.
12. Biegel, E., S. Schmidt, and V. Muller, *Genetic, immunological and biochemical evidence for a Rnf complex in the acetogen Acetobacterium woodii.* Environmental Microbiology, 2009. **11**(6): 1438–1443.

13. Schmidt, S., E. Biegel, and V. Muller, *The ins and outs of Na+ bioenergetics in Acetobacterium woodii.* Biochimica Et Biophysica Acta-Bioenergetics, 2009. **1787**(6): 691–696.
14. Kumagai, H., et al., *Membrane localization, topology, and mutual stabilization of the rnfABC gene products in Rhodobacter capsulatus and implications for a new family of energy-coupling NADH oxidoreductases.* Biochemistry, 1997. **36**(18): 5509–5521.
15. Green, J., and M.S. Paget, *Bacterial redox sensors.* Nat. Rev. Microbiol., 2004. **2**(12): 954–966.
16. Koo, M.S., et al., *A reducing system of the superoxide sensor SoxR in Escherichia coli.* EMBO J., 2003. **22**(11): 2614–2622.
17. Demple, B., *Redox signaling and gene control in the Escherichia coli soxRS oxidative stress regulon—a review.* Gene, 1996. **179**(1): 53–57.
18. Lee, P.E., B. Demple, and J.K. Barton, *DNA-mediated redox signaling for transcriptional activation of SoxR.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(32): 13164–13168.
19. Saaf, A., et al., *Divergent evolution of membrane protein topology: the Escherichia coli RnfA and RnfE homologues.* Proc. Natl. Acad. Sci. USA, 1999. **96**(15): p. 8540–8544.
20. Gifford, C.M., J.O. Blaisdell, and S.S. Wallace, *Multiprobe RNase protection assay analysis of mRNA levels for the Escherichia coli oxidative DNA glycosylase genes under conditions of oxidative stress.* J. Bacteriol., 2000. **182**(19): 5416–5424.
21. Chen, P.R., et al., *An oxidation-sensing mechanism is used by the global regulator MgrA in Staphylococcus aureus.* Nat. Chem. Biol., 2006. **2**(11): 591-595.
22. Gralnick, J., and D. Downs, *Protection from superoxide damage associated with an increased level of the YggX protein in Salmonella enterica.* Proc. Natl. Acad. Sci. USA, 2001. **98**(14): 8030–8035.
23. Pomposiello, P.J., et al., *SoxRS-regulated expression and genetic analysis of the yggX gene of Escherichia coli.* J. Bacteriol., 2003. **185**(22): 6624–6632.
24. Gralnick, J.A., and D.M. Downs, *The YggX protein of Salmonella enterica is involved in Fe(II) trafficking and minimizes the DNA damage caused by hydroxyl radicals: residue CYS-7 is essential for YggX function.* J. Biol. Chem., 2003. **278**(23): 20708–20715.
25. Gifford, C.M., and S.S. Wallace, *The Genes Encoding Endonuclease VIII and Endonuclease III in Escherichia coli are Transcribed as the Terminal Genes in Operons.* Nucleic Acids Research, 2000. **28**(3): 762–769.

Appendix I

Clean Deletion of the *nth* Gene from the Chromosome of CC104 *Escherichia Coli*

Introduction

The methods described here are adapted from those published previously [1]. Briefly, overlap extension PCR is used to create a construct in which regions upstream and downstream of a gene of interest (*nth*) are amplified and fused together, devoid of the *nth* gene. This construct is then cloned into the plasmid pKOV [1], which is then transformed into the recipient strain, CC104. During this transformation reaction, the pKOV plasmid can integrate itself onto the CC104 chromosome by homologous recombination. A second recombination event then permits the plasmid backbone to excise itself from the chromosome, leaving only the construct that is devoid of *nth* gene on the chromosome.

Materials

All strains and plasmids were obtained from the Coli Genetic Stock Center at Yale University (New Haven, CT). All enzymes were obtained from New England Biolabs and primers were ordered from Integrated DNA Technologies.

Preparation of *nth* Deletion Construct by Overlap Extension PCR

In order to prepare the Δ*nth* construct, regions upstream and downstream of the *nth* gene were PCR amplified and these products were used as templates in a second PCR experiment, resulting in a product of ~ 2000 bp that omits the *nth* gene (Figure I-1). This product was ligated into the BamHI and NotI sites of pKOV.

The pKOV vector is temperature sensitive, encodes chloramphenicol resistance, and contains the *sacB* gene, which confers sensitivity to sucrose [2].

Plasmid Integration by Homologous Recombination

Electrocompetent CC104 cells were transformed with the pKOV plasmid containing the Δ*nth* construct and then recovered at 30°C. The temperature sensitivity of pKOV necessitates this low recovery temperature. During recovery, the plasmid has the opportunity to recombine onto the CC104 chromosome as shown in Figure I-2.

After recovery, the transformation reactions were plated onto Luria-Bertani (LB) + chloramphenicol medium and incubated at either 30°C or 42°C. Colonies incubated at 42°C are unable to replicate the temperature-sensitive pKOV plasmid and will only maintain their chloramphenicol resistance if this plasmid is integrated onto the chromosome. After incubation, the ratio of colony number on the 42°C plate to colony number on the 30°C plate can be used to calculate the integration efficiency.

Excision of pKOV from the Genome of CC104 to Create a Δ*nth* Strain

Integration of the pKOV plasmid onto the CC104 chromosome leaves the cells with two copies of the upstream and downstream *nth* regions, one from the WT cell and one from the pKOV plasmid. To create the CC104 Δ*nth* strain, one of these copies of the region must be excised. This excision process is also

recombination-based, as diagrammed in Figure I-3. Cells that have excised the pKOV plasmid backbone can be selected based on their tolerance of sucrose. The *sacB* gene of the pKOV plasmid encodes sensitivity to sucrose. If a cell has excised the plasmid backbone from its chromsome, then the cell should be able to grow in the presence of 5% sucrose. If the cell has not excised the pKOV backbone, then sucrose is lethal to the cell. The sucrose selection was conducted by taking cells that had been incubated at 42°C in the previous step, growing them in liquid culture, and then serially diluting these cultures and plating them onto LB + sucrose medium.

Colony PCR to Confirm the Absence of the *nth* Gene

Cells that are able to grow in the presence of sucrose have excised the pKOV plasmid from their chromosome, but they still may not be Δ*nth* strain. As Figure I-3 shows, the orientation of the chromosomal DNA during recombination may create a Δ*nth* strain, or it may cause the cell to revert to WT. Consequently, a final screen is needed to determine which of the sucrose-tolerant cells are truly Δ*nth* mutants. To verify that the *nth* gene has been deleted, PCR was used with primers that bind upstream and downstream of the region containing the *nth* gene. The size of the PCR product indicated whether the colony examined was WT or Δ*nth*. In this reaction, a ~ 2000 bp PCR product was expected of Δ*nth* cells, whereas a ~ 2600 bp product was expected of WT cells (Figure I-4). Of several colonies

examined, three appeared to have the Δ*nth* genotype when examined by colony PCR (Figure I-4).

