Elucidating the Role of [4Fe4S] Clusters in DNA Replication and Repair Proteins

Thesis by

Phillip L. Bartels

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy in Chemistry

California Institute of Technology

Division of Chemistry and Chemical Engineering

Pasadena, CA 91125

(Defended February 5, 2018)

© 2018

Phillip Leon Bartels

All Rights Reserved

ORCID ID: 000-0002-9688-6592

Acknowledgements

Many people have provided invaluable support and assistance throughout the course of this work, and I am pleased to acknowledge them here. First, I owe a great deal of thanks to my advisor, Prof. Jackie Barton, for taking me on as a graduate student and allowing me to work on so many interesting projects. In addition to providing guidance and helpful feedback on my research, you also gave me first-hand experience in the publishing and review processes. Altogether, you have taught me a great deal about how science is done, and I am deeply grateful.

I must also thank my committee members, Profs. Harry Gray, Judy Campbell, and Dave Tirrell, whose feedback and support regarding both research and the progress of my career has been invaluable. All of you have helped make my science the best that it can be, and I greatly appreciate your assistance.

In addition to my committee, I would also like to thank Profs. Peter Dervan and Andre Hoelz for providing me the opportunity to TA for a freshman chemistry seminar for four sequential years. I learned a great deal in this time, and the experience in working with undergraduates was invaluable. Dr. Jeff Mendez, for whom I served as a TA for a freshman lab course in my first year, also deserves thanks for teaching me how to organize lab courses and fairly evaluate students.

Thanks are also due to several collaborators on various projects. First among these, I wish to thank Prof. Peter Burgers and his graduate student, Joe Stodola, for their assistance in the polymerase project that forms three chapters of this thesis. Your insights and suggestions greatly improved this project. I would also like to thank Joe for sending numerous protein samples from St. Louis over the last six years; I realize how much effort went into this, and I appreciate your willingness to sacrifice so much valuable protein for my strange-sounding experiments. With

respect to this project, thanks are also due to Carrie Stith, who was of great help in experiments with polymerase mutants.

Much gratitude is also owed to all of the collaborators on the MUTYH project that constitutes the longest single chapter of this thesis. In particular, Dr. Kevin McDonnell and Dr. Stephen Gruber deserve special thanks for providing us the opportunity to work on a clinically relevant mutant protein. I would also like to thank Kevin specifically for providing large amounts of purified protein to work with. Elizabeth O'Brien was also of great help in purifying DNA and running electrochemical experiments.

Prof. Frank Crespilho, Nicole Nuñez, Prof. Sheila David, Yang Ha, Prof. Keith Hodgson, and Prof. Ed Solomon were all of great help in our collaborative efforts with graphite electrochemistry and X-ray absorbance spectroscopy. This work would never have happened without Frank's expertise with graphite electrodes, and I owe him much thanks for teaching me these techniques and for many fruitful discussions. Sheila was a great source of information about MutY, and her feedback on numerous drafts of these papers was greatly appreciated. I am grateful to Nicole for providing MutY for experiments; this made our first paper much more complete. I thank Yang for carrying out the XAS experiments as well as provoking many interesting discussions. Among Barton lab members, Andy Zhou and Anna Arnold deserve thanks for producing enormous quantities of highly concentrated protein for electrochemistry.

Lastly, I would like to thank Aaron Ehlinger and Prof. Walter Chazin for providing purified RPA for our experiments in DNA charge transport. This effort has only begun, and should prove to be very interesting.

All of these collaborations have been very fruitful, and each project has benefitted enormously from the input of so many insightful minds and talented hands.

Numerous staff at Caltech have provided indispensable aid and training with specialized instrumentation. These include Mona Shahgholi in the mass spec lab, Angelo DiBilio and Paul Oyala with the EPR instruments, and Melissa Melendes, Nils Asplund, Steven Martinez, and Guy DeRose in the clean room at the Caltech's Kavli Nanoscience Institute.

Thanks are due to all of the Barton group members with which I have had the honor of overlapping with; your feedback and support made the Barton lab a wonderful place to work. There have been too many to thank each individually, but several deserve special mention. First, I would especially like to thank Ariel Furst for being an excellent friend and coworker whose support inside and outside lab was invaluable. I am grateful to Tim Mui for training me in the lab and for being a great friend in both his time as a student and afterwards. Ariel and Tim have also introduced me to a world of high-quality wines, and they deserve an additional special thanks for this alone. Alyson Weidmann was both a wonderful co-TA for three years, and a good friend outside the lab. Catrina Pheeney trained me in the clean room and introduced me to device fabrication; this skill set is incredibly useful to have, and I am deeply grateful to her for taking the time to teach me. Andy Zhou and Rebekah Silva made the North side of the lab a fascinating place to work with many fascinating discussions about science and life. Elizabeth O'Brien has been of great help across a number of projects, and her assistance in the clean room made a demanding group job much easier.

None of this work would have been possible without the Barton lab administrative assistants. I have had the good fortune to work with three of them: Mo Renta, Elisha Jung-Okawa, and Elizabeth Garcia. All have been fantastic to work with, and were incredibly helpful in dealing with the surprising difficulties of processes like paper submission. A special thanks is

owed to Mo for introducing me to the Barton lab and providing a nice break from science whenever it was necessary.

Outside of lab, I would like to thank Ed Wagner, Nick Trojanowski, and James Schlabach for remarkable and rewarding friendships that have lasted since college. During graduate school, I spent a great deal of time at the Caltech gym, where I met Sylvia and Tony Delgado, Jack Lampley, Robert Storts, Guillermo Ramirez, Jorge Palamara, and Dave Eggeleston. I could not imagine a more entertaining group with which to spend the early morning hours, and their friendship outside of the gym has meant a great deal.

Finally, I wish to thank my family for their support through this time. First, my parents, Leon and Connie Bartels, are owed a special thanks for supporting me through college and providing me the opportunity to work at Caltech and pursue a career in science. My aunt Kathleen Bartels deserves thanks for providing me a place to spend Thanksgiving and the occasional weekend throughout graduate school. I also wish to thank my grandfather, Walter Bartels, for being a source of valued support since childhood and over the course of my education.

Abstract

[4Fe4S] clusters, redox cofactors, have been discovered in DNA processing enzymes ranging from bacterial base excision repair glycosylases to eukaryotic DNA polymerases. Bacterial repair proteins are activated toward redox activity when bound to DNA and can take advantage of DNA-mediated charge transport (DNA CT) to search the genome for lesions. DNA CT involves the rapid transport of charges through the π-stacked base pairs and is sharply attenuated in the presence of lesions, mismatches, or other stacking perturbations. Thus, [4Fe4S] repair proteins use this chemistry to rapidly redistribute to target lesions and communicate with one another over long distances.

The general function of [4Fe4S] clusters in bacterial DNA repair has received much attention, but previous efforts have left several critical questions unanswered. First, while the redox potential of these proteins is affected by DNA binding, the relative importance of the negatively-charged DNA, the protein environment surrounding the cluster, and solvent has remained unclear. Second, the importance of [4Fe4S] clusters and DNA CT to human disease has never been directly addressed. The biological consequences of this chemistry are certainly a pressing issue, as numerous disease-relevant mutations in the human homologues of well-studied repair proteins have been recorded. Finally, the existence of [4Fe4S] clusters in eukaryotic DNA replication proteins in general, and in the B-family DNA polymerases in particular, was entirely unexpected. The function of the [4Fe4S] cluster in replication proteins was far from obvious, and the functional differences from repair proteins made them difficult to explain even in the context of CT signaling. Herein, these questions have been addressed using a combination of electrochemical, spectroscopic, and biochemical approaches.

First, we describe the use of pyrolytic graphite edge electrodes (PGE) and S K-edge X-ray absorption spectroscopy (XAS) to address the influence of protein environment, DNA, and solvation on the [4Fe4S] cluster redox potential in the bacterial base excision repair glycosylases endonuclease III (EndoIII) and MutY. The PGE surface is rough and favorable for protein binding; electron transfer can be further enhanced in the presence of carbon nanotubes. Electrochemical signals for EndoIII and MutY in the absence of DNA are large and reproducible, and a potential shift upon DNA binding is observed. With respect to studying proteins in the absence of DNA, the PGE electrode represents a significant advance over previously used highly-oriented pyrolytic graphite (HOPG), which is hydrophobic and difficult to prepare. To test the effect of protein environment on redox potential, a series of EndoIII point mutants were prepared in which the charge within 5 Å of the cluster was reversed or added in. None of these mutations induced a significant shift in redox potential relative to wild type, arguing that DNA electrostatics are the dominant factor in potential modulation. In parallel, XAS studies were performed on EndoIII and MutY in the presence and absence of DNA, and in the presence and absence of solvent. Ligating cysteinyl thiols and inorganic S atoms in the [4Fe4S] cluster absorb at different intensities in XAS depending on solvent environment and local electrostatics; these changes, in turn, directly correlate to redox potential. By XAS, DNA was found to induce a significant shift in absorbance, and thus potential; the removal of solvent had a smaller effect. Together, these studies provide new approaches for the study of DNA-binding [4Fe4S] proteins and reveal the critical role of DNA in tuning the redox potential.

Second, we report on a novel mutation in human MUTYH identified from a colorectal cancer patient and confirmed to be pathological. MUTYH is responsible for repairing certain lesions induced by oxidative stress and is thus frequently implicated in cancer. This new variant,

C306W, contains a mutation in one of the cysteines that ligates the [4Fe4S] cluster. Electrochemistry, activity and DNA binding assays, and spectroscopic analyses were performed for C306W alongside wild type MUTYH and two other disease-relevant mutants, Y179C and G396D, with an unaltered cluster environment. From this work, it is now clear that C306W can still bind a cluster, but it is susceptible to oxidative degradation to the $[3Fe4S]^{+}$ state upon redox signaling in an aerobic environment. Consequently, enzymatic activity is very low, and DNA binding is poor. Overall, this represents the first complete characterization of the [4Fe4S] cluster in a human homologue of MutY, and the first demonstration of pathology resulting from a mutation that primarily affects the [4Fe4S] cluster.

Moving into DNA replication proteins, we report on the characterization of the [4Fe4S] cluster in yeast DNA polymerase (Pol) δ, the eukaryotic lagging strand polymerase. Pol δ shows reversible electrochemical signals at a midpoint potential indistinguishable from EndoIII under the same conditions, and EPR spectroscopy confirms use of the $[4Fe4S]^{3+/2+}$ couple. The electrochemical signal is attenuated on DNA containing an abasic site or a CA mismatch, confirming that Pol δ is capable of DNA-mediated signaling. Bulk electrolysis and photooxidation were used to oxidize Pol δ under anaerobic conditions, and activity assays were carried out using oxidized or untreated protein. Oxidation stalls replication activity, while electrochemical reduction of oxidized samples restores activity to untreated levels. These results thus reveal that cluster oxidation serves as a reversible switch regulating Pol δ activity, suggesting an *in vivo* role in responding to replication stress, especially oxidative stress. In an effort to address these possibilities, we have carried out preliminary efforts in the characterization of two potentially CT-deficient mutants, W1053A and Y1078A. Both mutants

were found to be too structurally unstable to proceed with *in vivo* experiments, but they can serve to guide future efforts in this direction.

Finally, a strategy to examine charge transport through RPA-bound single-stranded DNA is reported. RPA is the eukaryotic single-stranded binding protein and forms a protective coat around vulnerable unwound DNA at replication forks. Given the importance of redox signaling in replication proteins, we aimed to use photooxidation experiments to determine if CT through RPA is a viable pathway; if so, this would open up a large set of long-range transfer pathways to [4Fe4S] proteins in replication. These efforts are ongoing, but the experimental strategy and initial efforts are discussed.

Published Content and Contributions

1. Bartels, P.L.; O'Brien, E.; Barton, J.K. DNA Signaling by Iron-Sulfur Cluster Proteins. pp. 405 – 423 in *Iron-Sulfur Clusters in Chemistry and Biology, 2nd ed.* (ed. Rouault, T.), DeGruyter, Berlin/Boston.
doi: http://10.1515/9783110479850-015

   Review article with P. Bartels as primary author.

2. Bartels, P.L.; Zhou, A.; Arnold, A.R.; Nuñez, N.N.; Crespilho, F.N.; David, S.S.; Barton, J.K. Electrochemistry of the [4Fe4S] Cluster in Base Excision Repair Proteins: Tuning the Redox Potential with DNA. *Langmuir*, **2017**, *33*, 2523 – 2530.
doi: http://10.1021/acs.langmuir.6b04581

   P. Bartels performed all electrochemical experiments with assistance from F. Crespilho and A. Zhou. A. Zhou, A. Arnold, and N. Nuñez prepared proteins for experiments.

3. Ha, Y.; Arnold, A.R.; Nuñez, N.N.; Bartels, P.L.; Zhou, A.; David, S.S.; Barton, J.K.; Hedman, B.; Hodgson, K.O.; Solomon, E.I. Sulfur K-Edge XAS Studies of the Effect of DNA Binding on the [Fe_4S_4] Site in EndoIII and MutY. *J. Am. Chem. Soc.* **2017**, *139*, 11434 – 11442.
doi: http://10.1021/jacs.7b03966

   Y. Ha performed all XAS experiments and DFT analysis. A. Arnold, N. Nuñez, and A. Zhou purified DNA and proteins. P. Bartels performed UV-visible, CD, and EPR spectroscopy.

4. McDonnell, K.J.; Chemler, J.A.; Bartels, P.L.; O'Brien, E.; Marvin, M.L.; Ortega, J.; Stern, R.H.; Raskin, L.; Li, G.; Sherman, D.H.; Barton, J.K.; Gruber, S.B. A Human MUTYH Variant Linking Colonic Polyposis to Redox Degradation of the $[4Fe4S]^{2+}$ Cluster. **2018**, *Submitted, Nature Chemistry*

   P. Bartels carried out all electrochemistry and UV-visible, circular dichroism, and EPR spectroscopy. E. O'Brien assisted with electrochemistry. K. McDonnell and J. Chemler identified the C306W mutation, purified proteins, and carried out characterization of the latter (activity assays, biolayer interferometry, and iron quantification).

5. Bartels, P.L.; Stodola, J.L.; Burgers, P.M.J.; Barton, J.K. A Redox Role for the [4Fe4S] Cluster of Yeast DNA Polymerase δ. *J. Am. Chem. Soc.*, **2017**, *139*, 18339 – 18348.
doi: http://10.1021/jacs.7b10284

   P. Bartels synthesized DNA, carried out all electrochemical and spectroscopic characterization, and ran and analyzed activity assays. J. Stodola and P. Burgers provided purified proteins for experiments.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

List of Tables and Schemes

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 7

Chapter 8

Chapter 1

Introduction: DNA Signaling by Iron-Sulfur Cluster Proteins

Adapted from: Bartels, P.L.; O'Brien, E.; Barton, J.K. DNA Signaling by Iron-Sulfur Cluster Proteins. pp. 405 – 423 in *Iron-Sulfur Clusters in Chemistry and Biology, 2nd ed.* (ed. Rouault, T.), DeGruyter, Berlin/Boston.

[4Fe4S] Clusters in DNA Processing Enzymes and DNA Charge Transport

The multi-metal [4Fe4S] clusters in mitochondrial and cytosolic proteins have long been known to serve critical functions ranging from electron transfer to catalysis (*1*). In an exciting turn, the last three decades have expanded the known cellular distribution of [4Fe4S] clusters to the nucleus, where they occur in DNA-processing enzymes throughout all domains of life (*2-12*). **Table 1.1** illustrates many of these new DNA processing enzymes containing [4Fe4S] clusters. As is evident from the table, [4Fe4S] clusters occur in proteins involved in all aspects of DNA processing. These enzymes are structurally and functionally diverse, acting in a range of pathways from base excision repair (BER) and nucleotide excision repair (NER) to DNA replication. Many of these proteins are known to be disease relevant, and so determining the function of the cluster has been a high priority (*13, 14*). In all of these proteins, the [4Fe4S] cluster is non-catalytic and largely redox-inert in solution, and only secondarily involved in maintaining structural integrity (*15, 16*). Despite some evidence to the contrary and the metabolic expense of inserting such a complex cofactor (*15 -17*), the cluster was typically assigned a structural role. This area remained highly contentious for years, but work in the Barton lab has demonstrated that DNA binding activates these proteins toward redox chemistry through an electrostatically-induced potential shift (*18*). The activation of these proteins for redox activity when bound to DNA enables them to take advantage of a fundamental property of DNA known as DNA charge transport (DNA CT).

DNA CT involves the transport of electrons or holes through the π-stacked base pairs; this process remains functional over long distances as long as stacking is not perturbed (i.e., by lesions, mismatches, or proteins that sharply kink the DNA) (*18*). Indeed, the sensitivity of DNA CT as a damage reporter led to much interest in possible biological applications of DNA CT, and

sparked initial interest in studying the mysterious [4Fe4S] proteins. The connection between [4Fe4S] proteins and DNA CT has proven to be extraordinarily fruitful from both perspectives. Over the course of numerous studies described later in this chapter, DNA CT has been found to help [4Fe4S] proteins search the genome for damage in a timely manner and to coordinate their activities by long-range redox signaling. The development of this model has since begun to provide insight into a range of poorly-understood mutations in the [4Fe4S] domain of repair proteins, and the importance of signaling among these proteins to disease-causing mutations underscores the need for a better understanding of the chemistry of [4Fe4S] clusters in DNA processing (*8,13,14*).

Until recently, the study of [4Fe4S] proteins was largely limited to bacterial repair proteins, with the behavior of their eukaryotic counterparts assumed to be similar. The work described in Chapters 2 and 3 describes some additional experiments to clarify details regarding bacterial repair proteins, while Chapter 4 discusses the first investigation of DNA CT in one of the eukaryotic homologues of bacterial BER proteins, human MUTYH. In addition, it has only recently become apparent that [4Fe4S] clusters are not uncommon in eukaryotic DNA replication proteins, and these include DNA primase and the B-family DNA polymerases. Chapters 5 through 7 describe multiple approaches to the study of one of these enzymes, the yeast lagging strand polymerase, DNA polymerase (Pol) δ. With the realization that [4Fe4S] proteins utilize DNA CT, it also became interesting to see if the single-stranded binding protein RPA could facilitate CT on ssDNA and serve to expand the possible signaling routes of redox-active proteins. Preliminary work with RPA thus constitutes Chapter 8. In this chapter, an outline of the studies that led to the current model for redox signaling in DNA repair is provided along with the

developments that led to an expansion and generalization of the model to include DNA replication proteins.

Table 1.1 DNA-processing [4Fe4S] proteins from the 3 domains of life.