After this PCR reaction, the samples that appeared to be Δ*nth* were isolated, re-streaked, and analyzed by colony PCR again. This process was repeated at least three times to ensure that the samples were not contaminated with WT CC104. Finally, sequencing was used to verify that the isolated strains were truly Δ*nth*.

References:

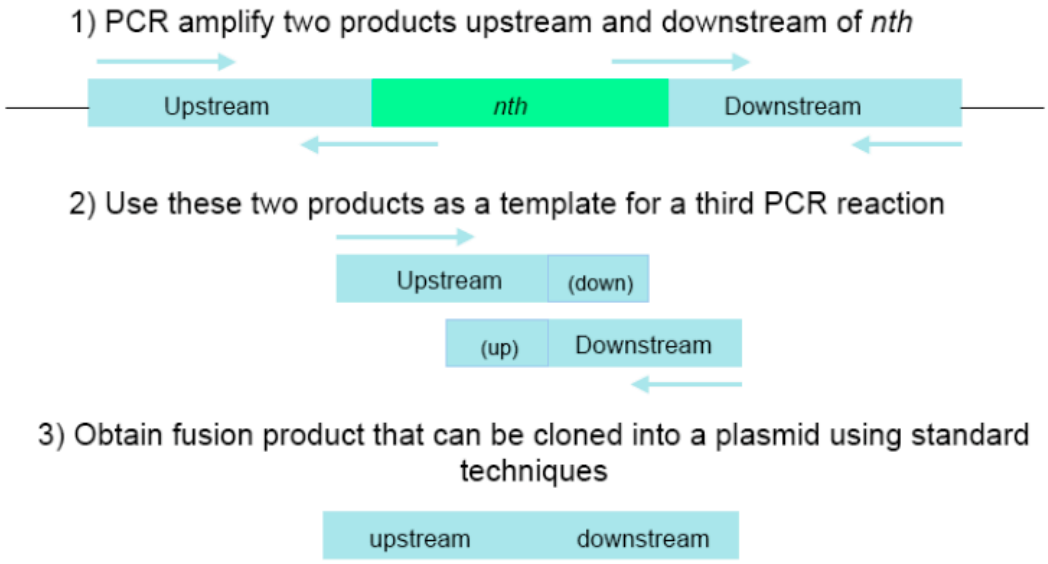
1. Link, A.J., D. Phillips, and G.M. Church, *Methods for Generating Precise Deletions and Insertions in the Genome of Wild-Type Escherichia coli: Application to Open Reading Frame Characterization.* Journal of Bacteriology, 1997. **179**(20): 6228–6237.
2. Dedonder, R., *Levansucrase from Bacillus subtilis.* Methods in Enzymology, 1966. **8**(C): 500–505.

Table I-1: Primers used in this protocol

Name	Sequence	Purpose
CARJKB172	5'-actagtgcggccgcgaactgcgcctttctgactggatc-3'	Primer for overlap extension to create Δ*nth* construct, and for verifying absence of *nth* on CC104 chromosome. It binds to a sequence 1000 bp upstream of *nth* gene.
CARJKB173	5'-ggcgcgccggatcccataaccaatcgccagcacaatagc-3'	Primer for overlap extension to create Δ*nth* construct, and for verifying absence of *nth* on CC104 chromosome. It binds to a sequence 1000 bp downstream of *nth* gene.
delnthfusfor	5'- gcattgccaaacggtgaaacagggaatgtctgatgaagaaaaggggtaacaccgattaccccattg-3'	Primer for overlap extension to create Δ*nth* construct. It binds to a sequence immediately following the *nth* gene.
delnthfusrev	5'-caatggggtaatcggtgttacccctttttcttcatcagacattccctgtttcaccgtttggcaatgc-3'	Primer for overlap extension to create Δ*nth* construct. It binds to a sequence immediately preceding the *nth* gene.
CARJKB174	5'-acattgtttgacggtgcaga-3'	Sequencing Δ*nth* strains
CARJKB175	5'-ggtcaatatttgttgcgtgtg-3'	Sequencing Δ*nth* strains
CARJKB180	5'-gcaatggcacattgtttgac-3'	Sequencing Δ*nth* strains

Figure I-1: Experimental schemes for clean deletion construct preparation. A) Schematic of how overlap extension PCR can be used to create a plasmid insert that contains a section of the CC104 chromosome devoid of *nth*. B) PCR results indication that regions 1000 bp upstream and downstream of the *nth* gene were successfully amplified and fused together through PCR.

A)



B)

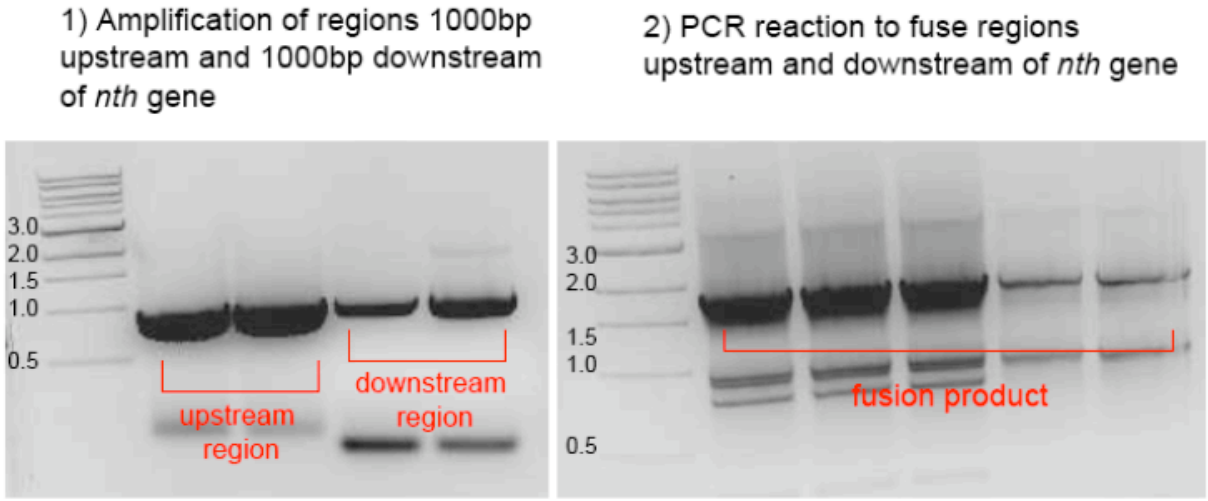


Figure I-2: Diagram of how pKOV can integrate onto the chromosome of CC104

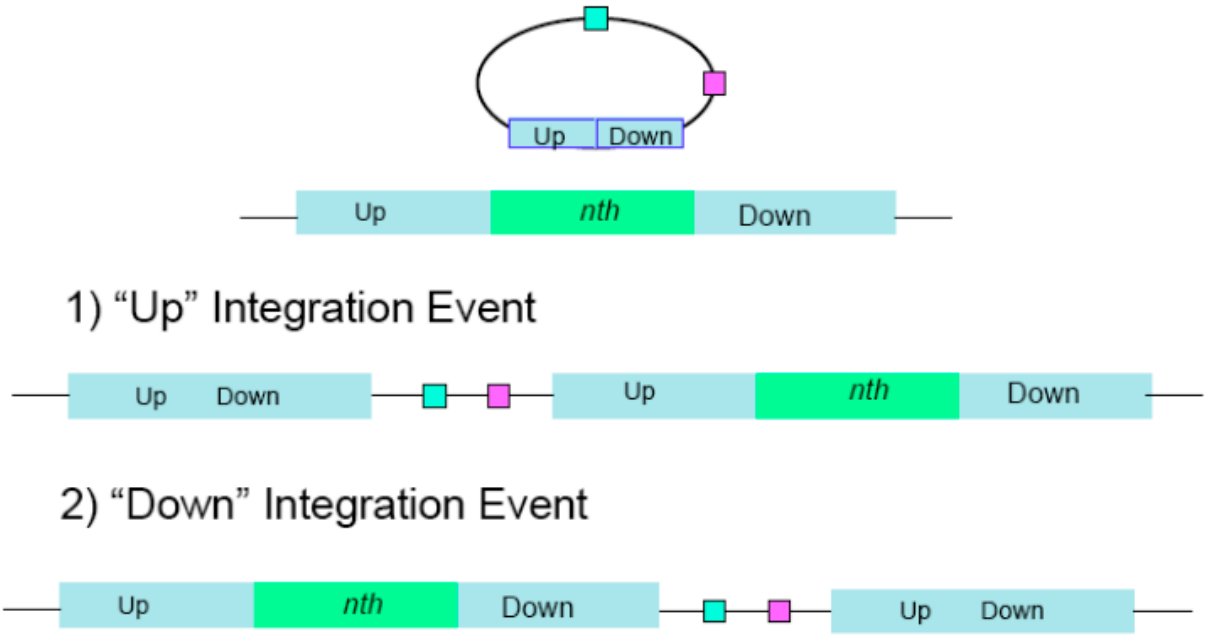


Figure I-3: Scheme for recombination-based excision of the pKOV plasmid. The plasmid should excise from the chromosome of CC104, yielding a single copy of the chromosomal region of interest. Note that several different products are available after this recombination/excision event, and further screens are necessary to select the strain of interest.

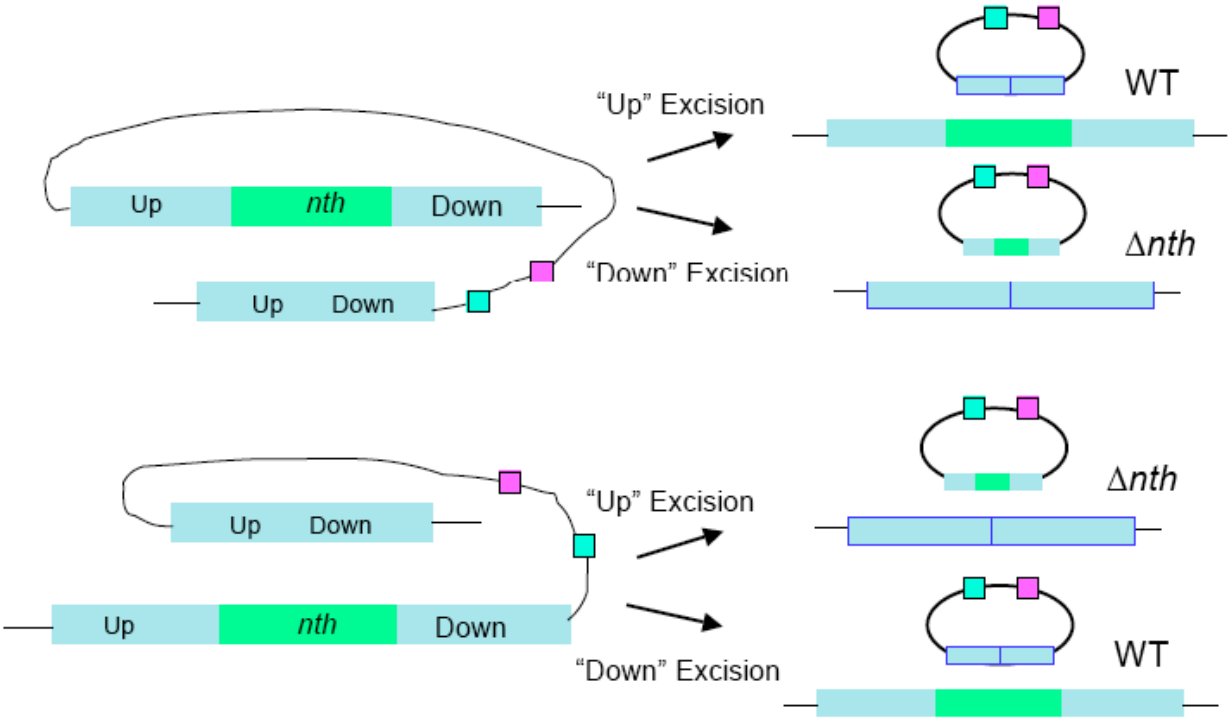
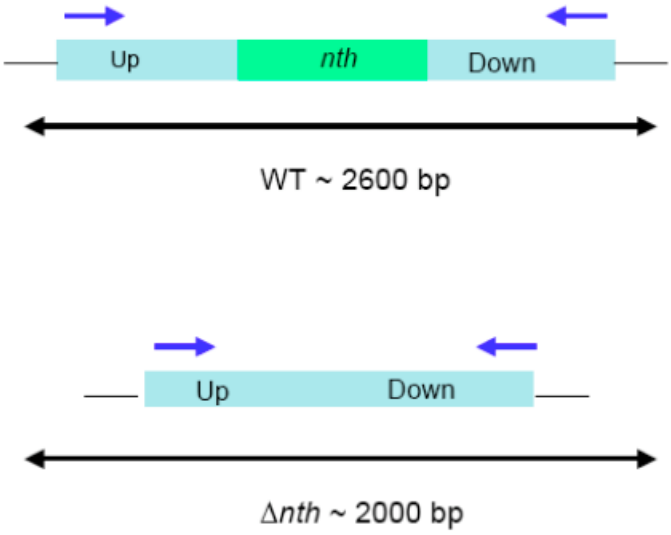


Figure I-4: Schematic for verification of clean deletion construct. A) Scheme of how colony PCR can be used to determine the presence or absence of the *nth* gene on the chromosome of CC104. B) Results of the screen of several colonies that detected three samples (colonies 19, 22, 23) that are possibly Δ*nth*. The lane labeled "WT" was a control performed with WT CC104, and the lane labeled "pKOV" was a control performed with the pKOV vector containing the Δ*nth* construct. The lane labeled "blank" was a cell-free control, and all other lanes are results using colonies taken from LB-sucrose plates.