Pathway	[4Fe4S] Proteins			Function
	Bacteria	Archaea	Eukarya	
Base Excision Repair (BER)	EndoIII, MutY	UDG, Mig	Ntg2, MUTYH, DME	DNA glycosylases that excise oxidized or misincorporated bases
Nucleotide Excision Repair (NER)	-	XPD	Rad3/XPD	Helicases that unwind DNA surrounding bulky lesions
DNA Replication	-	-	DNA Primase, DNA Polymerase (Pol) α, δ, ε, ζ	RNA priming of ssDNA and 5' - 3' synthesis of primed DNA
Replication Coupled Repair	AddAB, DinG	-	Dna2, FANCJ, ExoV	Varied; helicases and nucleases that expose ssDNA for homologous recombination or, in the case of DinG, unwind R-loops
Telomere Maintenance/Meiotic Crossover	-	-	Rtel1, Chl1	Helicases that unwind specialized DNA structures
Transcription	-	RNA Polymerase	Elp3	Template-directed RNA synthesis

DNA-Mediated Signaling in Base Excision Repair

The BER pathway involves the targeting and removal of damaged or misincorporated bases from DNA by one of several specialized DNA glycosylase enzymes (*2*). The resulting abasic (AP) site is then exposed by an endonuclease that nicks the phosphate backbone, allowing the short gap to be filled in by a DNA polymerase and then sealed by a DNA ligase. Within this pathway, [4Fe4S] clusters are present in several glycosylases of the helix-hairpin-helix family; the *Escherichia coli* enzymes endonuclease III (EndoIII) and MutY were the first well-characterized examples (*3-6*). EndoIII is a bifunctional glycosylase responsible for excising oxidized pyrimidines and nicking the DNA backbone at the site of damage, while MutY is a monofunctional glycosylase that removes adenine mispaired with 8-oxoguanine. Homologues of both proteins are present in nearly all organisms, from bacteria to man, and the [4Fe4S] domain is conserved throughout (*4, 6*). Early studies on EndoIII by Cunningham showed the cluster to be largely insensitive to both oxidation and reduction, leading to the eventual assignment of a structural role (*3, 15*). In the case of MutY, however, the cluster was found to be unnecessary for structural integrity, and a possible substrate-sensing role for the cluster was proposed instead (*5*). However, this mechanism could not explain the role of the [4Fe4S] cluster in proteins other than MutY, and recent work demonstrating full activity in the MutY homologues of anaerobic organisms that lack a cluster entirely provides a further argument against this possibility (*16*). Amidst this perplexing situation, support for a functional role for the cluster arose, due to strict conservation of [4Fe4S] clusters in these proteins despite the metabolic expense associated with cluster production and loading into target apoproteins (*17*).

Unexpectedly, the key to understanding the role of the cluster turned out to be a fundamental property of the DNA substrate itself: ground state B-form DNA can conduct charge

due to the π-stacked arrangement of the aromatic base pairs, which have a similar spacing and arrangement to that of conductive graphite sheets (*18*). This remarkable property was demonstrated in the ground state through electrochemical experiments where DNA containing a covalent alkane-thiol linker at one end was tethered to a gold electrode, and a redox-active intercalator appended to the opposite end of the duplex served as an electron donor/acceptor upon the application of a potential (Figure 1.1). Using cyclic voltammetry (CV) and square wave voltammetry (SQWV), rapid, long-range charge transport (CT) has been observed over distances up to 34nm (100bp of duplex DNA) with rates comparable to those measured in a 17-mer (*19*). CT is efficient even with multiple breaks present in the phosphate backbone, but just a slight perturbation to base stacking, such as the presence of a CA mismatch, has been shown to sharply attenuate charge transport yields (Figure 1.1). Well-stacked base pairs are thus a requirement for DNA CT. The biological accessibility of DNA on this platform has been demonstrated by experiments measuring restriction enzyme activity electrochemically, so that proteins are able to recognize their cognate sequence and carry out reactions on the DNA duplex on the electrode. Overall, CT renders DNA an effective redox sensor of DNA integrity in cells, and this concept, combined with the propensity of biological systems to use all available resources at their disposal, led to a series of experiments designed to test the redox activity of DNA-bound [4Fe4S] proteins.

To determine if otherwise redox-inert [4Fe4S] proteins could become activated to carry out DNA CT upon binding the DNA polyanion, [4Fe4S] proteins were added to DNA-modified gold electrodes, with the DNA-bound [4Fe4S] enzyme taking the place of a redox probe (Figure 1.2). In a revealing study, EndoIII and MutY from *E. coli* and Uracil DNA glycosylase (UDG) from *Archeoglobus fulgidus* were each incubated in buffered solution at physiological pH on

DNA-modified electrodes and scanned by CV (*20*). Remarkably, a reversible redox signal was observed for all three proteins and at quite similar potentials. DNA CT from the protein to electrode surface was severely attenuated both in the absence of DNA and when the DNA duplex contained an intervening abasic site proximal to the electrode surface, verifying that the signal was DNA-mediated. Midpoint potentials ranging from 75 mV versus NHE (EndoIII) to 95 mV (MutY and UDG) placed all three enzymes at the lower end of the 100-300 mV range reported for high potential iron proteins (HiPIPs), and well above the -200 to -600 mV range expected for ferredoxins (*21*). These electrochemical results, combined with data from EPR spectroscopy of native and chemically oxidized DNA-bound proteins, led to an assignment of the observed signal to the $[4Fe4S]^{3+/2+}$ couple utilized by HiPIPs rather than the $[4Fe4S]^{2+/1+}$ couple favored by ferredoxins.

To understand how this redox activation occurred, it was necessary to directly compare the redox potential of the [4Fe4S] cluster in the presence and absence of DNA. The DNA-dissociated [4Fe4S] enzymes were previously shown to be resistant to a change in cluster redox state even in the presence of powerful chemical oxidants *(3)*. This redox insensitivity suggested that the DNA-dissociated proteins have high reduction potentials outside the physiological range, requiring an electrode with a wider available potential window than gold. The 2V scanning window of highly-oriented pyrolytic graphite (HOPG) satisfied this requirement, and it could furthermore be modified with DNA by appending a pyrene linker to the end of the duplex to form a non-covalent bond with the surface (*22*). On the bare electrode, CV and SQWV revealed an irreversible signal for EndoIII with an oxidative peak centered at 250 mV versus NHE, just outside the physiologically relevant potential range (Figure 1.3). An irreversible $[4Fe4S]^{2+/1+}$ reduction was also observed around -300 mV, supporting the initial assignment of the DNA-

bound signal to the $[4Fe4S]^{3+/2+}$ couple. In the presence of DNA, however, the $[4Fe4S]^{3+/2+}$ couple underwent a shift in potential of ≥ -200 mV and became much larger and more reversible; signal loss in the presence of an abasic site confirmed that this process was DNA-mediated (Figure 1.3). The potential shift, in turn, corresponded thermodynamically to an increase in DNA binding affinity of 3 orders of magnitude for the oxidized $[4Fe4S]^{3+}$ form of the protein, relative to the reduced $[4Fe4S]^{2+}$ form. The lack of significant conformational differences between the DNA-dissociated and DNA- bound structures of EndoIII and MutY (*16, 23*) suggested that this potential shift is due to the electrostatic effects resulting from binding to the polyanionic backbone of DNA; the DNA polyanion tunes the potential of the bound [4Fe4S] cluster. More recent results examining EndoIII and MutY along with electrostatic mutants in electrochemistry experiments on graphite support this notion, and are the topic of Chapters 2 and 3 (*24*). In addition, limitations to the HOPG system and their solutions are discussed in detail in Chapter 2.

Overall, these electrochemical experiments revealed several critical details about the redox properties of [4Fe4S] proteins in BER. First, DNA binding activated the proteins for redox activity under physiological conditions by negatively shifting the potential, and this negative shift meant that the oxidized form of the protein would necessarily bind DNA with a much greater affinity than the native, reduced form. Second, EndoIII, MutY, and UDG all displayed DNA-mediated redox signals centered around 85 mV vs. NHE, and thus similar DNA-bound redox potentials for the cluster. With no other obvious redox partners, it was reasonable to consider whether these DNA-bound enzymes might be using DNA CT to signal to each other. CT between even distant DNA-bound proteins is certainly temporally feasible, as DNA charge equilibration takes place on the nanosecond timescale (*25*) while proteins diffuse along DNA on the microsecond to millisecond timescale.

One clear partner for redox chemistry was the guanine radical. Under conditions of oxidative stress, guanine radicals are generated in the DNA duplex, and indeed MutY recognizes and repairs 8-oxoguanine/A mismatches, generated after the formation of oxidized guanine radicals in DNA. Guanine damage generated by long-range oxidation of guanine through DNA-mediated CT has been measured biochemically and occurs over long molecular distances (*26,27*)**.** Monitoring guanine radical formation spectroscopically was used to determine rates of DNA CT; DNA CT occurs on the nanosecond timescale and is rate-limited by the base pair motions (*25*). In fact, EPR and transient absorption spectroscopies were used to characterize DNA CT between the guanine radical and MutY, resulting in formation of the oxidized $[4Fe4S]^{3+}$ cluster (*28*). We have also demonstrated that guanine radicals can transcriptionally activate SoxR from a distance (*29*). SoxR is an iron-sulfur protein that acts as a sensor of oxidative stress in bacteria and activates a series of genes to respond to the stress **(**Figure 1.4**)**. Thus, under conditions of oxidative stress, it is likely that the guanine radical can be a source for oxidation of the BER enzymes with [4Fe4S] clusters by DNA CT and a means potentially to signal the need to activate necessary repair.

The fact that cellular DNA is not linear but wrapped around histones in chromatin brought up an important concern about the feasibility of long-range signaling *in vivo*: can DNA CT still occur in DNA wrapped around histones in a nucleosome core particle? This issue was addressed by an experiment that isolated DNA cleavage at sites of guanine oxidation in nucleosome-wrapped DNA using a rhodium photooxidant covalently tethered to one end of the DNA (*30*). The occurrence of damage, even at sites distant from the photooxidant, demonstrated that curvature of the DNA is unimportant as long as local π-stacking is unperturbed. The wrapping of DNA around histones produces very gradual curvature, not the kinking of DNA,

which is known to interfere with DNA CT. Local π-stacking perturbations can arise from the binding of certain proteins, as has been shown with the transcription factor TATA binding protein, which kinks the DNA duplex at a sharp angle and effectively shuts off CT (*31*). This is not, however, the typical binding mode of DNA-binding proteins; helix-turn-helix proteins do not interfere with DNA CT and thus many proteins that coat the DNA in the cell are not expected to affect long range signaling through DNA CT.

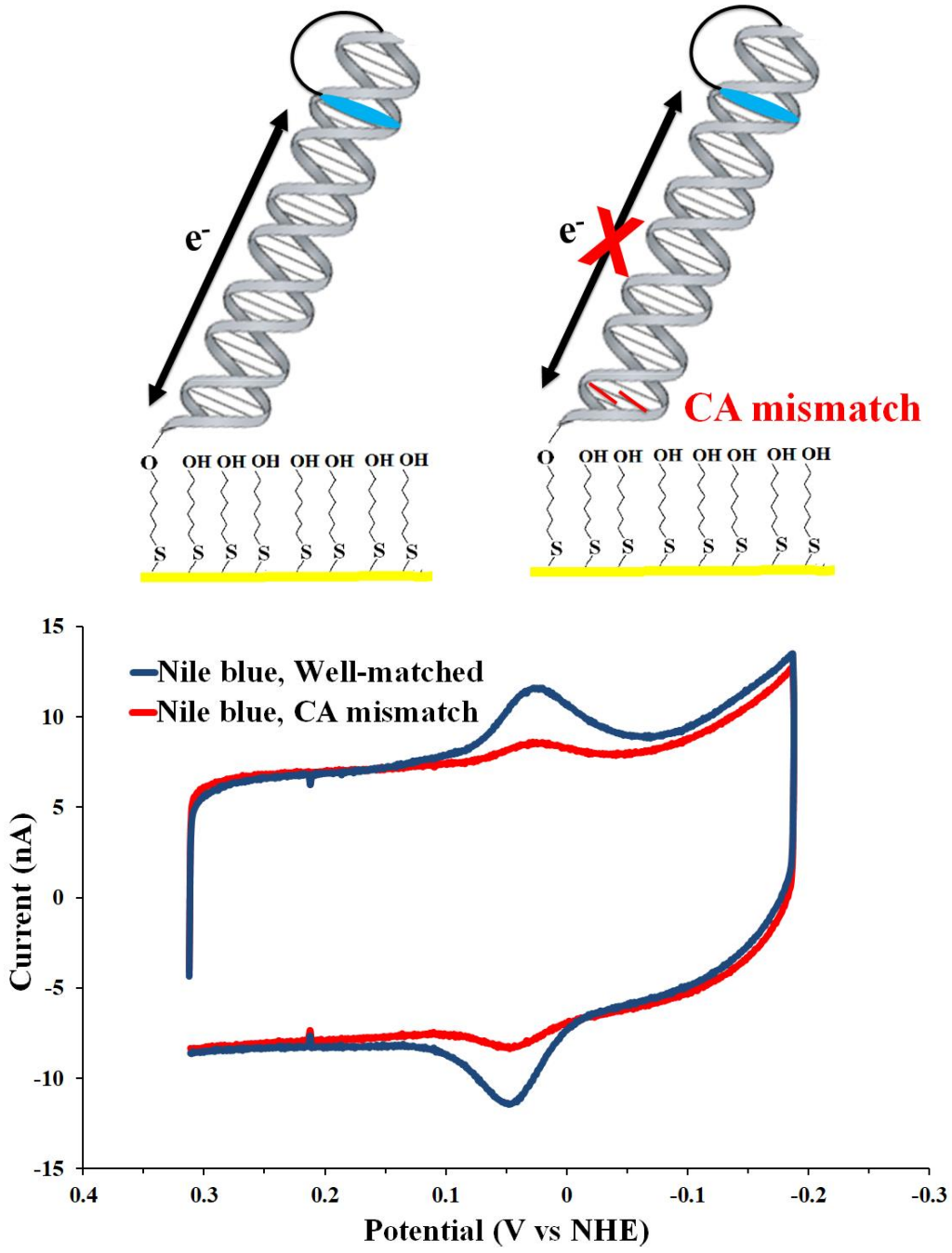


Figure 1.1 Redox probe electrochemistry on DNA-modified gold electrodes. As illustrated by Nile blue, DNA-intercalating redox probes can participate in DNA-mediated charge transport on this platform to produce a reversible signal by cyclic voltammetry (CV; blue). The presence of a single modest disruption, such as a CA mismatch, markedly attenuates CT (red).

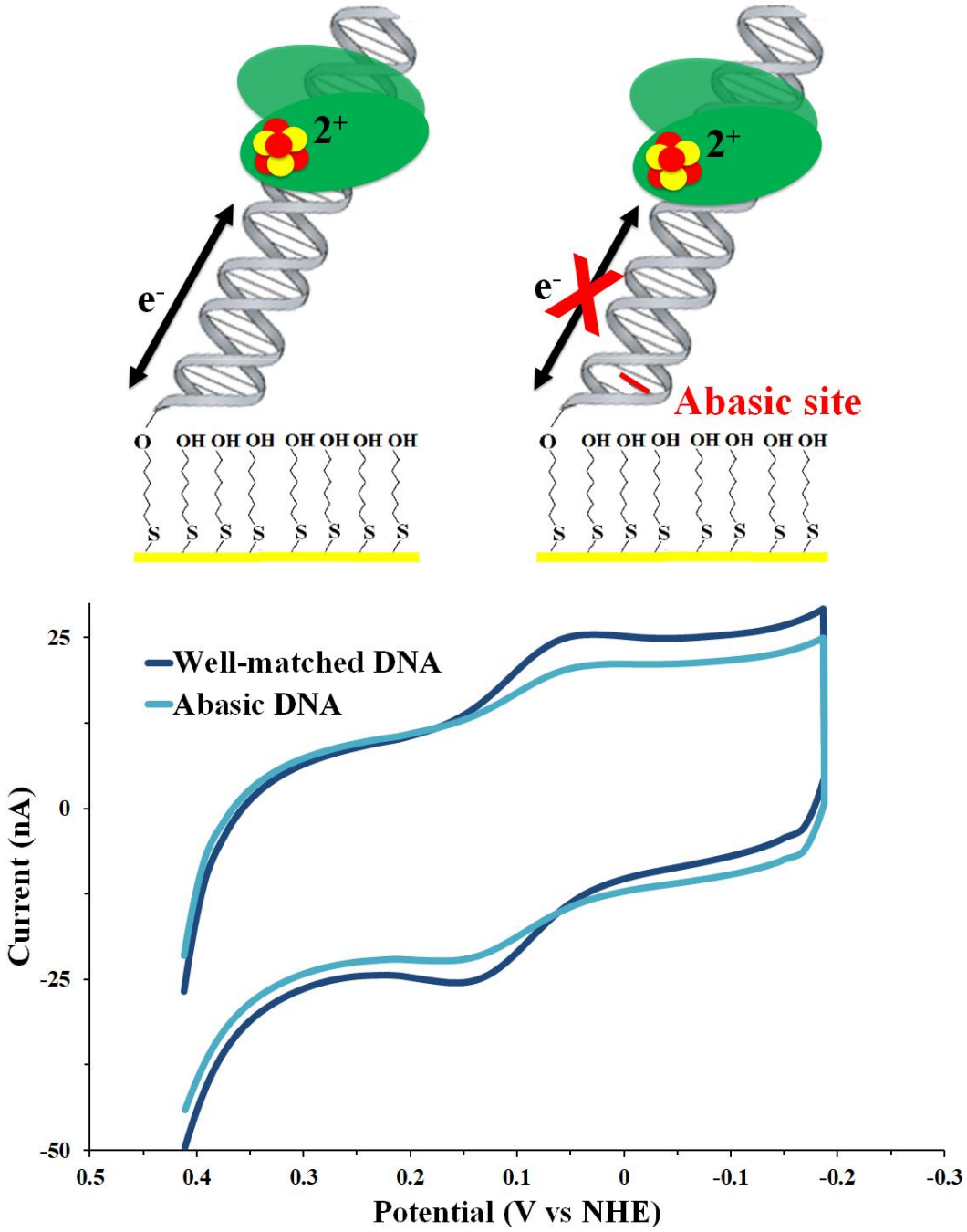


Figure 1.2 Protein electrochemistry on DNA-modified gold electrodes. When a [4Fe4S] cluster protein binds to DNA on a self-assembled monolayer, electron transfer between the cluster and the electrode is highly efficient (left cartoon, dark blue CV). If an abasic site or mismatched base pair is incorporated into the DNA sequence, CT is disrupted and the signal is effectively shut off (right cartoon, light blue CV).