A)



B)

Appendix II

Insertion of the *nth* Gene onto the Chromosome of *Escherichia Coli* at the λ Phage Attachment Site

Introduction

The methods described here are adapted from those previously published [1]. Briefly, EndoIII mutants of interested were cloned into a plasmid, pAH120 [1], which has the ability to integrate itself onto the chromosome of *Escherichia coli* at the λ phage attachment site. When this plasmid integrates, the allele of *nth* that it contains also integrates, thus creating a strain of CC104 that expresses a mutant version of the EndoIII protein.

Materials

All strains and plasmids were obtained from the Coli Genetic Stock Center at Yale University (New Haven, CT). All enzymes were obtained from New England Biolabs and primers were ordered from Integrated DNA Technologies.

Preparation of pAH120 with *nth* Alleles and Chromosomal Integration

EndoIII variants were PCR amplified from overexpression vectors and then ligated into the BamHI and NdeI sites of pAH120. Samples of pAH120 expressing different variants of *nth* were transformed into electrocompetent CC104Δ*nth* cells that were also expressing the plasmid pINT-ts. The latter plasmid is temperature sensitive and encodes integrase genes that help the pAH120 plasmid integrate onto the bacterial chromosome. Following transformation, the cells were incubated at

30ºC for one hour, during which time pAH120 may integrate onto the chromosome, and then at 42ºC for 30 minutes, during which time pINT-ts is cured. Cultures were then spread onto Luria-Bertani (LB) + kanamycin media and incubated at 37ºC. This final incubation selects for cells that have integrated pAH120, since this plasmid confers kanamycin resistance.

Colonies that emerge on these kanamycin selection plates were then screened using colony PCR to verify the presence of the *nth* variant on the chromosome. Colony PCR was performed with primers specific to the *nth* gene. A ~ 600 bp product is expected if the samples contain an *nth* variant. As Figure II-1 demonstrates, all colonies examined appeared to contain a variant of *nth*.

Verification That Only One Plasmid Integration Event Occurred

A further PCR screen is then needed to verify that the pAH120 plasmid integrated only once onto the chromosome of a given sample, and not multiple times. The integration events and PCR screen are diagrammed in Figure II-2. Primers were designed such that the size of a PCR product they produce indicates how many times the pAH120 plasmid integrated. Most of the colonies screened only contain one copy of the plasmid on their chromosome (Figure II-2).

References:

1. Haldimann, A., and B.L. Wanner, *Conditional-Replication, Integration, Excision, and Retrieval Plasmid-Host Systems for Gene Structure–Function Studies of Bacteria.* Journal of Bacteriology, 2001. **183**(21): 6384–6393.

Table II-1: Primers used in these experiments

Name	Sequence	Purpose
CARJKB164	5'-ggcgcgcccatatgaataaagcaaaacgcctggagatc-3'	Primer for cloning *nth* alleles into pAH120, NdeI site
CARJKB165	5'-ggcgcgccggatcctcagatgtcaactttctctttgtattc-3'	Primer for cloning *nth* alleles into pAH120, BamHI site
CARJKB176	5'-cgaattcaggcgctttttag-3'	Sequencing *nth* alleles after chromosomal insertion
CARJKB177	5'- tctgctggaaccactttcagt-3'	Sequencing *nth* alleles after chromosomal insertion
CARJKB178	5'-tgaaagtggttccagcagag-3'	Sequencing *nth* alleles after chromosomal insertion
CARJKB179	5'- gggagtgggacaaaattgaa-3'	Sequencing *nth* alleles after chromosomal insertion

Figure II-1: Results of colony PCR to verify presence of *nth* variants. PCR products were amplified from the chromosome of CC104 Δ*nth* after incubation with pAH120. All samples examined appear to have integrated the plasmid and the *nth* variant of interest.

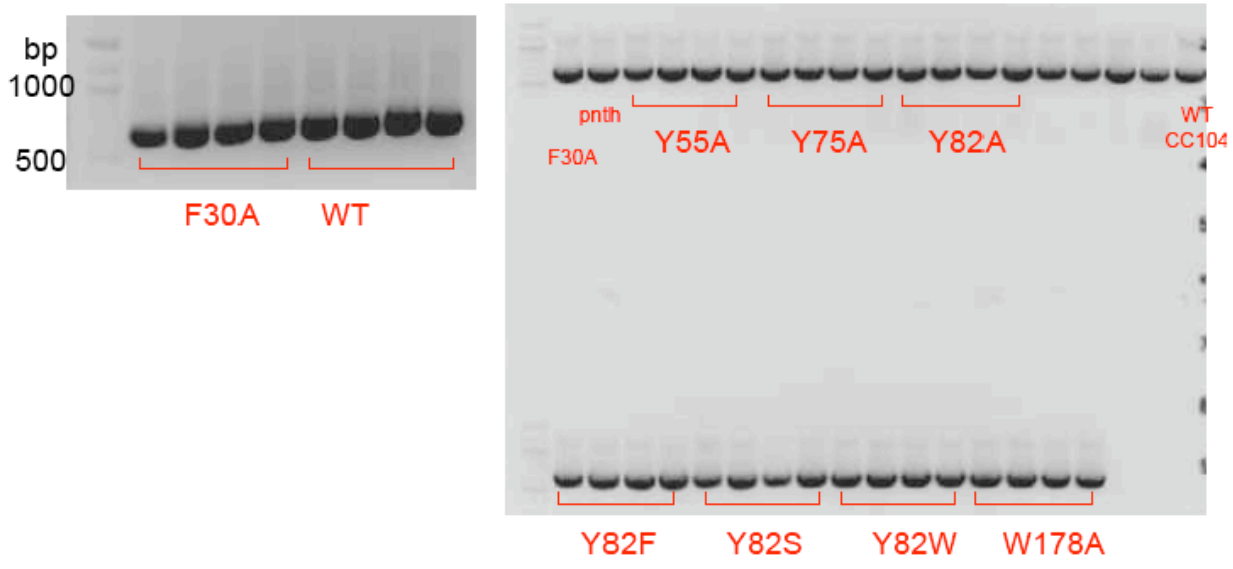
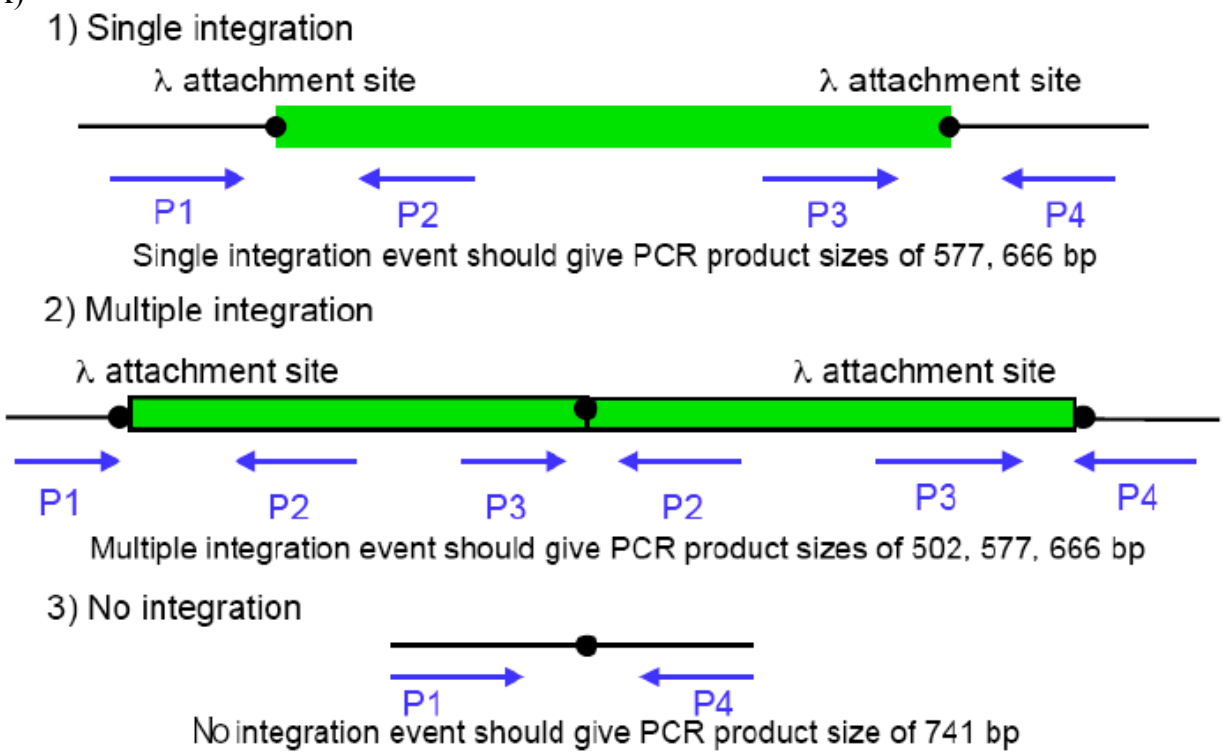
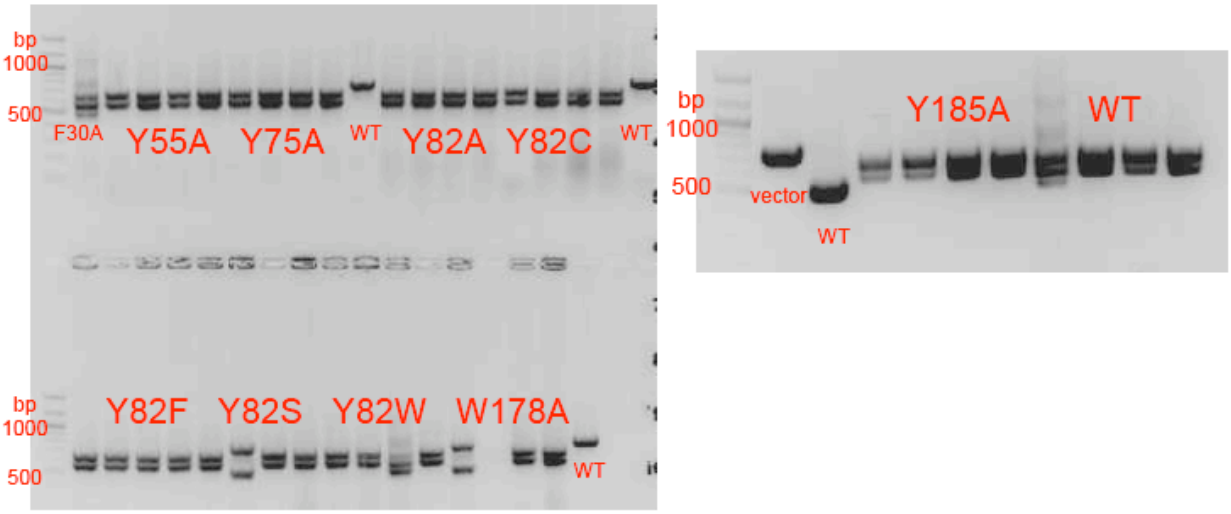


Figure II-2: Diagram of PCR experiments to verify single insertion. A) Diagram of how colony PCR can be used to verify that pAH120 only inserted once and not multiple times. Single integration, multiple integration events, and no integration events will yield products of different sizes. B) Results of colony PCR performed on several samples containing *nth* variants. Most samples display PCR products whose sizes indicate a single integration event.

A)



B)



Appendix III

Insertion of *nth* onto the Chromosome of *Escherichia Coli* Using the Tn7 Transposon

Introduction

The methods described here are adapted from those published previously [1]. The insertion process is also diagrammed in Figure III-1. Briefly, a suicide plasmid containing the Tn7 transposon [2] and the *nth* gene is prepared and then transformed into a donor strain, BW29427. Simultaneously, a separate sample of BW29427 is transformed with the plasmid pUX-BF13, which contains transposase genes that help the Tn7 transposon insert itself onto the *E. coli* chromosome. These two strains of BW29427 are then mated with a recipient, CC104.

During the mating, the two plasmids should enter cells of CC104, and the Tn7 transposon and *nth* gene will be inserted onto the CC104 chromosome. The mating reaction is then re-streaked onto media on which the BW29427 donor cells cannot grow, thus selecting for the CC104 recipient cells. The remaining steps of the protocol describe how to examine these CC104 cells to ensure that they received the Tn7 and *nth* genes, that they have been cured of the donor plasmids, and that the transposition reaction occurred at the correct genomic location.

Materials

The strain CC104 was obtained from Professor Jeffrey Miller at UCLA (Los Angeles, CA), CC104 *nth*$^-$ was obtained from Dr. Amie Boal, BW29427 was obtained from Professor Barry Wanner (Purdue University). The plasmids pURR24 and pUX-BF13 were obtained from Professor Dianne Newman at the

California Institute of Technology. Diaminopimelic acid (DAP), ampicillin, and spectinomycin were obtained from Sigma-Aldrich. Primers were ordered from Integrated DNA Technologies. Enzymes were purchased from New England Biolabs. All media and buffers were prepared according to standard procedures for *E. coli* [3].

Preparation of Donor Strains for Conjugation

The *nth* gene was amplified from a template provided by Dr. Amie Boal (Caltech). The PCR product was ligated into the XmaI site of pURR24 to create pURR24nth. The plasmid pURR24 contains a miniTn7 transposon and it is also a suicide plasmid, meaning it cannot be replicated by many strains of *E. coli*. Consequently, these strains can easily be cured of the plasmid once transposition has taken place.

The plasmid pURR24nth was then transformed into BW29427 *E. coli*. A separate sample of BW29427 was transformed with the plasmid pUX-BF13, which contains the transposase genes that mediate the transposition of Tn7 and *nth* onto the chromosome of an *E. coli* cell. BW29427 cells are unable to make diaminopimelic acid (DAP), a reagent that *E. coli* cells use to maintain the rigidity of their cell walls [4]. The Luria-Bertani (LB) media for BW29427 was supplemented with 300 μM DAP.