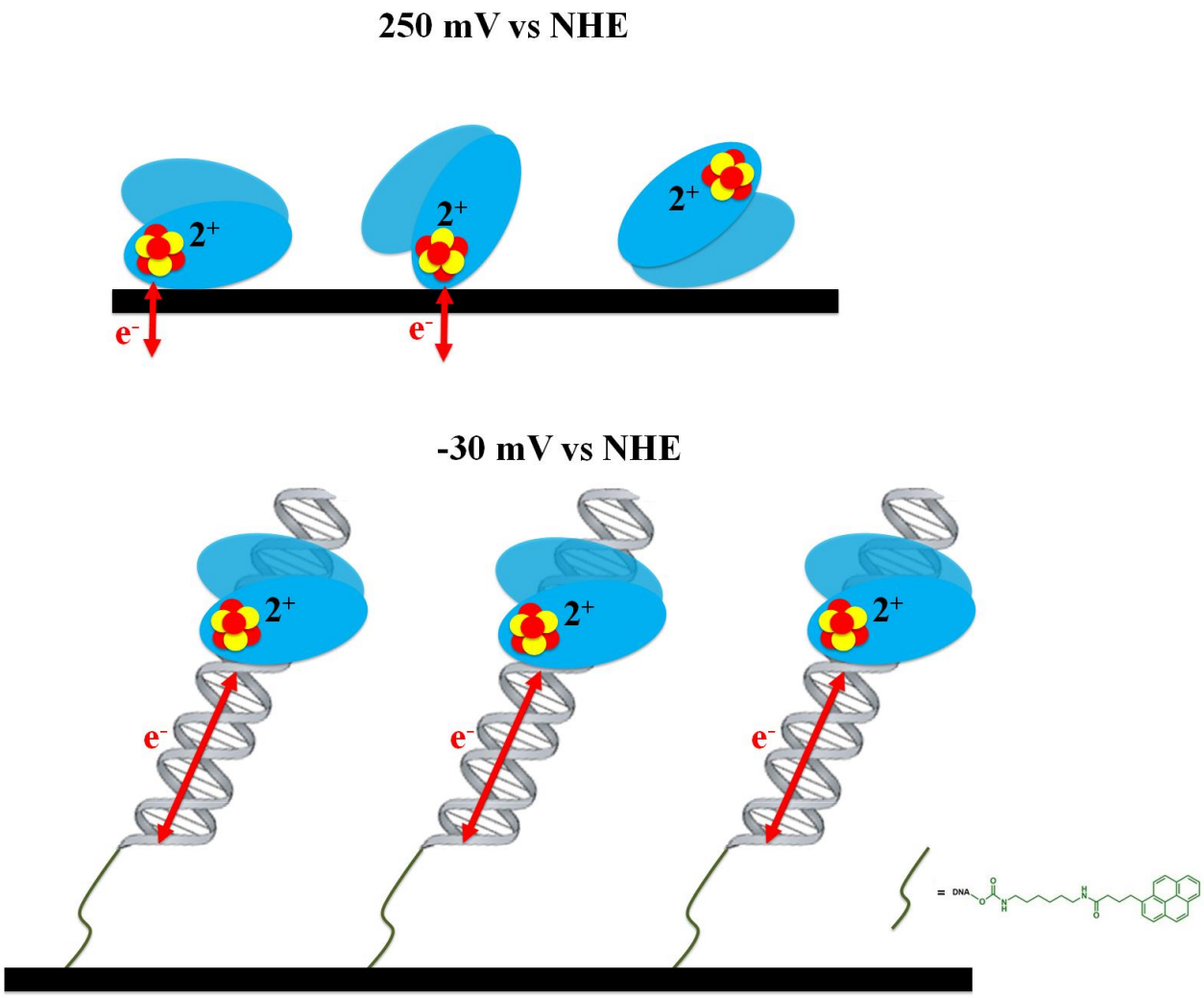


Figure 1.3 EndoIII electrochemistry on a highly oriented pyrolytic graphite (HOPG) electrode in the presence and absence of DNA. When EndoIII was incubated in solution on HOPG, an irreversible signal with a potential of 250 mV vs NHE by square wave voltammetry was observed (top). Importantly, the $[4Fe4S]^{2+/1+}$ couple could also be observed around -300 mV vs NHE, in agreement with the assignment of the high potential couple to the $[4Fe4S]^{3+/2+}$. In contrast, when a film of pyrene-modified DNA was present, the EndoIII signal was reversible and the potential shifted over 200 mV to -30 mV vs NHE (bottom). Due to the lack of significant conformational changes upon DNA binding, this effect was attributed primarily to the electrostatic effects associated with binding to the polyanionic DNA backbone.

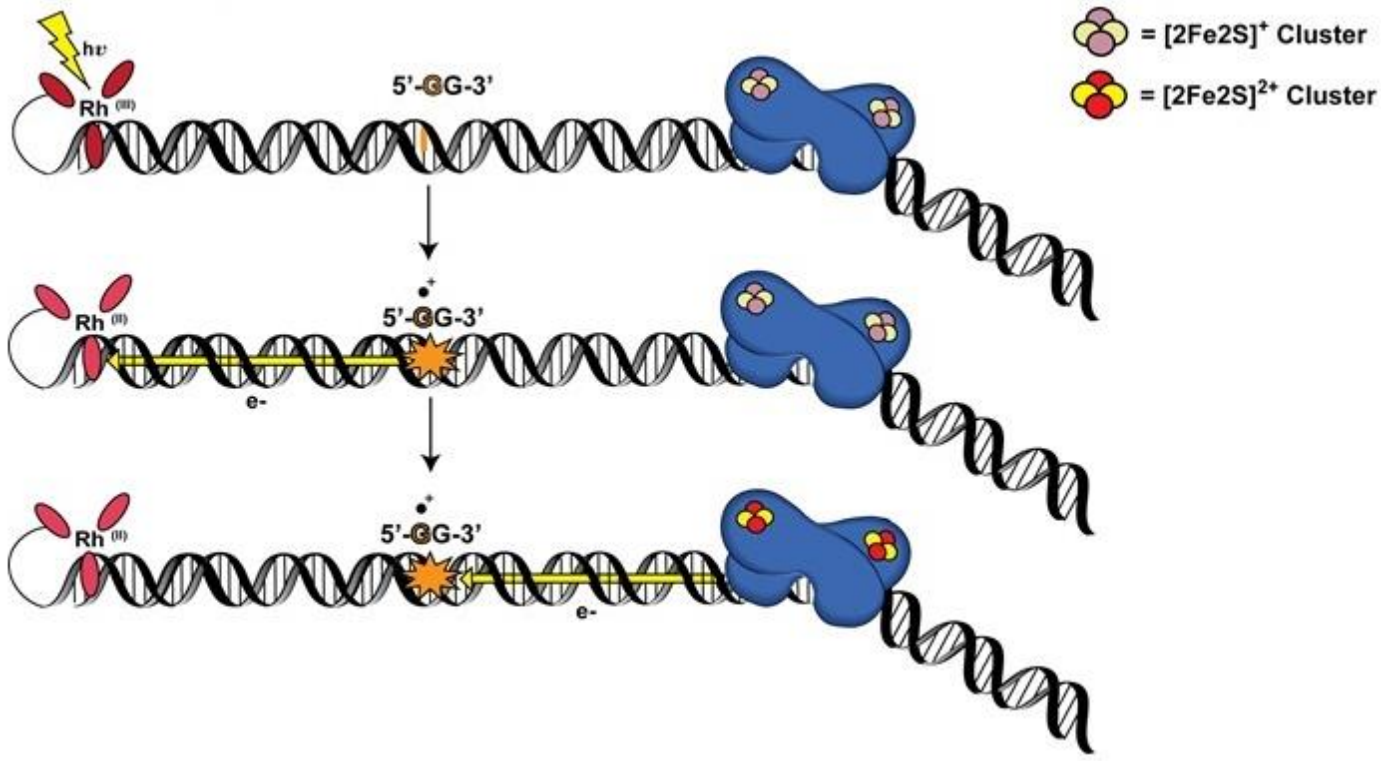


Figure 1.4 Transcription factor SoxR, activated in response to oxidative stress in the cell, contains a [2Fe2S] cluster, which is oxidized from the $[2Fe2S]^{+}$ form to the $[2Fe2S]^{2+}$ form when turned on for activity. Using a DNA-intercalating Rh (III) photooxidant, a guanine radical is generated at the 5′- position in a 5′-GG-3′ doublet. The guanine damage can be repaired at a distance by bound SoxR, though DNA CT. The guanine damage in turn oxidizes SoxR to the $[2Fe2S]^{2+}$ form and turns on the oxidative stress response.

Assessing redox signaling by [4Fe4S] proteins *in vitro* and *in vivo*

In addition to exploring chemically whether the [4Fe4S] cluster of the repair proteins could be oxidized in a DNA-mediated reaction, we became interested in visualizing the process. To do so, an atomic force microscopy (AFM) assay was developed to assess the distribution of [4Fe4S] proteins on DNA (Figure 1.5, ref. *32*). Specifically, this assay involved the addition of WT EndoIII to a mixture consisting of a 3.8 kb DNA substrate, either completely well-matched (WM) or containing a single CA mismatch (MM), and two smaller (2.2 and 1.6 kb) strands of WM DNA (from which the larger strand was composed). This solution was dried on a mica surface and imaged, with DNA-bound proteins distinguished by their greater height relative to DNA and free proteins, thereby providing a visual snapshot of the equilibrium binding distributions of proteins on the DNA. With respect to the distribution, it was predicted that redox signaling would cause tightly bound oxidized proteins to be trapped in the vicinity of a CT-attenuating mismatch, leading to an increase in binding density on mismatched DNA over well-matched DNA. The notion is that on the well-matched strand, there is extensive DNA-mediated CT between the DNA-bound proteins, facilitating the dissociation (with reduction) and reassociation (with oxidation) of proteins onto different strands. With a mismatch on the strand, DNA CT is inhibited, and thus little dissociation and redistribution occurs, leading ultimately to proteins being bound preferentially on the mismatched versus fully matched duplexes. Importantly, a CA mismatch is not a substrate for EndoIII, so there is no intrinsic reason for the proteins to localize to this strand. Indeed, we found that EndoIII binding density ratios for WM versus MM DNA (proteins bound per kb on long DNA/proteins bound per kb on short DNA) averaged to 1.6 for mismatched long DNA, indicating a preference for mismatched strands. In

control experiments where the long and short strands were both fully matched, the binding densities were always essentially the same.

Although EndoIII as purified is largely in the $[4Fe4S]^{2+}$ oxidation state, enough $[4Fe4S]^{3+}$ EndoIII must be present in a given sample to allow redistribution in the AFM assay. However, EndoIII and other BER proteins generally operate under conditions of oxidative stress in which one would expect a higher proportion of the proteins to be oxidized. To test conditions in which more oxidized $[4Fe4S]^{3+}$ protein is initially present, EndoIII/DNA mixtures were incubated with hydrogen peroxide prior to AFM imaging (*32*). Consistent with oxidative stress activating this process, oxidation resulted in an increase in the redistribution, with the binding density ratio on mismatched DNA increasing from 1.6 to 2.4. The protein is able to "find" the strand containing a single mismatch on a 3.8 kilobase duplex.

EndoIII mutants that were defective in carrying out DNA CT had been prepared and characterized, and it was of interest to see how these mutations would affect redistribution. Tyrosine and tryptophan residues are well known to facilitate electron transfer within proteins (*33*), and it was reasonable to consider that they might be involved in relaying electrons between DNA and the cluster in EndoIII. With this aim, a range of mutants were prepared and characterized in activity assays, and their CT properties were then investigated in electrochemical experiments and the AFM assay (*34*). Independent mutation of several aromatic residues in EndoIII, including F30, Y55, Y70, and Y82, resulted in proteins with full catalytic activity and an identical midpoint potential but differing extents of CT deficiency relative to WT, as measured by the current signal height per cluster in a cyclic voltammogram on a DNA electrode. Interestingly, we could correlate directly the efficiency of DNA CT with redistribution in the AFM assay; those proteins which showed poor electrochemical signals on DNA

electrodes, reflecting poor DNA CT, also showed low binding density ratios in the AFM assay, while those with high DNA CT efficiency showed high ratios for redistribution onto the mismatched strand. Thus proteins with efficient DNA CT could more effectively find the mismatched strand.

But does this signaling occur within the cell? To see if these redox-based exchanges between [4Fe4S] proteins occur in the cell, we took advantage of a genetic assay, our "helper function" assay designed to assess the effect of CT signaling on MutY activity (Figure 1.6, ref. *32*). This assay used *E. coli* strain CC104, which has a cytosine swapped for an adenine in the *lacZ* Glu-461 codon, preventing β-galactosidase activity and inhibiting growth in media with lactose as the primary carbon source. Because MutY specifically removes adenine mispaired with oxo-guanine, lowered MutY activity in the CC104 strain results in CG → AT transversions that restore β-galactosidase activity and growth on lactose (lac^+). Numbers obtained could range from ~20 revertants per 10^9 plated cells in the background to over 300 revertants per 10^9 cells when MutY was knocked out. To see if cooperation or signaling with other [4Fe4S] proteins might affect MutY activity, the EndoIII gene (*nth*) was knocked out in the CC104 strain; critically, EndoIII does not resolve A-oxoG mispairs, so it cannot function redundantly with MutY. Despite this, nth^- cells showed an average of 54 lac^+ revertants, more than a two-fold increase over background, indicating that EndoIII "helps" MutY find its targets. To verify that this effect was due to long range CT, CC104 nth^- strains were complemented with plasmids encoding either CT-deficient EndoIII Y82A or catalytically inactive but CT-proficient D138A. Y82A was unable to restore background transversion rates, which were indistinguishable from uncomplemented nth^-; in contrast, D138A, despite its inability to carry out the glycosylase reaction, was able to help MutY and lower transversion rates to background levels. These genetic

results provided a direct link between DNA CT and the observed ability of EndoIII to assist MutY in finding its targets.

These assays thus laid a foundation for considering how long range signaling through DNA CT might indeed function for communication and cooperation among [4Fe4S] cluster repair proteins within the cell. In our model, DNA repair proteins with [4Fe4S] clusters use long-range redox signaling to communicate on DNA as a first step in locating their targets (*32*). As illustrated in Figure 1.7, a redox-inert repair protein in the native $[4Fe4S]^{2+}$ oxidation state binds to DNA and becomes activated toward oxidation. If another distally bound protein is in the oxidized $[4Fe4S]^{3+}$ state, the newly-bound protein can reduce it at a distance via DNA CT; upon reduction, the binding affinity of the distal protein is lowered and the protein is free to diffuse to another region of the genome. When the intervening DNA between the two proteins is undamaged, this self-exchange reaction proceeds efficiently. However, if a mismatch or lesion is present between the proteins, CT is attenuated and the proteins can no longer communicate; both proteins then remain oxidized and bound to the DNA in the vicinity of the lesion, significantly reducing the range over which diffusion must occur and allowing repair of the entire genome on a biologically relevant time scale.

In addition to the novelty of the CT-based damage search, this model was particularly relevant because it presented a solution to the significant problem of how repair proteins manage to locate substrates on a time scale feasible for biological processes. Earlier models generally invoked some combination of one-dimensional and three-dimensional diffusion along DNA to explain this problem (*35*), but these mechanisms alone have been estimated to take far too long (over twice the cell's doubling time) for low-copy number proteins like MutY to search the ~4.5 Mb *E. coli* genome (*32*). If oxidative lesions were rare, this might not be a problem, but roughly

1000 such lesions occur per doubling time (*32*), and the situation is no more favorable in other organisms (*5*). However, when DNA-mediated CT scanning of the genome is factored into this process, the search time is significantly reduced even with short CT distances of ~200 bp, and substantially more so with longer CT distances (*32*). Critically, this mechanism does not preclude diffusional search methods, but instead simply provides a way for proteins to reach the vicinity of a lesion significantly more rapidly, and independent of the many other proteins associated with the DNA; once in the vicinity, diffusion over relatively short (~200 kb) distances would be used to locate and recognize the damaged base.

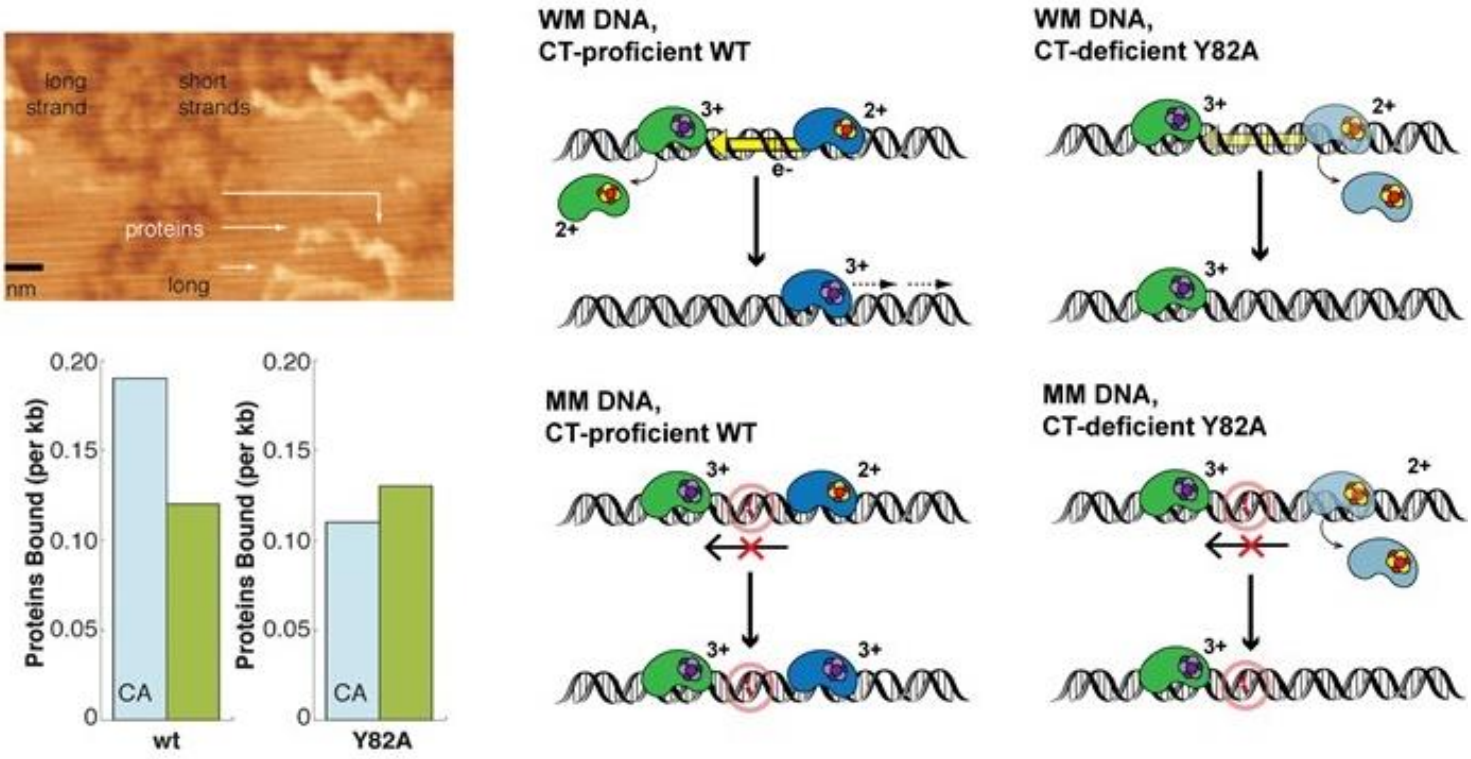


Figure 1.5 Atomic force microscopy (AFM) assay to assess protein redistribution by CT. DNA-bound proteins can be visually distinguished by their relatively great height on the surface, as seen in a sample image at top left. If CT signaling is occurring (middle), WT proteins like EndoIII preferentially bind to DNA containing a CA mismatch, leading to significantly more proteins bound to mismatched strands, as seen in the plot at bottom left. Conversely, when the assay is carried out with a CT-deficient mutant such as EndoIII Y82A (far right), no redistribution is observed (bottom left). Adapted from Reference 32.