Mating Reaction

Overnight cultures of BW29427 with pURR24nth, BW29427 with pUX-BF13, and CC104 *nth*$^-$ were prepared. A 100 μL aliquot of each culture was removed and placed in a sterile microcentrifuge tube. Control experiments were also performed in which only two of the three strains were mixed in the microcentrifuge tube. The tubes were centrifuged at 8000 rpm for two minutes. Their supernatant was decanted, and the cells were resuspended in 50 μL fresh Luria-Bertani (LB) medium, supplemented with DAP. This entire solution was then spread onto an LB + DAP plate and incubated for 3 hours at 37°C.

A sample of cells was collected from each of the plates using a sterile stick and re-suspended in 1 mL fresh LB (without DAP). This solution of cells was serially diluted, and various dilutions were plated onto LB + spectinomycin media and incubated at 37°C. Colonies should only appear if an integration event has occurred because this event would confer spectinomycin resistance. Transconjugant colonies were only observed on plates in which all three strains (BW29427 + pURR24nth, BW29427 + pUX-BF13, and CC104*nth*$^-$) had been mixed (Figure III-2).

Screening of Cells to Verify the Accuracy and Location of Transposition

After the plates have incubated overnight, the colonies that grow need to be examined to ensure that several processes have taken place: 1) The CC104 recipient cells have been cured of any plasmids, 2) Only the *nth* gene and Tn7 transposon

have integrated into CC104, not the entire pURR24 plasmid, 3) The *nth* gene and Tn7 transposon have integrated themselves into the CC104 chromosome, 4) The *nth* gene and Tn7 transposon inserted at the Tn7 attachment site [5, 6] and not elsewhere in the CC104 chromosome, and 5) Only one insertion event occurred, and the recipient cells do not contain multiple copies of the *nth* gene.

To screen for the absence of plasmids in the recipient CC104 *nth*$^-$ cells, a plasmid prep protocol can be performed on several colonies, and the results run on an agarose gel. No plasmids should be visible. When examining cells for the absence of pURR24 plasmid backbone, several screening methods can be applied. Select transconjugants can be re-streaked on LB media containg spectinomycin and 5% sucrose. The pURR24 plasmid backbone contains the *sacB* gene, which encodes sensitivity to sucrose [7]. Therefore, cells that contain pURR24 will not grow on sucrose medium. To enhance the cells' sucrose sensitivity, sodium chloride can be omitted from their growth medium. Simultaneously, transconjugants can be re-streaked onto LB-ampicillin media. Because the pURR24 backbone encodes ampicillin resistance, colonies that have removed it should be ampicillin sensitive. When subjected to these tests, the transconjugants examined displayed sucrose tolerance and ampicillin sensitivity (Figure III-3). PCR with primers specific to the pURR24 backbone further validated the plasmid's absence (Figure III-4).

To ensure that the Tn7 transposon and *nth* gene integrated successfully onto the chromosome of the CC104 recipient, PCR experiments were performed with

primers specific to the transposon-containing region of pURR24. The primers were designed such that the size of the PCR product indicates whether the Tn7 transposon and/or the *nth* gene have integrated onto the chromosome. PCR results indicated that the transconjugants screened did contain the Tn7 transposon and *nth* gene (Figure III-5).

When the Tn7 transposon and *nth* gene insert onto the chromosome of CC104, it is important that their orientation relative to other chromosomal genes be monitored. The screen for proper orientation is also accomplished by PCR (Figure III-5). Primers were designed such that the size of the PCR product indicates whether the gene or transposon is properly aligned on the chromosome. Several transconjugants with the appropriate alignment were obtained. The final screen of the transconjugants will verify that the *nth* gene and Tn7 transposon only inserted in single copy. This screen was not performed for the transconjugants made here, but can be accomplished by sequencing or southern blot.

References:

1. Choi, K.-H., and H.P. Schweizer, *1mini-Tn7 insertion in bacteria with single attTn7 sites: example Pseudomonas aeruginosa.* Nature Protocols, 2006. **1**(1): 153–161.
2. Peters, J.E., and N.L. Craig, *Tn7: Smarter that we Thought.* Nature Reviews: Molecular Cell Biology, 2001. **2**: 806–813.
3. Miller, J.H., *Experiments in Molecular Genetics*. 1972, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
4. Baumann, R.J., et al., *Inhibition of Escherichia coli Growth and Diaminopimelic Acid Epimerase by 3-Chlorodiaminopimelic Acid.* Antimicrobial Agents and Chemotherapy, 1988. **32**(8): 1119–1123.
5. Gringauz, E., et al., *Recognition of Escherichia-Coli AttTn7 by Transposon Tn7—Lack of Specific Sequence Requirements at the Point of Tn7 Insertion.* Journal of Bacteriology, 1988. **170**(6): 2832–2840.
6. Mckown, R.L., et al., *Sequence Requirements of Escherichia-Coli AttTn7, a Specific Site of Transposon Tn7 Insertion.* Journal of Bacteriology, 1988. **170**(1): 352–358.
7. Dedonder, R., *Levansucrase from Bacillus subtilis.* Methods in Enzymology, 1966. **8**(C): 500–505.

Table III-1: Primers used in these experiments

Name	Sequence	Purpose
CARJKB6	5'-cccgggtcaatggggtaatcggtgtt- 3'	Cloning of *nth* gene into pURR24, XmaI site
CARJKB8	5'- cgccccggggtggatcctcaatggggtaa-3'	Cloning of *nth* gene into pURR24, XmaI site
CARJKB14	5'-gcgagggctttactaagctg -3'	Verifying presence of *nth* gene in pURR24, and sequencing of this insert
CARJKB15	5'-tccagttatgctgtgaaaaagc-3'	Verifying presence of *nth* gene in pURR24, and sequencing of this insert
CARJKB32	5'-tcgctataatgaccccgaag-3'	Primer is specific to pURR24 backbone, used to verify pURR24 absence in a CC104 sample
CARJKB36	5'-cggtttgtcacatggagttg-3'	Primer matches a region just upstream of Tn7 attachment site, used to verify presence of Tn7 in CC104 sample
CARJKB37	5'-gcaggccaaccagataagtg-3'	Primer matches a region of the Tn7 transposon, used to verify transposon's presence in CC104 sample
CARJKB38	5'-tgctttttcacagcataactgg-3'	Primer matches a region of the Tn7 transposon, used to verify transposon's presence in CC104 sample
CARJKB39	5'- cgataacatgcacatcatcgag-3'	Primer matches a region just downstream of Tn7 attachment site, used to verify presence of Tn7 in CC104 sample
CARJKB42	5'-ttcggtcctccgatcgttgt-3'	Primer is specific to pURR24 backbone, used to verify pURR24 absence in a CC104 sample
CARJKB43	5'-ttctgctatgtggcgcggta-3'	Primer is specific to pURR24 backbone, used to verify pURR24 absence in a CC104 sample
CARJKB44	5'-ttttcacagcataactggactga-3'	Primer matches a region of the Tn7 transposon, used to verify transposon's presence in CC104 sample
CARJKB45	5'-gaagcgctggcagaagattt-3'	Primer matches a region just downstream of Tn7 attachment site, used to verify presence of Tn7 in CC104 sample

Figure III-1: Overview of gene insertion using the Tn7 transposon

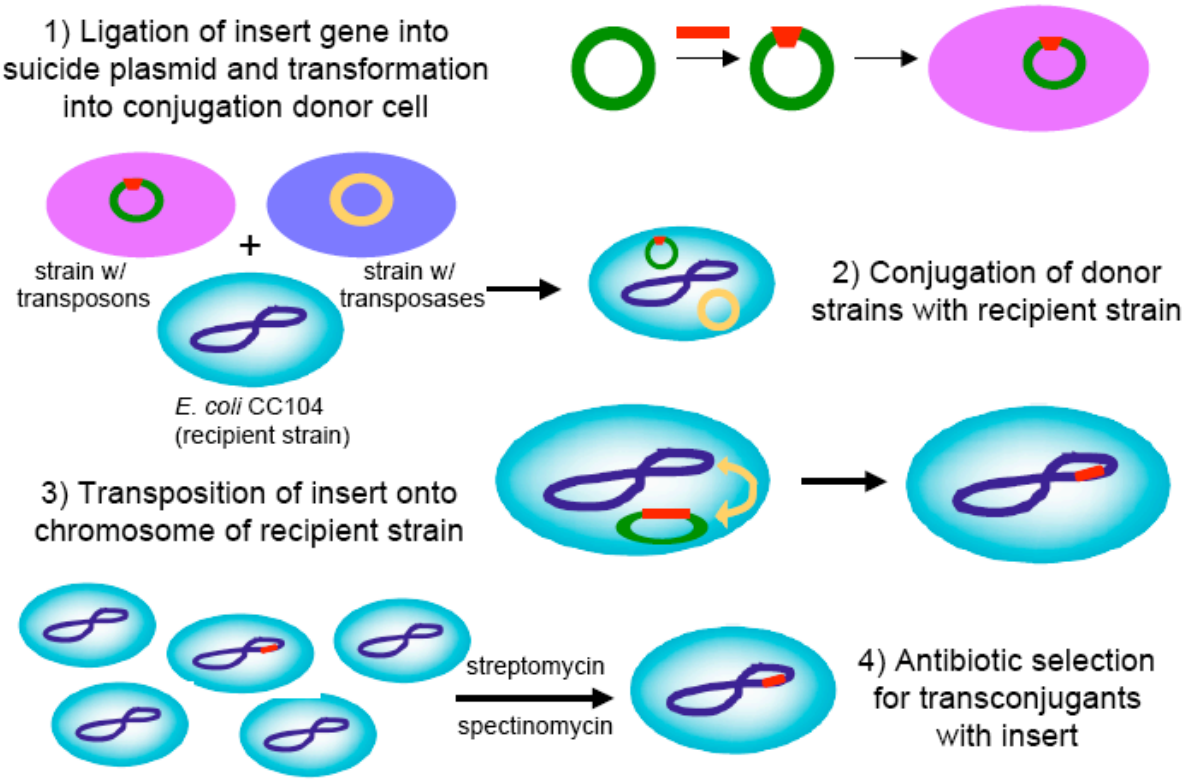


Figure III-2: Results of mating showing that transconjugant colonies form only when all three parent strains are used in the mating reaction

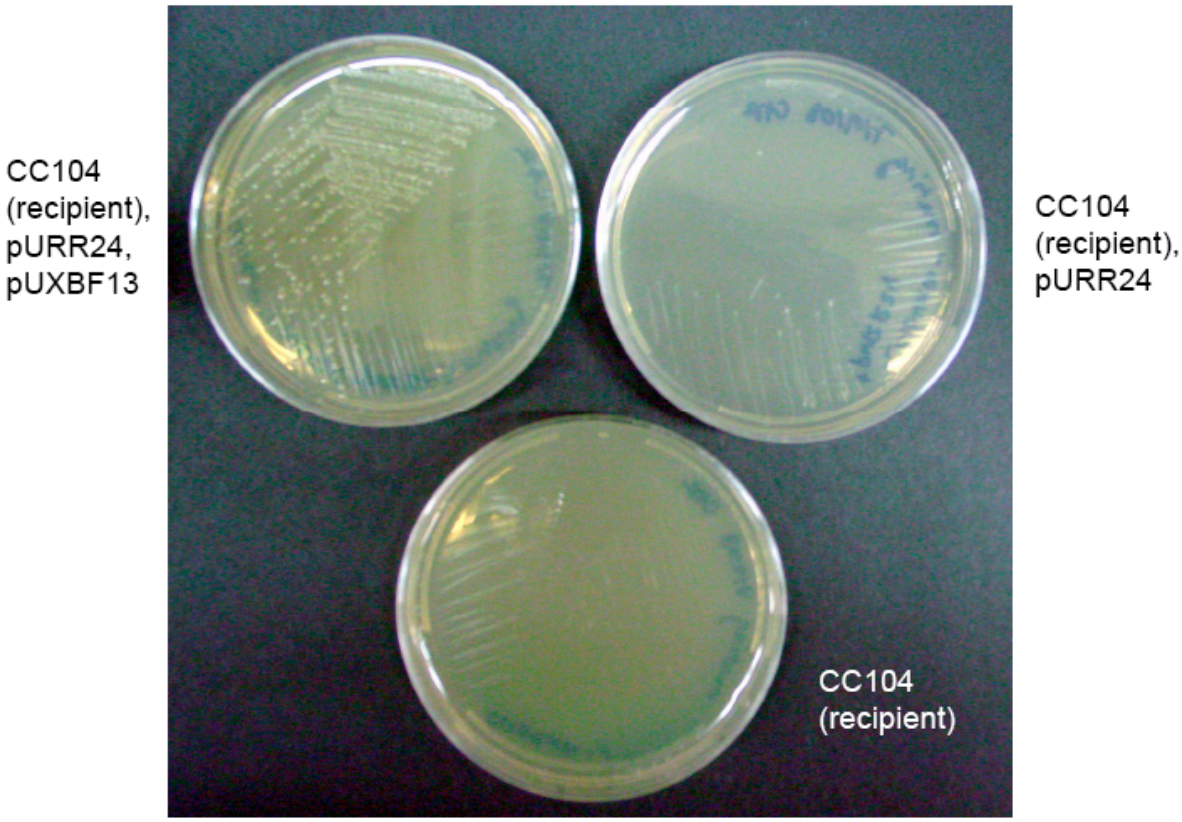


Figure III-3: Transconjugants were re-plated on sucrose-containing and ampicillin-containing media. The results are shown for sample number eight. If the pURR24 backbone is not present on the strain's chromosome, then the strain should be sucrose-tolerant and ampicillin-sensitive.