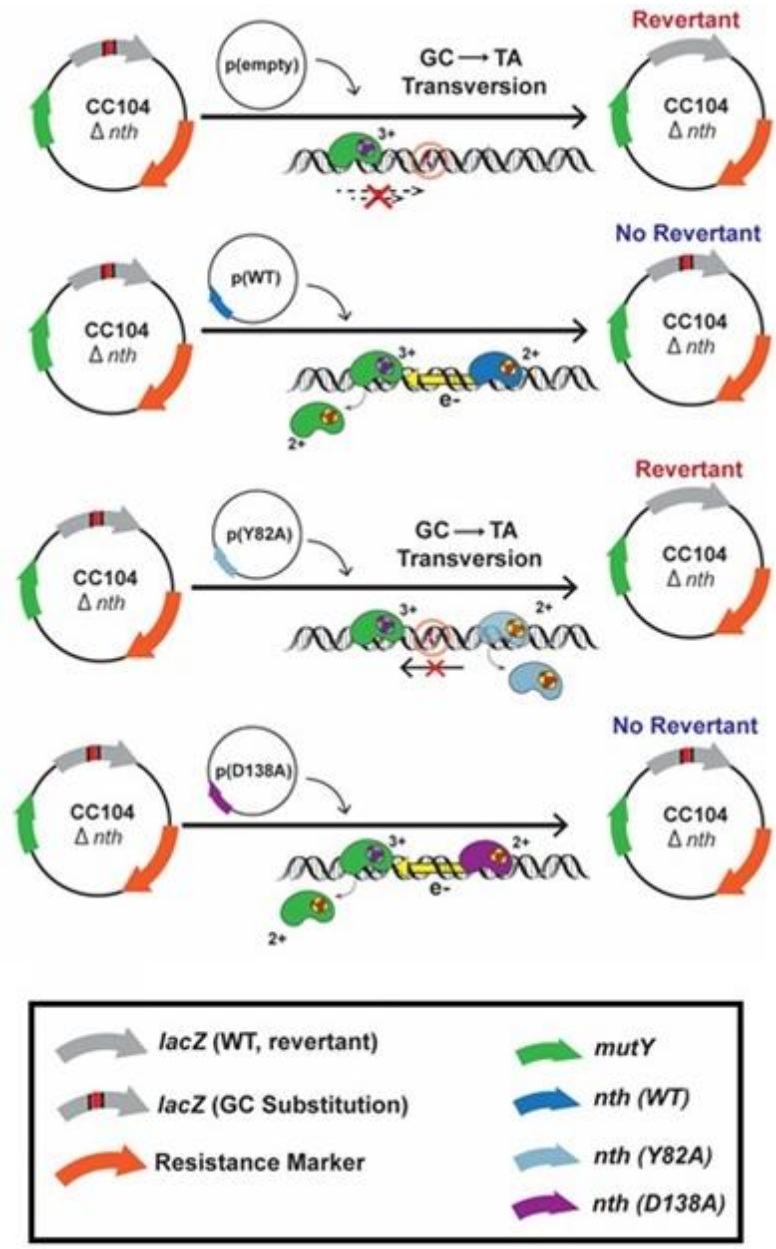


Figure 1.6 *E. coli* helper function assay for MutY activity. The CC104 strain used in this assay contains a GC substitution in the *lac*Z gene, rendering the cells unable to metabolize lactose. Oxidative stress generates 8-oxoguanine, which is readily mispaired with adenine during replication; repair by enzymes that target 8-oxo G ultimately cause a GC → TA transversion, reverting the *lac*Z gene back to WT. MutY, however, excises adenine mispaired with 8-oxoG, so its activity prevents reversions. Remarkably, when EndoIII (nth) is knocked out and an empty plasmid is added in (top), revertants are observed, indicating an impairment in MutY activity. WT EndoIII restores the efficiency of MutY (top middle), while the CT-deficient mutant Y82A is unable to rescue MutY activity (bottom middle). In contrast, the CT-proficient, but catalytically defective, mutant EndoIII D138A has the same restorative effect on MutY activity as WT (bottom), confirming that DNA-mediated redox signaling is the primary factor responsible for this result.

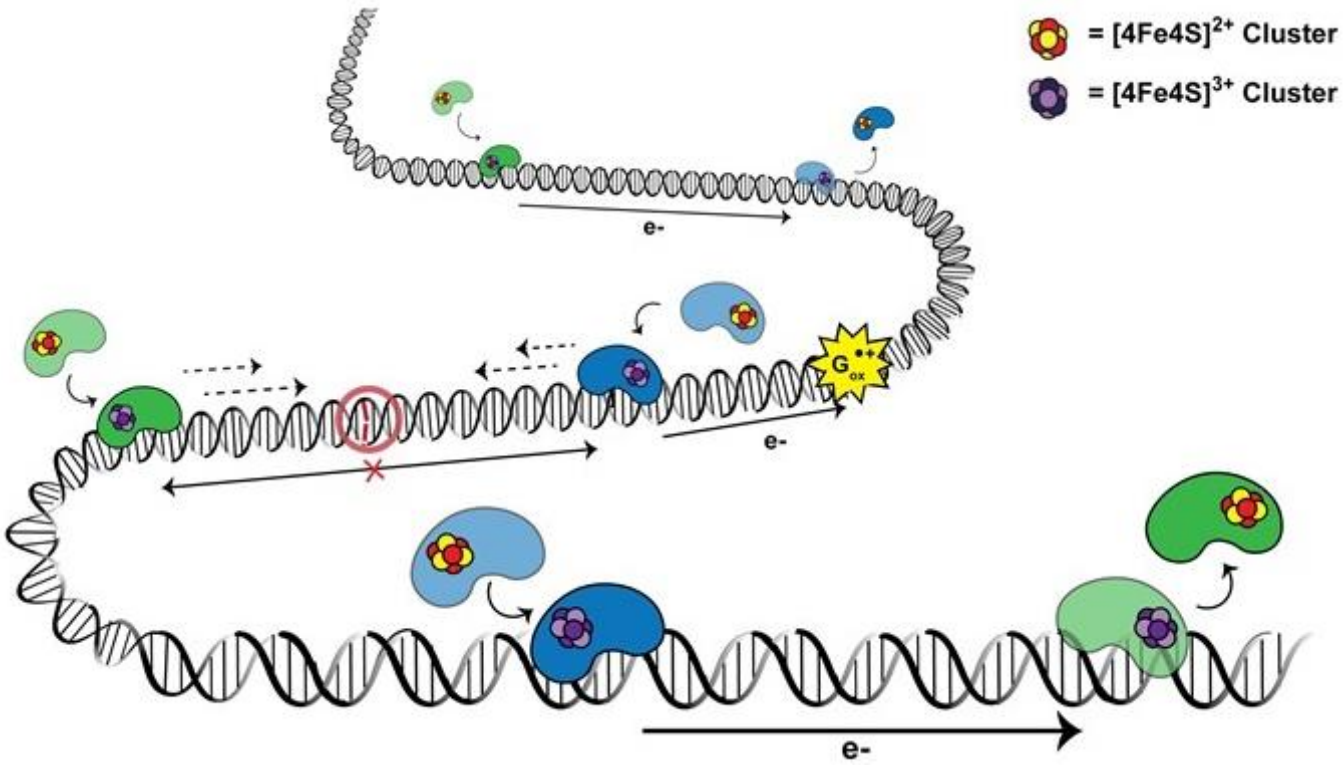


Figure 1.7 A model for DNA-mediated redox signaling between repair proteins. Enzymes with the cluster in the native $[4Fe4S]^{2+}$ first bind DNA, causing the cluster to become activated toward oxidation. Oxidative stress initiates the damage search when highly reactive species such as the guanine radical cation are formed; these can oxidize DNA-bound proteins in their vicinity. Oxidation of the cluster to the $[4Fe4S]^{3+}$ form leads to a 1000-fold increase in DNA binding affinity, so oxidized proteins remain bound and diffuse along the DNA. When another [4Fe4S] protein binds at a distant site, it can send an electron through the DNA base stack to reduce the oxidized protein. At this point, the reduced protein binds less tightly to DNA and diffuses away, while the newly oxidized protein continues the damage search. This process of redox exchange continues until a segment of DNA containing a lesion is approached. Since even subtle lesions can disrupt base stacking, CT is attenuated and any nearby oxidized proteins remain bound. Thus, DNA CT allows repair proteins to scan large sections of the genome and focus their time on areas containing damage.

DNA CT in other repair pathways

In addition to BER, [4Fe4S] proteins have also been found in NER pathways in archaea and, by homology, eukaryotes (*8*). NER involves the removal of bulky lesions such as thymine dimers by exposing the damage through helicase-mediated unwinding of the surrounding ~25 nt of DNA, after which an endonuclease excises the segment and the resultant gap is filled in by a DNA polymerase (*36*). In archaea and eukaryotes, DNA unwinding is dependent upon the ATP-stimulated activity of the [4Fe4S] helicase enzyme XPD. Although XPD is itself part of the transcription factor IIH (TFIIH) complex, it has nonetheless been isolated from the extremophile *Sulfolobus acidocaldarius* and structurally characterized. Upon incubation on a gold electrode modified with an appropriate DNA substrate (20-mer duplex with a 9-mer ssDNA overhang), a reversible, DNA-mediated signal centered at 80 mV vs. NHE is observed (*37*). This signal is comparable in both shape and potential to EndoIII and MutY, in line with predictions for the CT scanning model. In the case of XPD, however, the addition of ATP to stimulate DNA unwinding strongly enhances the current, while the non-hydrolysable analogue ATP γ-S does not. This signal enhancement indicates improved coupling of the cluster to the DNA during activity, a function which could be very important in coordinating with other proteins during NER, effectively "signaling" that the repair protein is functioning.

In humans, mutations in XPD are associated with several diseases, including xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy (*8*). The archaeal versions of two of these mutants, G34R and L325V, were characterized electrochemically; both were CT-deficient (*37*,*38*). To see if DNA-mediated signaling could occur between disparate pathways and proteins, *S. acidocaldarius* XPD and *E. coli* EndoIII were incubated together in the presence of DNA and imaged by AFM (*38*). As with experiments involving only EndoIII, the presence of a

CA mismatch resulted in an elevated DNA-binding density ratio; this effect was lost if WT EndoIII or XPD were incubated with a CT-deficient signaling partner, namely XPD L325V or EndoIII Y82A, respectively. Thus, XPD was able to help EndoIII localize to damaged DNA, but only if both proteins were CT-proficient. This experiment established two important general properties of CT between [4Fe4S] proteins: first, that long-range signaling can occur between proteins in distinct pathways, and second, that the proteins do not even have to be from the same organism in order to communicate in this manner. What is critical is that they both bind DNA, have similar DNA-bound redox potentials, and are well coupled into the DNA helix to carry out DNA CT.

E. coli DinG is a superfamily 2 helicase with homology to XPD that also contains a [4Fe4S] cluster (*7*), although DinG is primarily tasked with R-loop maturation rather than NER. R-loop maturation involves the helicase-mediated unwinding of RNA-DNA hybrids that result from collisions between transcription and replication machinery (*39*). We found that DinG behaved similarly to XPD on DNA-modified Au electrodes, displaying a virtually identical midpoint potential, and the increase in current upon the addition of ATP was even more dramatic than for XPD (*40*). Likewise, DinG showed a redistribution onto mismatched DNA in the AFM assay, both alone and in a mixture with WT EndoIII but not when combined with CT-deficient EndoIII Y82A. These assays supported the model developed for CT signaling in repair.

The real value in probing signaling by DinG, as an *E. coli* protein, was the ability to examine *in vivo* signaling in a bacterial system both with EndoIII and MutY. As an initial effort in elucidating signaling between pathways, the *lac*$^+$ helper function assay discussed above was employed to see if CT-active DinG could stimulate MutY activity in the same way as EndoIII. Remarkably, a DinG knockout did cause an increase in *lac*$^+$ reversions, despite the fact that

DinG and MutY are active in distinct repair pathways. This result was in agreement with the *in vitro* AFM studies showing communication between XPD and EndoIII but went further by demonstrating such communication *in vivo* within the same organism.

Importantly, we could ask also if EndoIII signaling was necessary for DinG activity. Here we used the InvA *E. coli* strain to test the effect of EndoIII CT on DinG activity (Figure 1.8, *41*). This assay relied upon the reversal of the frequently transcribed *rrnA* operon in *E. coli* to increase the frequency of replication/transcription collisions, causing stalled forks and generating R-loops; DinG correction of these R-loops was essential for cell survival. Indeed, cell growth was abolished when EndoIII was knocked out, and complementation with either WT or the enzymatically inactive, CT-proficient mutant EndoIII D138A restored survival. In contrast, complementation with the CT-deficient but enzymatically active EndoIII Y82A did not restore survival. Taken together, the MutY helper function and InvA strain survival assays showed that long-range signaling by DNA CT is critical to [4Fe4S] enzymes in disparate pathways: DinG signals BER proteins in their search for damage, and the BER proteins, in turn, can facilitate R-loop resolution by DinG.

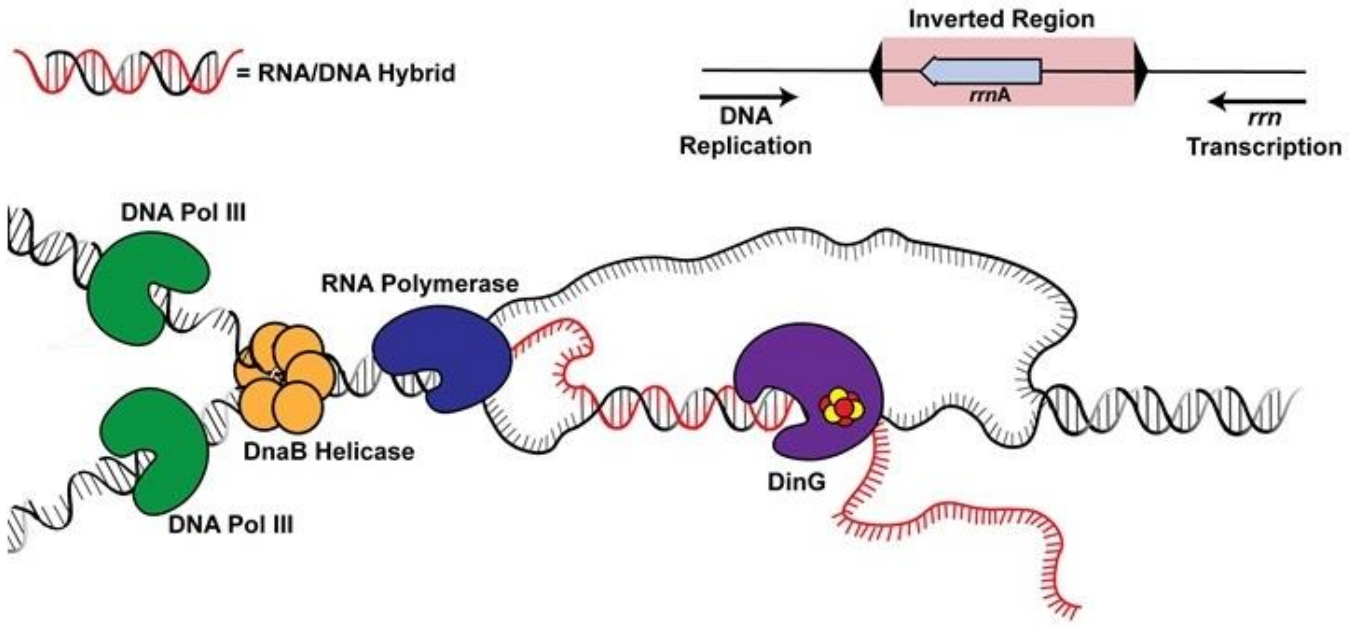


Figure 1.8 *E. coli* inverted A assay for DinG activity. In this assay, the highly transcribed *rrn*A operon has been inverted to increase the rate of replication/transcription collisions, generating R-loops, which the [4Fe4S] helicase DinG unwinds to resolve the stalled fork. Surprisingly, knocking out EndoIII prevented DinG from resolving R-loops, abolishing cell growth. As in the helper function assay (**Figure 1.6**), both WT and the CT-proficient mutant D138A rescued this effect, while CT-deficient EndoIII Y82A did not. Overall, this indicates that CT signaling between EndoIII and DinG is important in helping DinG localize to collision sites.

A role for CT in eukaryotic DNA replication?

In addition to BER and NER, conserved [4Fe4S] clusters have been identified in eukaryotic replication proteins, including yeast and human DNA primase and the yeast B-family DNA polymerases (Pols) α, δ, ε, and ζ (*9-10*) **(**Figure 1.9**).** In DNA replication, a replication bubble is generated by two helicase complexes unwinding DNA in opposite directions, and new DNA strands are synthesized in the 5′ → 3′ direction by DNA polymerases, which extend from a short (8-10 nucleotide) RNA primer generated by a primase enzyme (*42*). The directionality of the polymerases necessitates that one strand, the leading strand, undergoes continuous synthesis in the direction of the fork, and the other, the lagging strand, be formed in discontinuous 120-150 nt Okazaki fragments. Under normal conditions, eukaryotic polymerases divide the task of DNA replication, with Pol α adding 10-20 nt of DNA to the RNA primer before ceding the leading strand to Pol ε and the lagging strand to Pol δ (*43*).

Replication presents an additional puzzle beyond DNA repair, as these proteins do not scan the genome, but instead associate in complexes at replication forks (*44*). Faithful duplication of the genome requires a large amount of coordination among DNA-bound [4Fe4S] proteins, which may be facilitated by DNA-mediated redox signaling. Work in our laboratory has indeed shown this to be the case for DNA primase and the lagging strand polymerase Pol δ (*45*). Primase forms a complex with Pol α (Pol-Prim) *in vivo*, and together these proteins form a 20-30 nt RNA-DNA hybrid primer. The means by which the RNA and DNA segment lengths are precisely controlled, and the mechanism of the primase-Pol α handoff are currently poorly understood. In the case of DNA primase, oxidation of the cluster electrochemically results in markedly enhanced DNA binding, and redox activity has been proposed to act as a switch mediating the hand-off from primase to Pol α. In this model, DNA-bound primase in the

$[4Fe4S]^{3+}$ state would synthesize an 8-10 nt RNA primer, at which point reduction of the oxidized primase by the [4Fe4S] cluster of Pol α would terminate primer synthesis and facilitate the hand-off. Signaling between these particular proteins is a compelling possibility, given that primase and Pol α form a single complex flexible enough to position the [4Fe4S] clusters in primase and Pol α for CT signaling.

Upon completion of the RNA-DNA hybrid primer by Pol-Prim, the primer end is handed off to the clamp loader complex, which attaches the circular sliding clamp PCNA to the primed end; PCNA is then bound by Pol ε on the leading strand and Pol δ on the lagging strand to synthesize long stretches of DNA in a processive manner (*46*). Because the clamp loader does not contain a [4Fe4S] cluster, the Pol δ cluster must have a purpose other than primer handoff. Among other possibilities, a role for the cluster in sensing and coordinating the response to stress during DNA replication was particularly appealing. This intriguing problem led to extensive work with Pol δ, which is described in detail in Chapters 5 through 7.

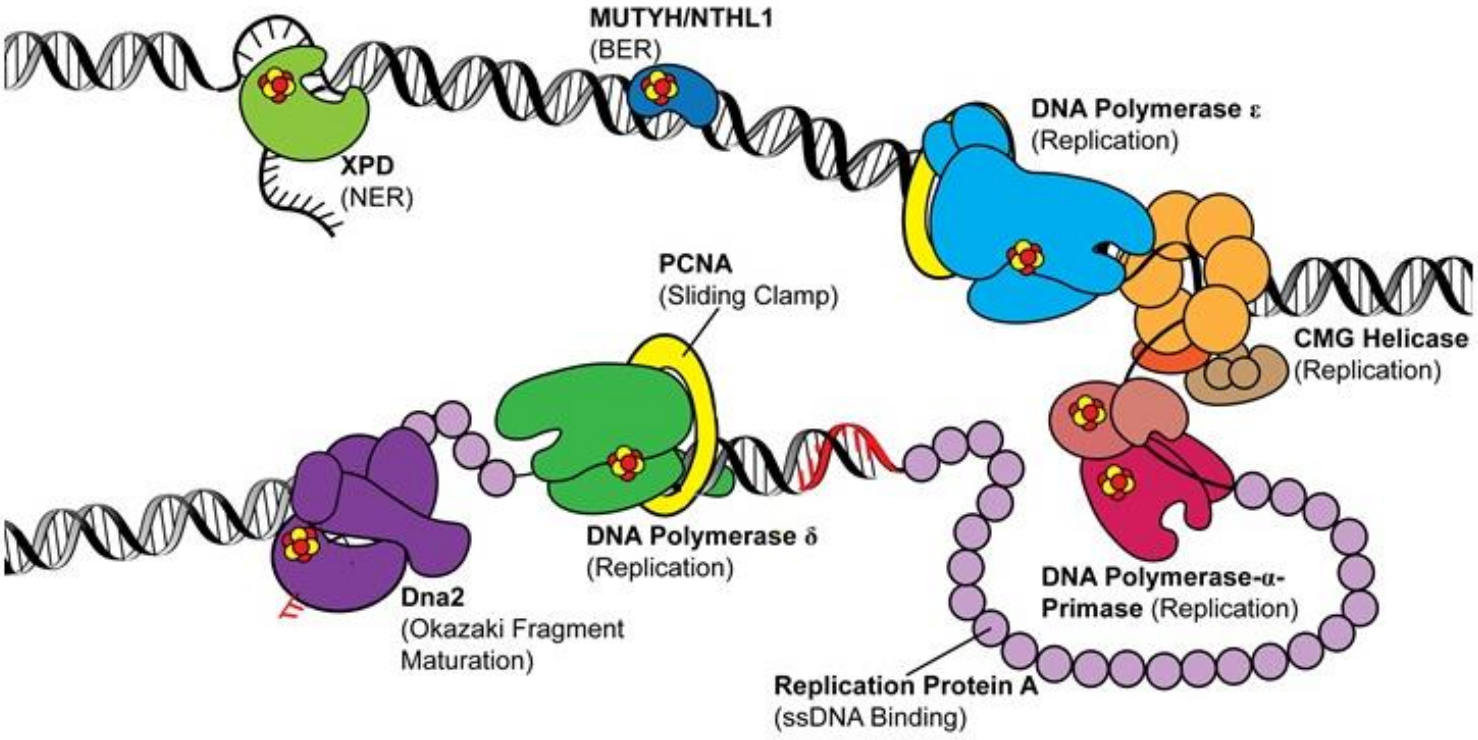


Figure 1.9 DNA-binding eukaryotic enzymes containing [4Fe4S] clusters are found both in replication and in repair pathways. The three polymerases (Polymerase-α-Primase, Polymerase ε, and Polymerase δ) responsible for copying genomic DNA from a parent template all contain a cluster. Polymerase-α-Primase initiates replication through RNA/DNA primer synthesis; Polymerase ε and Polymerase δ then take over replication on the leading and lagging strand of the replication fork, respectively. Dna2 helicase-nuclease, instrumental in Okazaki fragment processing, also contains a [4Fe4S] cluster in its nuclease domain. Repair enzymes, such as XPD helicase in the TFIIH complex (NER) and MUTYH/NTHL1 glycosylases (BER) search for and repair any mismatched/damaged bases on the nascent genomic DNA. This array of [4Fe4S] enzymes in several pathways, with the ability to communicate with one another through DNA CT, suggests that redox signaling may play a larger role in coordinating the complex and dynamic process of eukaryotic replication.