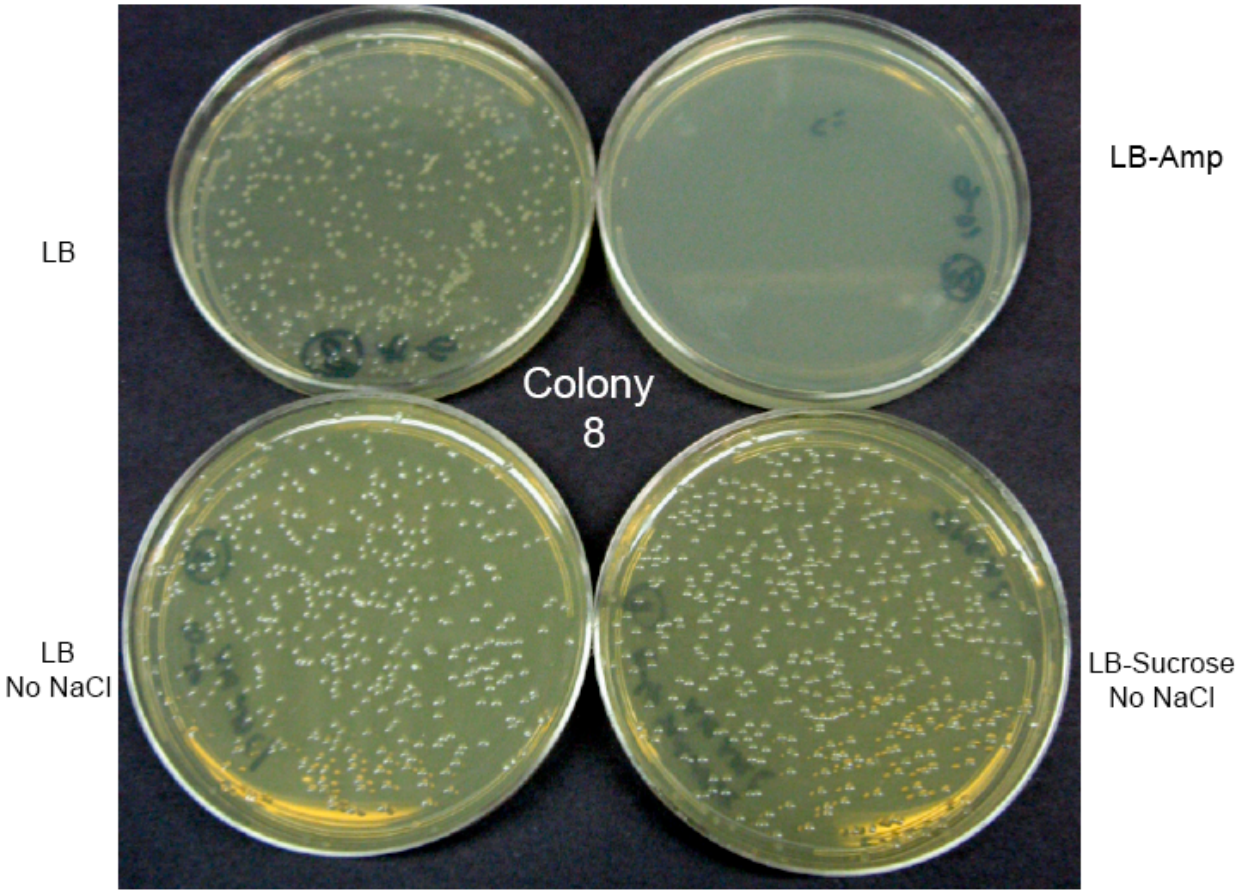


Figure III-4: Results of PCR experiments with primers specific to the backbone of pURR24. Since no product of the appropriate size is detected, this result corroborates those described in Figure III-3 and suggests that the pURR24 backbone is not present in most of the colonies examined.

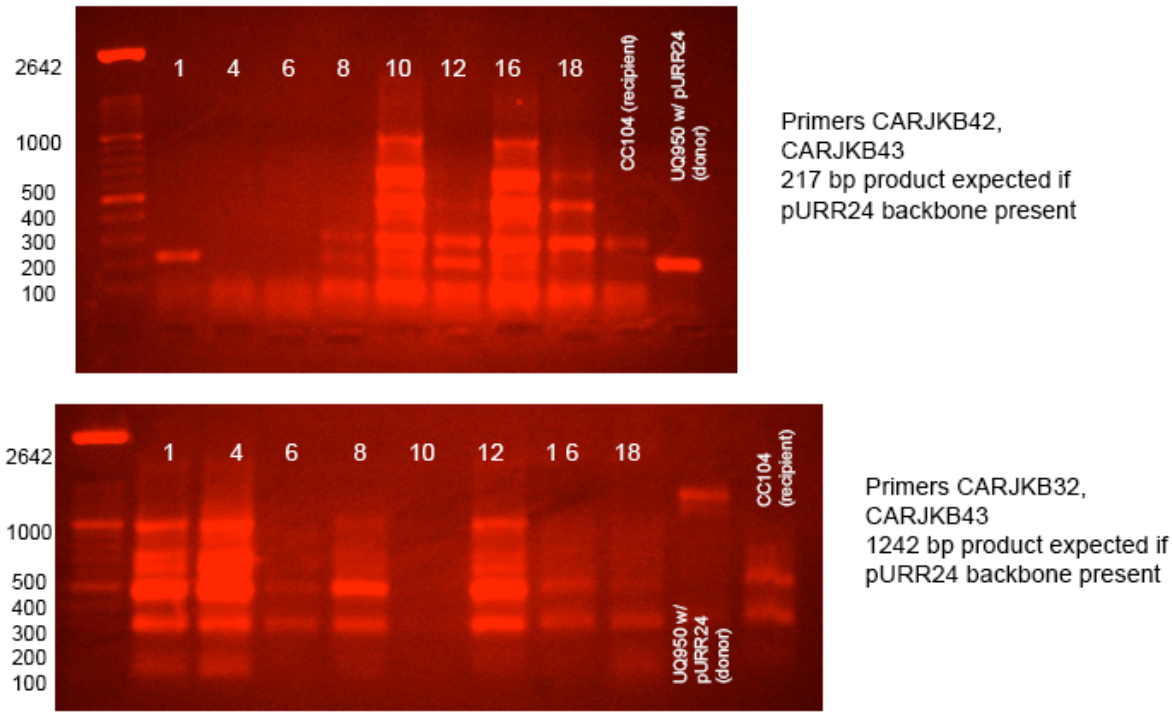


Figure III-5: Results of PCR experiments with primers specific to the Tn7 transposon and *nth* gene. These experiments demonstrate that the *nth* gene and Tn7 transposon are present on the chromosomes of the samples examined, and also that the insert is in the correct orientation relative to the genes that flank the Tn7 attachment site.