DNA-binding [4Fe4S] proteins in human disease

Many DNA processing proteins containing [4Fe4S] clusters are known to be involved in human disease, with non-catalytic mutations near the cluster being surprisingly prevalent. These proteins include the human homologues of MutY and EndoIII (MUTYH and NTHL1, respectively) and XPD; poorly understood mutations in the [4Fe4S] domain are also present in DNA primase and Pol δ. MUTYH is a critical player in colon cancer, specifically in MUTYH-associated polyposis (*13*). Several poorly characterized mutations in residues near the cluster have recently been recognized. While difficult to understand from conventional perspectives, preliminary electrochemical studies have revealed at least one of these mutants to be more sensitive to oxygen-mediated cluster degradation as described further in Chapter 4 (*48*). Cluster degradation would be especially devastating from the perspective of a CT-based damage search, compromising not only MUTYH but other repair proteins within the redox signaling network. Similarly, NTHL1 has been recognized as important in a variety of different cancers, although mutations in NTHL1 have not been studied to the extent that those in MUTYH have (*14*). XPD mutations are better known for their direct role in three major genetic disorders: trichothiodystrophy, Cockayne syndrome, and xeroderma pigmentosum. These disorders cause a photosensitivity phenotype that ultimately results in the development of cancer and/or accelerated aging. Cancer-relevant mutations in the [4Fe4S] domain have also been reported in both DNA primase and Pol δ (*49, 50*); however, mutations in replicative polymerases are relatively rare due to the essential nature of these enzymes, and those that have been studied generally occur in catalytic domains (*51, 52*). From these examples, it is clear, however, that non-catalytic [4Fe4S] clusters in DNA-processing enzymes are more relevant to disease than

previously suspected. Some work has already implicated defective CT capabilities as important, but many questions remain to be answered.

Conclusions

Our understanding of the role of [4Fe4S] clusters in DNA-processing enzymes has progressed from that of an unusually complex structural group to that of a critical element of rapid, long-range redox signaling along DNA. Considered from another perspective, the recognition of redox-signaling between [4Fe4S] proteins has demonstrated the importance of DNA-mediated CT to biological systems. Indeed, DNA CT is crucial for the identification of lesions by low-copy number proteins like MutY, which would otherwise be unable to find their targets on a relevant timescale. Although this body of work has come a long way toward explaining daunting problems regarding [4Fe4S] proteins and the role of DNA CT in the cell, it opens the door to numerous further questions. These include the nature and origin of the unprecedented potential shift itself, the role of the cluster in disease, and the cluster's function in DNA replication proteins that serve very different purposes than the [4Fe4S] protein studied up to this point.

The work described in following chapters builds upon this basis and represents an effort to answer some of these unresolved questions. This starts with problems in the same bacterial proteins described in this chapter, specifically addressing the role of charged amino acid residues near the cluster in modulating the redox potential and probing the molecular source of the large potential shifts seen on HOPG. Moving beyond bacteria, the role of the cluster in human repair proteins and its relevance to disease was directly addressed for the first time when a remarkable opportunity to study a novel human MUTYH mutation was provided by clinical researchers. Finally, the relevance of clusters in eukaryotic DNA replication, and in the B-family DNA

polymerases in particular, was considered in depth. Indeed, the polymerases, along with DNA primase, represent the first efforts at studying DNA CT in proteins outside of DNA repair. As will hopefully be made clear, despite the diverse functions of these proteins, they all use DNA CT as a regulatory switch to rapidly coordinate activity between pathways and under stress conditions. From a slightly different perspective, this work provides more support for the notion that DNA CT, far from being an esoteric mechanism useful only in detection devices, can serve as a sensor for the integrity of the genome and expedites the cellular response to DNA damage, oxidative stress, and other genomic insults.

References

1) Beinert, H.; Holm, R.H.; Münck, E. Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures. *Science* **1997**, *277*, 653 – 659.

2) Jacobs, A.L.; Schär, P. DNA glycosylases: in DNA repair and beyond. *Chromosoma* **2012**, *121*, 1 – 20.

3) Cunningham, R.P., Asahara, H., Bank, J.F., Scholes, C.P., Salerno, J.C., Surerus, K., Münck, E., McCracken, J., Peisach, J., and Emptage, M.H. Endonuclease III is an iron-sulfur protein. *Biochemistry* **1989**, *28*, 4450 – 4455.

4) You, H.J.; Swanson, R.L.; Doetsch, P.W. *Saccharomyces cerevisiae* Possesses Two Functional Homologues of *Escherichia coli* Endonuclease III. *Biochemistry* **1998**, *37*, 6033 – 6040.

5) Markkanen, E.; Dorn, J.; Hübscher, U. MUTYH DNA glycosylase: the rationale for removing undamaged bases from the DNA. *Frontiers in Genetics* **2013**, *4*, 1 – 20.

6) Porello, S.L.; Cannon, M.J.; David, S.S. A substrate recognition role for the $[4Fe\text{-}4S]^{2+}$ cluster of the DNA repair glycosylase MutY. *Biochemistry*. **1998**, *37*, 6465 – 6475.

7) Ren, B.; Duan, X.; Ding, H. Redox Control of the DNA-damage inducible Protein DinG Helicase Activity via Its Iron-Sulfur Cluster. *J. Biol. Chem.* **2009**, *284*, 4829 – 4835.

8) Fan, L.; Fuss, J.O.; Cheng, Q.J.; Arvai, A.S.; Hammel, M.; Roberts, V.A.; Cooper, P.K.; Tainer. J.A. XPD Helicase Structure and Activities: Insights into the Cancer and Aging Phenotypes from XPD Mutations. *Cell*, **2008**, *133*, 789 – 800.

9) **a.** Klinge, S.; Hirst, J.; Maman, J.D.; Krude, T.; Pellegrini, L. An iron-sulfur domain of the eukaryotic primase is essential for RNA primer synthesis. *Nat. Struct. Mol. Biol.* **2007**, *14*, 875 – 877. **b.** Weiner, B.E.; Huang, H.; Dattilo, B.M.; Nilges, M.J.; Fanning, E.; Chazin, W.J. An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase. *J. Biol. Chem.* **2007**, *282*, 33444 – 33451.

10) Netz, D.J.A.; Stith, C.M.; Stümpfig, M.; Köpf, G.; Vogel, D.; Genau, H.M.; Stodola, J.L.; Lill, R.; Burgers, P.M.J.; Pierik, A.J. Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat. Chem. Biol.* **2012**, *8*, 125 – 132.

11) Pokharel, S.; Campbell, J.S. Cross talk between the nuclease and helicase activities of Dna2: role of an essential iron-sulfur cluster domain. *Nucleic Acids Res.* **2012**, *40*, 7821 – 7830.

12) Lessner, F.H.; Jennings, M.E.; Hirata, A.; Duin, E.C.; Lessner, D.J. Subunit D of RNA Polymerase from *Methanosarcina acetivorans* Contains Two Oxygen-labile [4Fe-4S] Clusters: Implications for Oxidant-dependent Regulation of Transcription. *J. Biol. Chem.* **2012**, *287*, 18510 – 18523.

13) D'Agostino, V.G., Minoprio, A., Torreri, P., Marinoni, I., Bossa, C., Petrucci, T.C., Albertini, A.M., Ranzani, G.N., Bignami, M., and Mazzei, F. Functional analysis of MUTYH mutated proteins associated with familial adenomatous polyposis. *DNA Repair (Amst)*. **2010**, *9*, 700 – 707.

14) Weren, R.D.A. et al. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nat. Genet*. **2015**, *47*, 668 – 671.

15) Thayer, M.M.; Ahern, H.; Xing, D.; Cunningham, R.P.; Tainer, J.A. Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *The EMBO J.* **1995**, *14*, 4108 – 4120.

16) Transviña-Arenas, C.H.; Lopez-Castillo, L.M.; Sanchez-Sandoval, E.; Brieba, L.G. Dispensability of the [4Fe-4S] cluster in novel homologues of adenine glycosylase MutY. *FEBS J.* **2016**, *283*, 521 – 540.

17) Netz; D.J.A.; Mascarenhas, J.; Stehling, O.; Pierik, A.J.; Lill, R. Maturation of cytosolic and nuclear iron-sulfur proteins. *Trends in Cell Biol.* **2014**, *24*, 303 – 312.

18) Arnold, A.R.; Grodick, M.A.; Barton, J.K. DNA Charge Transport: from Chemical Principles to the Cell. *Cell Chem. Biol.* **2016**, *23*, 183 – 197.

19) Slinker, J.D.; Muren, N.B.; Renfrew, S.E.; Barton, J.K. DNA charge transport over 34 nm. *Nature Chem.* **2011**, *3*, 228 – 233.

20) Boal, A.K.; Yavin, E.; Lukianova, O.A.; O'Shea, V.L.; David, S.S.; Barton, J.K. DNA-Bound Redox Activity of DNA Repair Glycosylases Containing [4Fe-4S] Clusters. *Biochemistry* **2005**, *44*, 8397 – 8407.

21) Backes, G.; Mino, Y.; Loehr, T.M.; Meyer, T.E.; Cusanovich, M.A.; Sweeney, W.V.; Adman, E.T.; Sanders-Loehr, J. The environment of Fe_4S_4 clusters in ferredoxins and high-potential iron proteins: New information from x-ray crystallography and resonance Raman spectroscopy. *J. Am. Chem. Soc.* **1991**, *113*, 2055 – 2064.

22) Gorodetsky, A.A.; Boal, A.K.; Barton, J.K. Direct electrochemistry of Endonuclease III in the presence and absence of DNA. *J. Am. Chem. Soc.* **2006**, *128*, 12082 – 12083.

23) **a.** Fromme, J.C.; Verdine, G.L. Structure of a trapped endonuclease III-DNA covalent intermediate. *EMBO J.* **2003**, *22*, 3461 – 3471. **b.** Guan, Y.; Manuel, R.C.; Arvai, A.S.; Parikh, S.S.; Mol, C.D.; Miller, J.H.; Lloyd, R.S.; Tainer, J.A. MutY catalytic core, mutant, and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nat. Struct. Biol.* **1998**, *5*, 1058 – 1064. **c.** Woods, R.D.; O'Shea, V.L.; Chu, A.; Cao, S.; Richards, J.L.; Horvath, M.P.; David, S.S. Structure and stereochemistry of the base excision repair glycosylase MutY reveal a mechanism similar to retaining glycosylases. *Nucleic Acids Research* **2015**, *44*, 801 – 810.

24) Bartels, P.L.; Zhou, A.; Arnold, A.R.; Chaffee, N.; Crespilho, F.N.; David, S.S.; Barton, J.K. Electrochemistry of the [4Fe4S] cluster in base excision repair proteins: Tuning the redox potential with DNA. *Langmuir*, **2017**, *33*, 2523 – 2530.

25) O'Neill, M.A.; Becker, H.-C.; Wan, C.; Barton, J.K.; Zewail, A.H. Ultrafast Dynamics in DNA-Mediated Electron Transfer: Base Gating and the Role of Temperature. *Angew. Chem. Int. Ed. Engl.*, **2003**, *42*, 5896 – 5900.

26) Nunez, M.E.; Hall, D.B.; Barton, J.K. Long-Range Oxidative Damage to DNA: Effects of Distance and Sequence. *Chem. Biol. (Oxford, U. K.)*, **1999**, *6*, 85 – 97.

27) Hall, D.B.; Holmlin, R.E.; Barton, J.K. Oxidative DNA Damage through Long Range Electron Transfer. *Nature*, **1996**, *382*, 731 – 735.

28) Yavin, E.; Boal, A.K.; Stemp, E.D.A; Boon, E.M.; Livingston, A.L.; O'Shea, V.L.; David, S.S.; Barton, J.K. Protein-DNA Charge Transport: Redox Activation of a DNA Repair Protein by a Guanine Radical. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3546 – 3551.

29) Lee, P.E.; Demple, B.; Barton, J.K. DNA-mediated redox signaling for transcriptional activation of SoxR. *Proc. Natl. Acad. Sci. U. S. A.*, **2009**, *106*, 13164 – 13168.

30) Nunez, M.E.; Noyes, K.T.; Barton, J.K. Oxidative Charge Transport through DNA in nucleosome particles. *Chemistry & Biology* **2002**, *9*, 403 – 415.

31) Gorodetsky, A.A.; Ebrahim, A.; Barton, J.K. Electrochemical Detection of TATA Binding Protein at DNA-Modified Microelectrodes. *J. Amer. Chem. Soc.* **2008**, *130*, 2924 – 2925.

32) Boal, A.K.; Genereux, J.C.; Sontz, P.A.; Gralnick, J.A.; Newman, D.K.; Barton, J.K. Redox signaling between repair proteins for efficient lesion detection. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15237 – 15242.

33) Winkler, J.R.; Gray, H.B. Could tyrosine and tryptophan serve multiple roles in biological redox processes? *Philo. Trans. A: Math Phys. Eng. Sci.* **2015**, *373*, DOI: 10.1098/RSTA.2014.0178

34) Romano, C.A.; Sontz, P.A.; Barton, J.K. Mutants of the Base Excision Repair Glycosylase, Endonuclease III: DNA Charge Transport as a First Step in Lesion Detection. *Biochemistry* **2011**, *50*, 6133 – 6145.

35) Verdine, G.L.; Bruner, S.D. How do DNA repair proteins locate damaged bases in the genome? *Chemistry & Biology* **1997**, *4*, 329 – 334.

36) Marteijn, J.A.; Lans, H.; Vermeulen, W.; Hoeijmakers, J.H.J. Understanding nucleotide excision repair and its roles in cancer and aging. *Nature Reviews Mol. Cell Biol.* **2014**, *15*, 465 – 481.

37) Mui, T.P.; Fuss, J.O.; Ishida, J.P.; Tainer, J.A.; Barton, J.K. ATP-Stimulated, DNA-Mediated Redox Signaling by XPD, a DNA Repair and Transcription Helicase. *J. Amer. Chem. Soc.* **2011**, *133*, 16378 – 16381.

38) Sontz, P.A.; Mui, T.P.; Fuss, J.O.; Tainer, J.A.; Barton, J.K. DNA charge transport as a first step in coordinating the detection of lesions by repair proteins. *Proc. Natl. Acad. Sci. USA*, **2012**, *109*, 1856 – 1861.

39) Voloshin, O.N.; Camerini-Otero, R.D. The DinG Protein from *Escherichia coli* Is a Structure-specific Helicase. *J. Biol. Chem.* **2007**, *282*, 18437 – 18447.

40) Grodick, M.A.; Segal, H.M.; Zwang, T.J.; Barton, J.K. DNA-Mediated Signaling by Proteins with 4Fe-4S Clusters Is Necessary for Genomic Integrity. *J. Amer. Chem. Soc.* **2014**, *136*, 16470 – 16478.

41) Boubakri, H.; Langlois de Septenville, A.; Viguera, E.; Michel, B. The helicases DinG, Rep, and UvrD cooperate to promote replication across transcription units *in vivo*. *EMBO J.* **2010**, *29*, 145 – 157.

42) Leman, A.R.; Noguchi, E. The Replication Fork: Understanding the Eukaryotic Replication Machinery and the Challenges to Genome Duplication. *Genes (Basel)* **2013**, *4*, 1 – 32.

43) Lujan, S.A.; Williams, J.S.; Kunkel, T.A. DNA Polymerases Divide the Labor of Genome Replication. *Trends in Cell Biology* **2016**, *26*, 640 – 654.

44) Sun, J.; Shi, Y.; Georgescu, R.E.; Yuan, Z.; Chait, B.T.; Li, H.; O'Donnell, M.E. The architecture of a eukaryotic replisome. *Nat. Strut. Mol. Biol.* **2015**, *22*, 976 – 982.

45) **a.** O'Brien, E.; Holt, M.E.; Thompson, M.K.; Salay, L.E.; Ehlinger, A.C.; Chazin, W.J.; Barton, J.K. The [4Fe4S] cluster in human DNA primase functions as a redox switch using DNA charge transport. *Science* **2017**, *355*, 813. **b.** Bartels, P.L.; Stodola, J.L.; Burgers, P.M.J.; Barton, J.K. A redox role for the [4Fe4S] cluster of yeast DNA Polymerase δ. *J. Am. Chem. Soc.* **2017**, *139*, 18339 – 18348.

46) Georgescu, R.E.; Schauer, G.D.; Yao, N.Y.; Langston, L.D.; Yurieva, O.; Zhang, D.; Finkelstein, J.; O'Donnell, M.E. Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. *eLife* **2015**, DOI: 10.7554/eLife.04988

47) Leroy, C.; Mann, C.; Marsolier, M. Silent repair accounts for cell cycle specificity in the signaling of oxidative DNA lesions. *EMBO J.* **2001**, *20*, 2896 – 2906.

48) McDonnell, K.J.; Chemler, J.A.; Bartels, P.L.; O'Brien, E.; Marvin, M.L.; Ortega, J.; Stern, R.H.; Raskin, L.; Li, G.; Sherman, D.H.; Barton, J.K.; Gruber, S.B. A Human MUTYH Variant Linking Colonic Polyposis to Redox Degradation of the $[4Fe4S]^{2+}$ Cluster. **2018**, *Submitted.*

49) Forbes, S.A. et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **2011**, *39*, D945 - D950

50) Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. *Nat. Comm.* **2015**, doi: 10.1038/ncomms7604

51) Fuss, J.O.; Tsai, C.; Ishida, J.P.; Tainer, J.A. Emerging critical roles of Fe-S clusters in DNA replication and repair. *Biochimica et Biophysica Acta* **2015**, *1853*, 1253 – 1271.

52) Flohr, T.; Dai, J.; Büttner, J.; Popanda, O.; Hagmüller, E.; Thielmann, H.W. **1999**. Detection of mutations in the DNA polymerase δ gene of human sporadic colorectal cancers and colon cancer cell lines. *International J. of Cancer.* **1999**, *80*, 919 – 929.

Chapter 2

Electrochemistry of the [4Fe4S] Cluster in Base Excision Repair Proteins: Tuning the Redox Potential with DNA

Adapted from: Bartels, P.L.; Zhou, A.; Arnold, A.R.; Nuñez, N.N.; Crespilho, F.N.; David, S.S.; Barton, J.K. Electrochemistry of the [4Fe4S] Cluster in Base Excision Repair Proteins: Tuning the Redox Potential with DNA. *Langmuir*, **2017**, *33*, 2523 – 2530.

P. Bartels performed all electrochemical experiments with assistance from F. Crespilho and A. Zhou. A. Zhou, A. Arnold, and N. Nuñez prepared proteins for experiments.

Introduction

E. coli endonuclease III (EndoIII) is a DNA glycosylase that excises oxidized pyrimidines from DNA, functioning as part of the base excision repair (BER) pathway in order to maintain the integrity of the genome (*1*). EndoIII contains a $[4Fe4S]^{2+}$ cluster that is relatively insensitive to reduction and oxidation in solution (*2*); as a result, it was initially proposed that the cluster served only a structural role within the protein. MutY is another *E. coli* BER glycosylase, homologous to EndoIII, that also contains a $[4Fe4S]^{2+}$ cluster (*3*). MutY, found in organisms from bacteria to man, is involved in the repair of oxoG:A mismatches (*4*); in humans, inherited defects in MUTYH are associated with a familial form of colon cancer known as MUTYH-associated polyposis (MAP) and many MAP-associated variants are localized near the [4Fe4S] cluster (*4*). Furthermore, in the case of MutY, it has been shown that the cluster is not required for folding or stability (*3*), or direct participation in the intrinsic glycosidic bond hydrolysis catalysis (*5*), making the widespread presence of conserved, non-catalytic [4Fe4S] clusters difficult to explain.

Notably, the earliest studies with EndoIII and MutY looked only at free protein in solution, neglecting the effect of DNA binding on redox potential. Experiments carried out on DNA-modified electrodes have demonstrated that, in both EndoIII and MutY, the cluster undergoes a negative shift in potential associated with binding to the DNA polyanion and is activated toward reversible redox activity (*6*). In these experiments, DNA monolayers were formed on gold electrodes, and, upon addition of EndoIII or MutY, a reversible signal with a midpoint potential ranging from 60-95 mV versus NHE was observed. Importantly, the introduction of just a single mismatch or abasic site into DNA led to signal attenuation, showing that electron transfer between the protein and the electrode was through the π-stacked base pairs

in a process known as DNA-mediated charge transport (DNA CT) (*7*). In this process, charge is funneled from the electrode surface through the π-stack of the DNA bases to reach the redox probe (a protein in this case); the only requirement is that the probe must be electronically coupled to the DNA π-stack. Remarkably, the sensitivity to base stacking observed with EndoIII and MutY was comparable to that obtained using small molecules such as Nile blue or methylene blue that intercalate directly into the base stack.

The expanded potential window of highly-oriented pyrolytic graphite (HOPG) and the ability to form pyrene-modified DNA films on the surface made it possible to directly compare the potential of proteins in the presence and absence of DNA (*8*). Experiments with EndoIII revealed that DNA binding shifts the reduction potential of the $[4Fe4S]^{3+/2+}$ couple by -200 mV to favor oxidation. Thermodynamically, this shift corresponded to a large (~3 orders of magnitude) increase in the DNA binding affinity of the oxidized form of the protein. Crystal structures of EndoIII and MutY with and without DNA do not show any significant structural change upon DNA binding (*9*-*12*), so this dramatic result was attributed to a combination of electrostatic effects resulting from the negatively charged DNA backbone and decreased solvent accessibility of the cluster in DNA-bound protein, which is in agreement with the known sensitivity of [4Fe4S] clusters to their local environment (*13*). By demonstrating that DNA binding brought the redox potential of EndoIII into a biologically relevant window, this result served to explain the previously observed redox insensitivity of free EndoIII and provided evidence in favor of a redox role for the DNA-bound protein cluster.

Since these experiments were carried out, a wide range of DNA processing enzymes have been revealed to contain [4Fe4S] clusters with properties similar to EndoIII and MutY. These include the *Archaeoglobus fulgidus* uracil DNA glycosylase (UDG), archaeal and eukaryotic

versions of the nucleotide excision repair helicase XPD and the *E. coli* R-loop maturation helicase DinG (*14),* all of which were found to have similar DNA-bound potentials (~80 mV versus NHE) as measured on DNA-modified gold electrodes (*6,7*). The similar DNA-bound midpoint potentials and picosecond kinetics of DNA CT together suggested that DNA CT could provide a means for these enzymes to localize efficiently to the vicinity of their target lesions (*15*). Indeed, experiments carried out both *in vitro* and *in vivo* have led to the development of a model for DNA repair in which two [4Fe4S] cluster proteins use DNA CT to communicate with each other over long molecular distances via electron transfer self-exchange reactions (*7, 15*). As evidenced through the potential shift, DNA binding activates the proteins toward oxidation to the $[4Fe4S]^{3+}$ state (*8*). When the DNA intervening between the two proteins is undamaged, the self-exchange reaction can proceed efficiently, with the result that one of the DNA-bound proteins is reduced and its affinity for DNA lowered. This protein is then free to diffuse to another region of the genome. However, in the case of an intervening mismatch or lesion that impairs CT by disrupting π-stacking, this self-exchange reaction is inhibited. Both proteins then remain bound to the DNA in the vicinity of the lesion, significantly reducing the range over which the slower processes of diffusion must occur and facilitating repair of a relatively large genome on a biologically relevant time scale (*15*).

While DNA binding is clearly of critical importance to the redox activity of these enzymes, it is not clear that it represents the only way to modulate the potential. It was recently reported that carboxylic acid monolayers had a similar activating effect as DNA, although, in contrast with the above model, they identified the relevant couple as the $[4Fe4S]^{2+/+}$ rather than the $[4Fe4S]^{3+/2+}$ couple (*16*). With respect to the latter point, the high potential of the reversible DNA-bound signal on both gold and HOPG, EPR spectroscopy of oxidized DNA bound EndoIII

and MutY, and the observation of both couples in the expected potential regimes on HOPG support the original $[4Fe4S]^{3+/2+}$ assignment (*6,8,15*). Furthermore, this assignment is in agreement with the known potential ranges accessed by the $[4Fe4S]^{3+/2+}$ couple of HiPIPs (*17*). Regardless of redox couple assignment, the possibility of activation by other molecules remains an interesting point deserving further investigation.

In addition to other molecules, charged amino acid residues near the cluster might also be expected to affect the potential. This was explored in a recent study in which several EndoIII mutants, E200K, Y205H, and K208E, were prepared and extensively characterized on DNA-modified gold electrodes; although these residues are located within 5 Å of the cluster, all of the mutants had indistinguishable DNA-bound midpoint potentials (*18*). Overall, these observations suggested that DNA binding was the dominant environmental effect in modulating potential, but the narrow accessible potential window on gold prevented further investigation.

In this work, we used direct electrochemistry on carbon electrodes to address the capacity of DNA and other polyanions to activate [4Fe4S] proteins for redox activity and to assess the ability of local electrostatics to shift the potential of EndoIII in the absence of DNA. Because the hydrophobic surface of HOPG is unsuitable for protein adsorption and difficult to prepare (*8, 19-20*), we turned to the rougher, more hydrophilic pyrolytic graphite edge (PGE) electrode for these experiments, using the technique of thin film voltammetry to immobilize proteins in a stable layer on the electrode surface (*21-24*). To enhance signal sizes, we also included single-walled carbon nanotubes (CNT) when possible, taking advantage of their high conductivity and the additional 3-dimensional surface area they can provide for protein binding (*21*). In summary, this platform provided an ideal and reliable way to improve our understanding of the factors important to tuning the potential of DNA processing enzymes containing [4Fe4S] clusters.

Materials and Methods

EndoIII Overexpression and Purification

WT *E. coli* EndoIII was overexpressed in BL21star-(DE3)pLysS cells containing a pET11-ubiquitin-His_6-*nth* construct and purified as detailed previously (*18*), with the exception that the final buffer contained 10% rather than 20% glycerol (20 mM sodium phosphate, pH 7.5, 0.5 mM EDTA, 150 mM NaCl, 10% glycerol). For electrochemical experiments, glycerol was removed from the protein solution using a HiPrep 26/10 desalting column (GE Healthcare) equilibrated with a buffer consisting of 20 mM sodium phosphate, pH 7.5, 0.5 mM EDTA, 150 mM NaCl. Following buffer exchange, the protein was concentrated in two steps. First, 10,000 molecular weight cutoff (MWCO) Amicon Ultra 15 mL centrifugation filter units (Millipore) were used to concentrate each protein solution to a total volume of 1 mL or less. Samples were then transferred to 10,000 MWCO Amicon Ultra 0.5 mL centrifugation filter units (Millipore) and concentrated until the initially yellow protein solutions were very dark in color (approximately 300 μL final volume from 6 L of bacterial culture). Protein purity was confirmed by SDS-PAGE. Immediately following concentration of the sample, the [4Fe4S] cluster loading ratio was calculated by dividing the total [4Fe4S] cluster concentration as determined from the UV-visible absorbance spectrum using ε_{410} = 17,000 $M^{-1}cm^{-1}$ by the total protein concentration as measured in a Bradford assay; typical cluster loading ratios for WT EndoIII were 70-75%.

MutY Overexpression and Purification

MBP (Maltose Binding Protein)-MutY fusion protein was expressed and purified using a slightly modified version of a previously reported protocol (*25*). Modifications to the protocol included changes in “buffer A” to a resuspension buffer (20 mM sodium phosphate, pH 7.5, 200 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF, and 10% glycerol) and use of an amylose

column to eliminate the necessity of a streptomycin sulfate and ammonium sulfate precipitation. During the amylose preparation, the sample was washed with amylose wash buffer (20 mM sodium phosphate, pH 7.5, 200 mM NaCl, 1mM EDTA pH 8) and eluted in amylose elutant buffer (20 mM sodium phosphate, pH 7.5, 200 mM NaCl, 1 mM EDTA pH 8, 10 mM maltose). The resultant fractions were concentrated using an ultrafiltration cell with a 10,000 MWCO filter with stirring at 4^{o}C. Protein was then diluted 10-fold in heparin buffer A (20 mM sodium phosphate, pH 7.5, 1 mM EDTA, 5% glycerol in water), applied to a Pharmacia Hi-trap heparin column on an AKTApurifier FPLC system, and eluted using a 10% linear gradient in heparin buffer A to 100% heparin buffer B (20mM sodium phosphate, pH 7.5, 1mM EDTA, 5% glycerol, and 1 M NaCl in water). MBP-MutY eluted at 450 mM NaCl (45% heparin buffer B). Purity of protein samples was confirmed via 12% SDS page stained with SYPRO orange. The [4Fe4S] cluster loading was determined using the UV-visible absorbance at 410 nm (ε_{410} = 17,000 $M^{-1}cm^{-1}$) and at 280 nm (ε_{280} = 143,240 $M^{-1}cm^{-1}$); samples were typically 65-75% loaded.

DNA Preparation

DNA strands for EndoIII experiments were purchased from Integrated DNA Technologies, with sequences as follows:

20-mer: 5′-GTG AGC TAA CGT GTC AGT AC-3′

Complement: 5′-GTA CTG ACA CGT TAG CTC AC-3'

Single-stranded DNA oligomers (5 μmol) were resuspended in MilliQ water and purified by ethanol precipitation. Briefly, 1000 μL of cold 200 proof ethanol and 50 μL of 3 M NaCl were added to 100 μL single-stranded DNA in water and vortexed; DNA solutions were then frozen in

liquid nitrogen for rapid precipitation and spun at 16,000 RCF (25 minutes) to form a pellet which was then re-dissolved in EndoIII storage buffer (20 mM sodium phosphate, pH 7.5, 0.5 mM EDTA, 150 mM NaCl). Single-stranded DNA was quantified by UV-vis using ε_{260} values calculated using the Integrated DNA Technologies oligo analyzer tool; these were 197,800 $M^{-1}cm^{-1}$ for the 20-mer strand and 190,200 $M^{-1}cm^{-1}$ for its complement. Equimolar amounts of each strand were then annealed by incubation at 90°C for 5 minutes followed by slow cooling to ambient temperature.

For MutY experiments, DNA substrates containing oxoG (8-oxo-guanine) or FA (2′-fluoro-adenine) were synthesized at the University of Utah DNA and Peptide Synthesis Core Facility and unmodified strands were ordered from Integrated DNA Technologies. The following DNA duplexes were used:

15-mer: 5′-GGA GCC A**X**G AGC TCC-3

15-mer Complement: 3′-CCT CGG T**Y**C TCG AGG-5′

30-mer: 5′-CGA TCA TGG AGC CAC **X**AG CTC CCG TTA CAG-3′

30-mer Complement: 3′-GCT AGT ACC TCG GTG **Y**TC GAG GGC AAT GTC-5′

X = G or oxoG and **Y** = C, FA, or A

Oligonucleotides containing the central oxoG or FA were deprotected and cleaved from the column by incubation in NH_4OH; 2-mercaptoethanol was added into oxoG samples to prevent

oxidation. The cleaved DNA substrates were dissolved in H_2O, filtered with a 0.2 μm filter, and HPLC purified using a Beckman Gold Nouveau system with a Waters AP1DEAE 8HR column; a 10-100% gradient of 90:10 H_2O/acetonitrile with 2 M NH_4Ac was used in purification. Isolated fractions were dried down and de-salted using SEP-PAK cartridges, and DNA integrity was confirmed using MALDI-MS. All DNA substrates were stored dried in the -20^oC freezer prior to annealing.

Electrochemistry

All electrochemical experiments were performed on an edge-plane pyrolytic graphite electrode (Pine Research Instrumentation) with a geometric surface area of 0.196 cm^2. To generate a rough surface suitable for protein binding, the electrode was abraded with 400 grit sandpaper and cleaned by sonication for 1 minute each in ethanol and water. After sonication, the absence of electroactive impurities was verified by scanning in EndoIII storage buffer (20 mM sodium phosphate, 0.5 mM EDTA, 150 mM NaCl, pH 7.5) or MutY storage buffer (20 mM sodium phosphate, 1 mM EDTA, 150 mM NaCl, pH 7.6) as appropriate.

Single-walled carbon nanotubes (CNTs) were found to enhance the signal size of adsorbed protein, so they were included in the formation of all thin films unless otherwise noted. Protein thin films were formed from several (typically 3-6) alternate layers of 10 μL single-walled carbon nanotubes (CNT) in water (0.25 mg/mL) and 10 μL EndoIII (150 μM in storage buffer) or MutY (50 μM) in a 1:1 mix with aqueous CNTs. Each layer was gently dried under an argon gun, and the process was repeated until the surface was coated by a viscous film, which was then secured with 5% Nafion in water (diluted from 10% in water as purchased) to prevent dispersal (*21*). For experiments including DNA, CNTs generally hindered electrochemical signals, so these films were formed in their absence. Poly-L glutamate (MW 50-100 kDa) was

used to assess the effects of a negatively charged non-substrate on potential. Although Nafion also carries a negative charge at the pH values used, it was applied only to the top of a multilayer film to form a binding layer, minimizing interactions with the electroactive protein; in contrast, poly-L glutamate and DNA were incorporated directly into the thin film with protein to maximize any possible interactions.

After thin film formation, 50 μL of EndoIII or MutY storage buffer was pipetted on top of the film and an Ag/AgCl reference in 3 M NaCl and Pt auxiliary electrode were submerged in the resulting droplet. Reduction potential, current, and charge measurements were then taken by cyclic voltammetry (CV), square wave voltammetry (SQWV), and differential pulse voltammetry (DPV); all experiments were conducted at ambient temperature (20 °C). Electroactive area was determined by plotting the scan rate dependence of the CV current generated by 1 mM $[Ru(NH_3)_6]Cl_3$ in storage buffer and applying the Randles-Sevcik equation (*26*),

$$I_p = [(0.4463(F^3/RT)^{1/2}](n^{3/2})(A)(D^{1/2})(C^\circ)\nu^{1/2} \qquad (1)$$

I_p is the peak current in amperes, F is Faraday's constant (96485 $C \cdot mol^{-1}$), R is the universal gas constant (8.314 $J \cdot (mol \cdot K)^{-1}$), T is temperature in K, n is the number of electrons transferred per CV peak, A is electrode area in cm^2, D is the diffusion coefficient in $cm^2 \cdot s^{-1}$ (9.0×10^{-6} for $[Ru(NH_3)_6]^{3+}$; *27*), C° is bulk protein concentration in $mol \cdot cm^{-3}$, and ν is the scan rate in $V \cdot s^{-1}$. Potentials were converted to NHE by adding 0.212 V to the value measured by Ag/AgCl, using the value of 0.209 mV value at 25 C given by the supplier, BASi®, and applying a temperature correction (*28*). To prevent leakage of NaCl into the buffer and subsequent wandering of the

reference potential, the glass frit of the electrode was immersed in a gel loading pipet tip containing 3 M NaCl with 4% dissolved agarose, and dried in this mix overnight. CNTs, 10% aqueous Nafion, poly-L glutamate, and $[Ru(NH_3)_6]Cl_3$ were purchased from Sigma-Aldrich, while the Ag/AgCl reference electrode in 3 M NaCl was purchased from BASi®.

Results

Direct Electrochemistry of WT EndoIII and MutY

To examine variations in EndoIII potentials with various substitutions or binding partners, we first prepared WT EndoIII thin films on a PGE electrode. All films were anchored to the surface by Nafion and capped with 50 μL of storage buffer for scanning. In the absence of DNA, a quasi-reversible signal was observed by CV (Figure 1). The signal under these conditions was relatively small (3 ± 1 μC reductive peak, -4 ± 1 μC oxidative peak), and the reductive peak partially overlapped the much larger wave of oxygen reduction, making it challenging to quantify. By adding 0.25 mg/mL CNTs to form a protein/CNT/Nafion thin film, the peak areas increased by an order of magnitude to reach 16 ± 3 μC and -18 ± 5 μC reductive and oxidative peaks, respectively, while the peak potentials remained unaltered from those without the CNTs at 74 ± 20 mV and 162 ± 18 mV vs NHE (Figure 1).

The addition of EndoIII and CNTs was associated with a large increase in the capacitance; much of this increase was due to CNTs, as seen in CNT/Nafion thin films, but the protein itself certainly contributed (Figure 2). Notably, the high conductivity of CNTs amplifies redox events at the surface; a CNT/Nafion film shows reversible peaks around 200 mV vs NHE and -80 mV vs NHE, both of which show no splitting and are likely attributable to the reversible reduction of surface oxides on the edge plane and even on the CNTs themselves (Figure 2) (*29-31*). Indeed, the 200 mV peaks were invariably present, although smaller, in buffer alone and the -80 mV peaks varied in size based upon the freshness of the CNT suspension applied, consistent with this assertion. The presence of protein on the surface markedly suppressed both of these peaks, and the EndoIII signal differed from the background both by its potential, which was

essentially identical to that measured in the absence of CNTs, and in the occurrence of peak splitting; the latter suggested a slower process, in agreement with reports of other proteins adsorbed on carbon (*21*).

WT MutY thin films were prepared just as with EndoIII, although the stock concentrations were somewhat lower (~50 μM for MutY compared to ~150 μM for EndoIII). MutY displayed a quasi-reversible signal similar to EndoIII on CNT/Nafion thin films, with CV peak potentials centered at 100 ± 9 mV for the reductive peak and 162 ± 3 mV for the oxidative peak (Figure 1). Notably, the potentials were within error of the values obtained for EndoIII. The respective peak areas were 2.3 ± 0.3 μC and -3.4 ± 0.1 μC, about an order of magnitude smaller than EndoIII and indicative of lower surface coverage.

For both EndoIII and MutY, the current exhibited a linear dependence on the scan rate (Figure 3), confirming that the protein was adsorbed to the electrode surface rather than diffusing in from solution; this relationship was present whether or not CNTs were included. Surface coverage was initially determined simply by converting the total CV peak charge at a scan rate of 100 mV/s into pmol using Faraday's constant and dividing by the geometric surface area of the electrode. Because the PGE surface is uneven, the geometric surface area can underestimate the electroactive area by a factor as large as 10^4 (*32*). Indeed, using the geometric area of the electrode (0.196 cm^2) gave a surface coverage of 550 ± 300 pmol/cm^2 for 75 μM EndoIII stock, over 10 times larger than reported for ferredoxin thin films on PGE (40 pmol/cm^2) (*32*) and over 100 times larger than CNT/Nafion/protein thin films on glassy carbon (2-6 pmol/cm^2) (*21*).

By taking the scan rate dependence of the current for $[Ru(NH_3)_6]^{3+}$ in EndoIII storage buffer and applying the Randles-Sevcik equation (Equation 1), the electroactive surface area was determined to be 1.0 cm^2, about 5 times larger than the simple geometric area (Figure 4). When

this correction was applied to a thin film formed from 75 μM EndoIII stock, a value of 108 ± 60 pmol/cm^2 was obtained, which is still high but much closer to previously published results on PGE (*32*). Applying the same correction to films formed from 25 μM MutY stock gave a coverage of 29 ± 6 pmol/cm^2, around 25% of that measured for 75 μM EndoIII. To facilitate a more direct comparison, surface coverage on thin films formed with 25 μM EndoIII was measured to be 51 ± 8 pmol/cm^2, indicating that MutY adsorption was absolutely less extensive than EndoIII. This result is not surprising, given that unmodified MutY is 39 kDa while EndoIII is only 24 kDa; the 42 kDa N-terminal MBP tag on MutY would only enhance this issue.

Adsorption of proteins to the electrode surface made it possible to estimate electron transfer rate (k_{ET}) and transfer coefficient (α) values using the Laviron method for diffusionless systems (Figure 5) (*33*), where α is a measure of transition state symmetry, taking on values between 0 and 1. The MutY signal was too small to measure the currents at high scan rates, but this analysis could be carried out for EndoIII. In the case of EndoIII, we obtained a k_{ET} of 3 ± 0.6 s^{-1} and α values of 0.4 and 0.6 for the reductive and oxidative peaks, respectively. Assuming that electron transfer is the only reaction taking place on the electrode, the values of α imply a quasi-reversible system. Importantly, electron transfer rates are similar to those reported for other redox-active enzymes/proteins adsorbed to carbon electrodes in the presence of CNTs and Nafion (*21*).

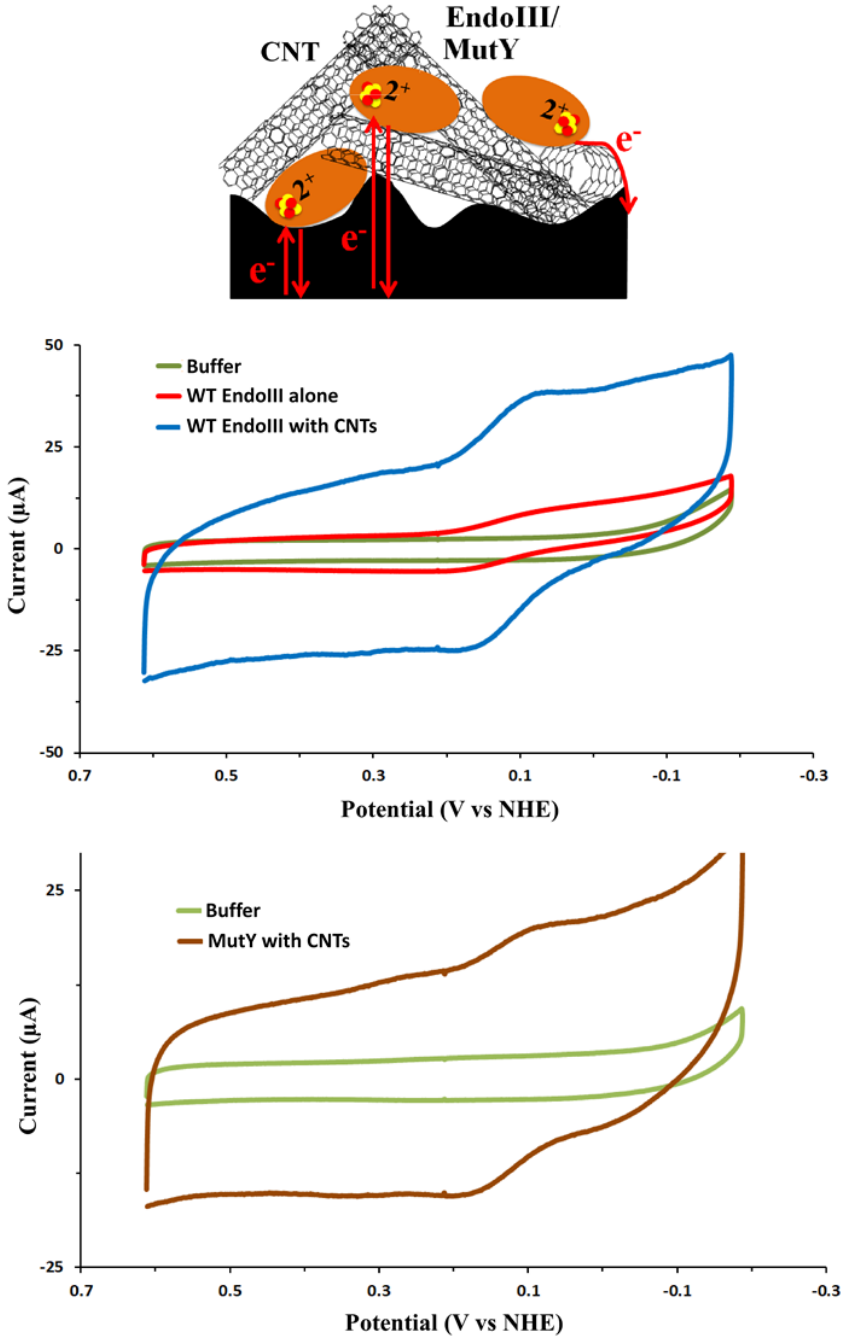


Figure 2.1 Representative CVs from EndoIII and MutY thin-films on a PGE electrode. A thin film containing only 75 µM EndoIII capped with Nafion gave a quasi-reversible signal with reductive and oxidative peaks centered, respectively, at 74 ± 11 and 162 ± 20 mV versus NHE (center; red trace). Notably, the addition of CNTs substantially amplified the signal, simplifying quantification (center; blue trace). Similarly, 25 µM MutY in the presence of CNTs yielded a signal with reductive and oxidative peaks at 100 ± 9.0 and 162 ± 3.0 mV versus NHE (bottom). CV scans were taken at a scan rate of 100 mV/s.

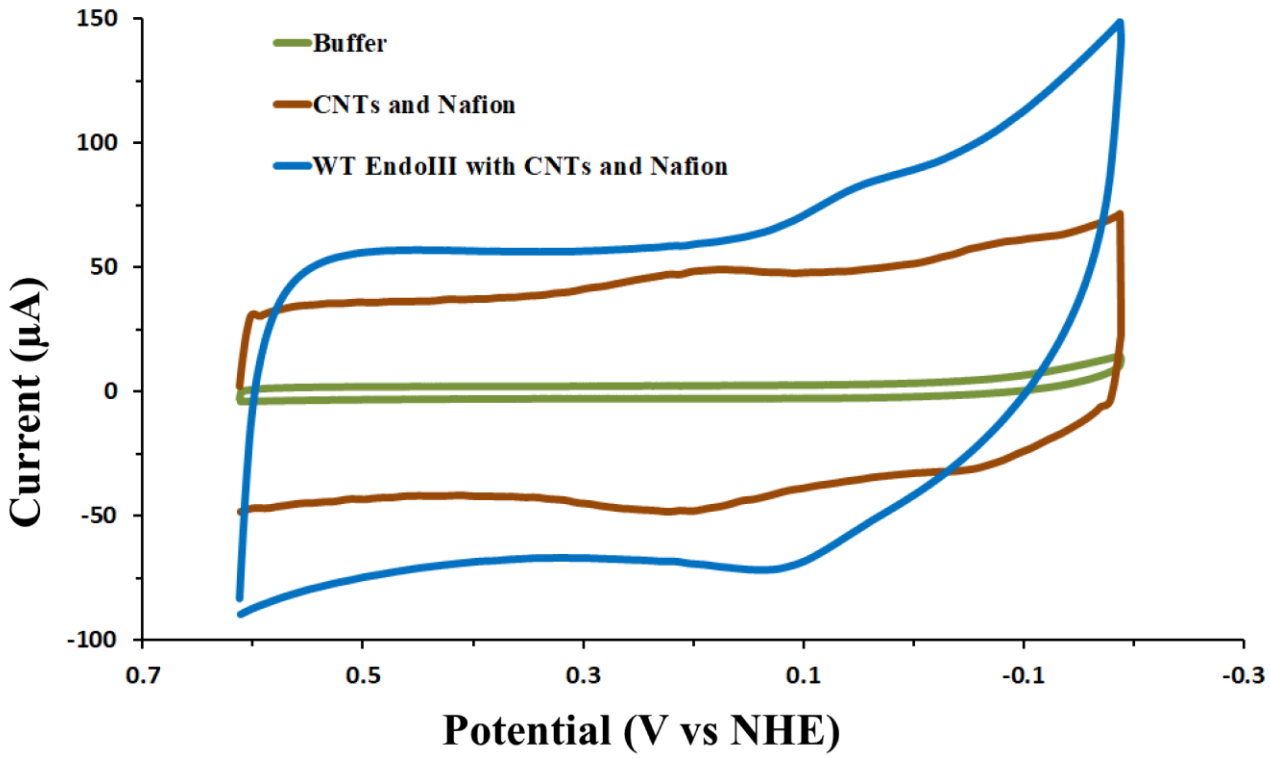


Figure 2.2 CV of storage buffer, a CNT/Nafion thin film, and a WT EndoIII/CNT/Nafion thin film. CNTs (from a 0.25 mg/mL stock in water) increased the capacitance dramatically relative to that observed with storage buffer alone and strongly amplified the signal around 200 mV vs NHE that can be attributed to oxides present on the electrode surface. In the presence of CNTs, an additional signal occurs around -80 mV vs NHE which is likely the result of oxides on the CNTs themselves. Notably, none of the signals observed with CNTs alone shows any peak splitting, consistent with very rapid electron transfer taking place at the electrode surface. The incorporation of WT EndoIII (75 µM stock) into the film suppresses the surface signals and results in a new, reversible signal centered around 130 mV vs NHE. In contrast to the unsplit surface signals, the EndoIII signal is split by ~85 mV, with reductive and oxidative peaks around 75 and 160 mV vs NHE, respectively. The much greater capacitance in the EndoIII CV shown here relative to that in Figure 1 is due to the presence of 3 additional CNT/protein layers; this made the CNT alone/EndoIII signals more easily compared, but the peaks were broadened and less readily quantified. CVs were taken at a scan rate of 100 mV/s. Storage buffer consisted of 5 mM sodium phosphate, pH 7.5, 150 mM NaCl, and 0.5 mM EDTA.

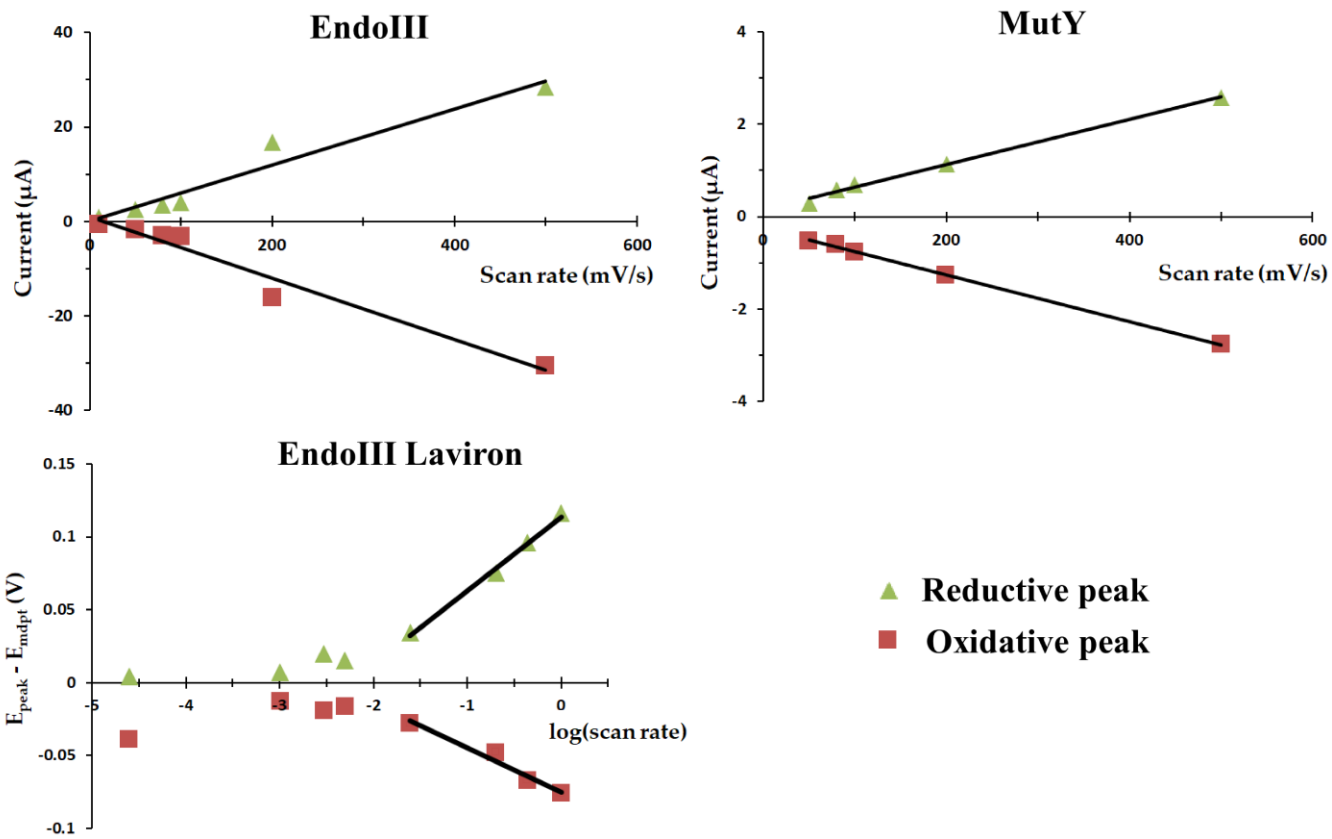


Figure 2.3 Scan rate dependence of the CV current for EndoIII and MutY, and peak splitting dependence on the natural logarithm of scan rate for EndoIII. (Top) Both EndoIII and MutY currents showed a linear dependence on the scan rate, confirming adsorption to the electrode surface. (Bottom) The large EndoIII signals made it possible to estimate electron transfer rates (k_{ET}) and coefficients (α) by fitting a line to the log(scan rate) dependence of the peak splitting at very high scan rates according to Laviron's method for diffusionless systems; small signal sizes made this analysis impossible for MutY.

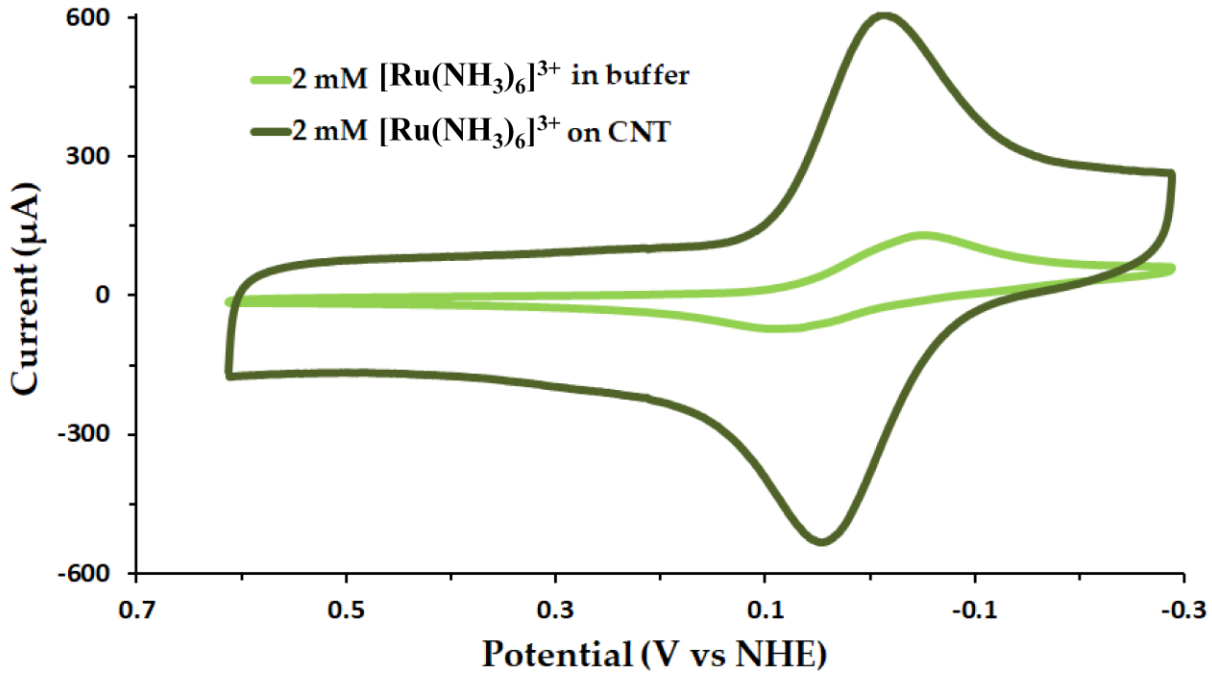


Figure 2.4 $[Ru(NH_3)_6]Cl_3$ electrochemistry on a PGE electrode. 2 mM $[Ru(NH_3)_6]^{3+}$ in EndoIII storage buffer was used first to determine the true electroactive area of the electrode and second as an independent confirmation of redox potential measurements with proteins. CNTs markedly amplified the signal size and decreased the peak splitting, but had no effect on the midpoint potential.

WT EndoIII and MutY Electrochemistry in the Presence of DNA

Having found thin films on PGE to facilitate the direct electrochemistry of free [4Fe4S] proteins, we proceeded to see if the DNA-directed potential shift observed on HOPG could be replicated. With this aim in mind, a solution containing EndoIII and 20-mer duplex DNA in a 1:1 ratio in storage buffer was prepared, incubated on ice for 30 minutes and dried to form a thin film on the electrode surface. Inclusion of CNTs in the film resulted in a noisy, widely split (> 100mV) signal that had the same potential as DNA-free protein, indicating that the protein was not DNA-bound. Thus, CNTs were excluded from subsequent experiments with DNA present. Under these conditions, EndoIII showed a very small signal by CV that was much more readily visualized and quantified using square wave voltammetry (SQWV) (Figure 5, Table 1). The potential as measured by SQWV was 64 ± 8 mV, representing a negative shift of ~60 mV from the DNA-free value of 130 ± 8 mV. This result is in overall agreement with earlier results supporting the stabilization of the oxidized $[4Fe4S]^{3+}$ form upon DNA binding, but the shift seen here is markedly smaller (*8*). This difference can likely be attributed to less than 100% of the protein being DNA bound; free protein likely contributed to the observed peak, leading to an underestimate of the shift. This heterogeneity could be further complicated by the expected random orientation of DNA and protein at the electrode. Indeed, this is a fundamental difference between the HOPG and PGE setups: the signal on HOPG was DNA-mediated, while that on PGE was not.

Electrochemistry of MutY thin films prepared in the presence and absence of DNA (a 30-mer duplex, in this case) gave similar results to EndoIII. In the absence of DNA and CNTs, the SQWV peak was too small to measure the potential with confidence, but it was in the same region as EndoIII (~130 mV vs NHE) (Figure 5); when DNA was present, the potential shifted to

85 ± 3 mV vs NHE (Figure 5, Table 1). In an effort to obtain larger signal sizes by increasing the proportion of DNA-bound protein on the surface, the experiment was repeated with DNA containing an FA:oxoG substrate trap, which mimics the A:oxoG target of MutY but inhibits N-glycosidic bond cleavage due to the electron withdrawing effect of the fluorine at the sugar 2′ position (*34*). No significant differences were observed, although the potential of 80 ± 6 mV indicated a similar DNA-bound signal, suggesting that any increase in DNA binding affinity afforded by the substrate trap was not sufficient to overcome the combined effects of DNA-free proteins on the surface and surface passivation by the DNA itself. In considering these results, it is important to note that, while a shift could be observed for both proteins, surface passivation by non-tethered DNA blocked a portion of the available electroactive sites and made the signals small and noisy. Taken together, these results demonstrate the limitations of the PGE surface, but, importantly, they confirm that large potential shifts can be observed on the PGE electrode even under suboptimal conditions.

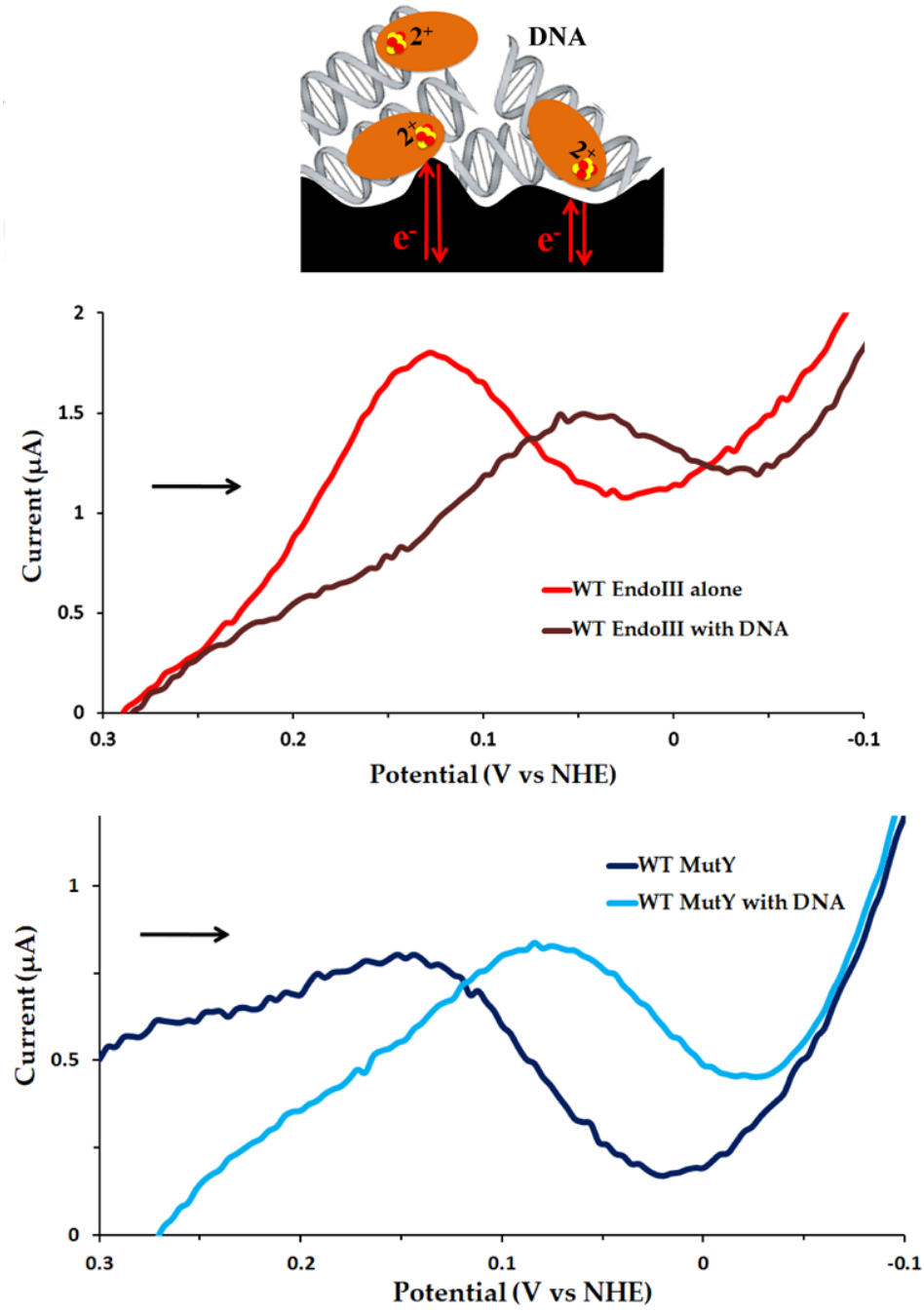


Figure 2.5 EndoIII and MutY thin-film voltammetry on a PGE electrode in the presence of DNA. By SQWV, the presence of DNA resulted in a -65 mV shift in the potential of 75 μM EndoIII (center), with a similar result for 25 μM MutY (bottom). Signals were very small due to surface passivation, but they were still readily apparent by SQWV. Unfortunately, CNTs led to inconsistent and unstable signals, likely interfering with DNA binding, so they could not be used to enhance the signals. Thin films were formed from several layers of a pre-mixed 1:1 protein/DNA solution, and were capped with Nafion. SQWV scans were taken at a frequency of 15 Hz with 0.025 V amplitude, and scans were from positive to negative potentials (indicated by the arrow).

Table 2.1 Potentials (versus NHE) of SQWV reductive sweeps for WT and mutant EndoIII and WT MutY in the presence and absence of DNA.

[a] Enzyme	[b] E_{SQWV} (mV), - DNA	[b] E_{SQWV} (mV), + DNA
WT EndoIII	130 ± 8	[c] 64 ± 8
WT EndoIII + poly-L glutamate	110 ± 9	-
EndoIII Y82A	120 ± 14	-
EndoIII E200K	130 ± 7	-
EndoIII Y205H	125 ± 11	-
EndoIII K208E	141 ± 12	-
WT MutY	[d] > 130 mV	[c] 85 ± 3 [c, e] 80 ± 6

a. When used, CNTs were added from a 0.25 mg/mL stock, and all experiments used protein storage buffer as the supporting electrolyte (20 mM sodium phosphate, 0.5 mM EDTA, 150 mM NaCl, pH 7.5 for EndoIII, 20 mM sodium phosphate, 1 mM EDTA, 150 mM NaCl, pH 7.6 for MutY).
b. Error is the standard deviation of the mean for 3 or more separate experiments.
c. No CNTs present.
d. The MutY SQWV peaks (Figure 2) were too small and noisy to measure the potential with confidence.
e. OG:FA MutY substrate trap DNA

WT EndoIII Electrochemistry in the Presence of Poly-L Glutamate

We next sought to determine if DNA is unique in its ability to shift the potential of the [4Fe4S] cluster by forming thin films with poly-L glutamate (MW 50-100 kDa). Due to the small size of MutY signals, EndoIII was used for this purpose. Like DNA, poly-L glutamate is polyanionic, but it is not a specific target for EndoIII binding. In these experiments, poly-L glutamate was preincubated with EndoIII in a 1:1 ratio and added to thin films just as with DNA. Unlike with DNA, the measured potentials and peak shapes were indistinguishable whether or not CNTs were included. Interestingly, the presence of poly-L glutamate did not result in a significant potential shift by CV, with the reductive and oxidative peaks centered at 81 ± 1 mV and 143 ± 1 mV, respectively (Figure 3). At 110 ± 9 mV, the SQWV potential is, accounting for error, about 10 mV lower than that of free EndoIII, but is still substantially more positive than the DNA-bound potential (Table 1). The CV values are within error of that for free WT EndoIII (Figure 6, Table 1), and, given that EndoIII does not specifically bind to poly-L glutamate, this result should not be too surprising. It is likely that the [4Fe4S] cluster is not exposed to the negative charges of unbound poly-L glutamate to the extent that it would be when the protein is tightly bound to DNA, so no significant shift can be observed. This is also in line with the insusceptibility of free EndoIII to oxidation or reduction originally noted by Cunningham (*2*). Overall, these results indicate that the potential of the EndoIII [4Fe4S] cluster can be appreciably altered only when the protein binds directly to a polyanion, such that the cluster experiences the full effect of the negative charges. Thus, DNA appears uniquely able to affect the potential of the cluster both by its charges and by its function as a binding substrate for these types of [4Fe4S] proteins.

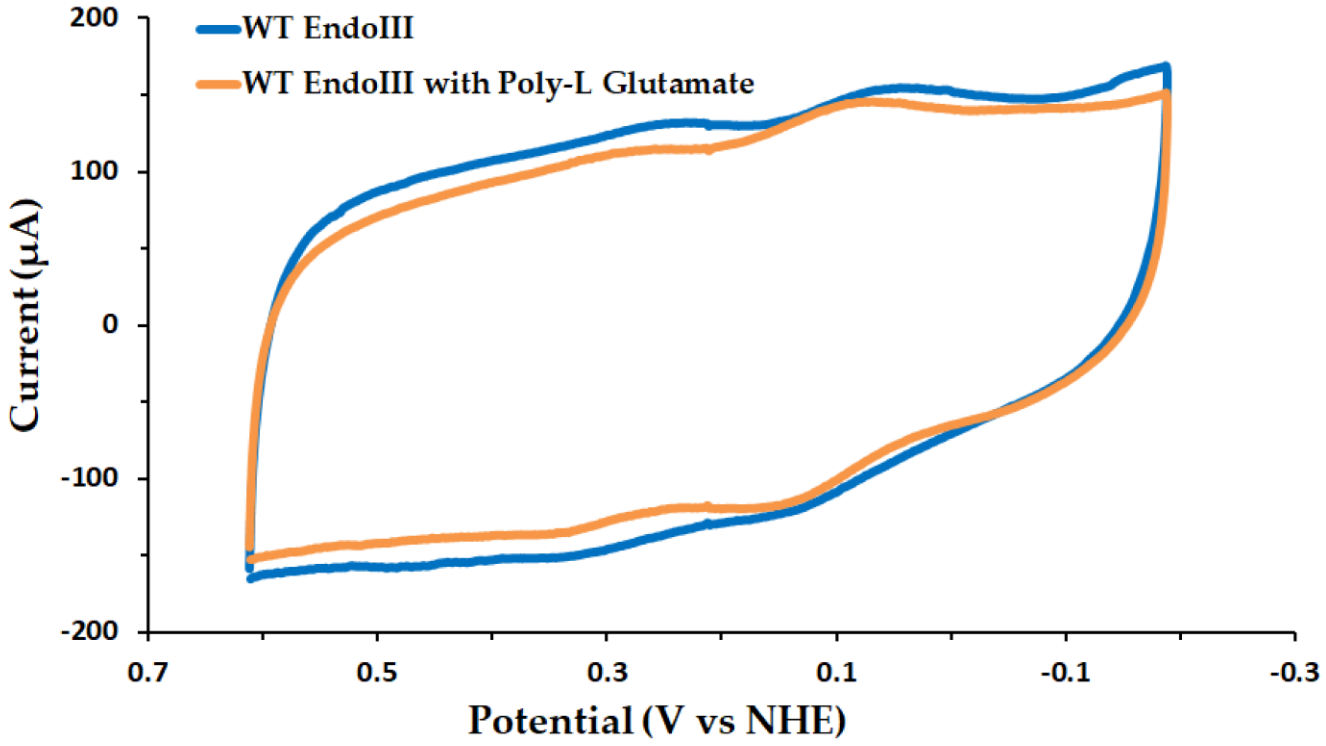


Figure 2.6 CV of WT EndoIII alone and in the presence of poly-L glutamate. In these experiments, poly-L glutamate was pre-incubated with EndoIII (~20 minutes, final EndoIII concentration 75 μM) prior to thin film formation in the presence of CNTs. Unlike DNA, poly-L glutamate caused no significant potential shift even at 6 mM glutamate. CVs were taken at a scan rate of 100 mV/s; DPVs were taken at an amplitude of 0.05 V with a 0.5 s pulse period.

Direct Electrochemistry of EndoIII Mutants

Since EndoIII was observed to undergo a significant potential shift upon DNA binding, but not in the presence of the non-substrate poly-L glutamate, we reasonably assumed that only charges in the immediate vicinity of the cluster could shift the potential. Following this line of reasoning, we sought to determine if altering the charge on amino acid residues near the cluster might have a similarly large effect on the potential. Mutation of nearby residues to give a net positive or negative charge was of further interest because it could, in principle, shift the cluster potential in both positive and negative directions, unlike the unidirectional shift associated with DNA binding.

In order to investigate these possibilities, the following EndoIII point mutants were prepared (*18*): E200K, Y205H, and K208E. All of these residues are ~5 Å from the cluster, and these mutations span nearly the full range of possible single-charge alterations, going from a single negatively charged residue to a positively charged one (E200K), a neutral residue to a positive residue (Y205H), and a positive residue to a negative one (K208E). Positively charged mutants would be expected to be more repulsive to the $[4Fe4S]^{3+}$ state, favoring reduction, while negatively charged mutants would be more attractive, stabilizing the oxidized form of the protein. While shifts might be expected, the precise extent cannot be readily predicted; a study on outer-sphere effects in HiPIP [4Fe4S] protein resulted in shifts as large as 150 mV, while surface mutations in certain ferredoxins led to no changes in potential (*35*, *36*). As a further control, we included the mutant Y82A, which is known to be deficient in DNA CT on DNA-modified gold electrodes, but which was not expected to differ from WT EndoIII in the absence of DNA (*18*).

As expected, the Y82A SQWV potential fell within error of WT at 120 ± 14 mV vs NHE; furthermore, the signal was of comparable size to WT, confirming that thus mutation is only defective at DNA-mediated signaling. Surprisingly, all of the charge-altered mutants exhibited DNA-free potentials on PGE thin films within error of WT (Figure 7, Table 1). However, the standard deviation was relatively high, with the largest value (for WT) around 20 mV. Since conclusions are limited by the largest error among the species investigated, these data are consistent with two possibilities: either there is no change between WT and any of the mutants, or a small shift on the order of 10 mV is present. In either case, it is clear that the effect of changing a single charge, even one in very close proximity to the cluster, is much less dramatic than that of DNA binding.

To confirm the accuracy of potential measurements, thin films formed with 2 mM $[Ru(NH_3)_6]^{3+}$ with and without EndoIII were also examined (Figure 4). The $[Ru(NH_3)_6]^{3+}$ midpoint potential was consistently ~10 mV vs NHE in the presence and absence of CNTs and Nafion, as well as in the presence and absence of EndoIII. Just as with EndoIII, CNTs did markedly sharpen the peaks and facilitate larger signals, independently confirming their effect on species in a thin film. Overall, these controls verified that the measured potentials were not affected by CNTs or other thin film components, indicating that the observed lack of variation between mutants was due to properties of the proteins themselves. Overall, direct electrochemistry of these EndoIII mutants confirm that DNA binding is a dominating effect relative to single charge reversals in the amino acid sequence (*18*).

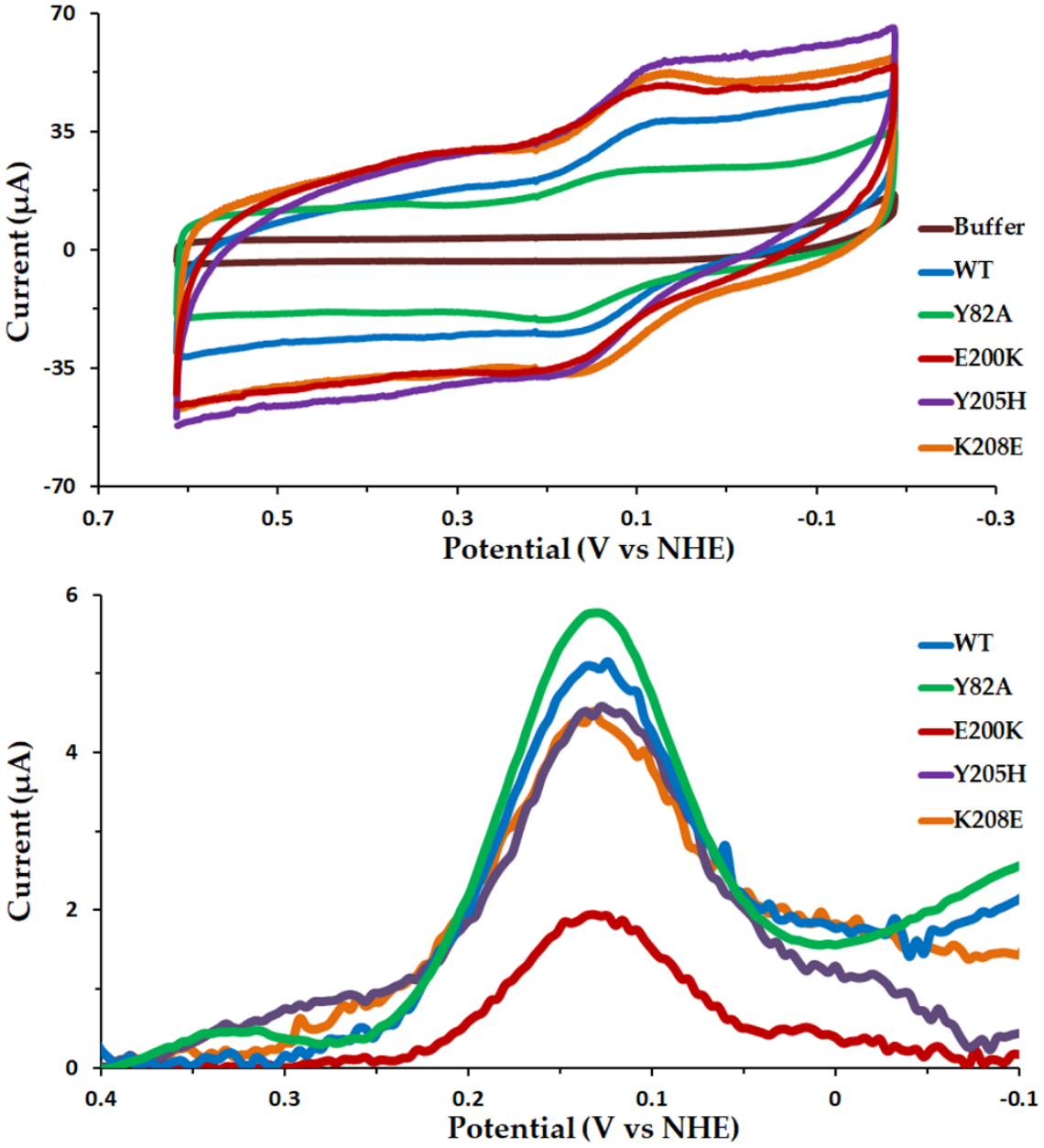


Figure 2.7 Thin-film voltammetry of WT EndoIII and the mutants Y82A, E200K, Y205H, and K208E. 75 μM WT and mutant EndoIII in protein/CNT/Nafion thin films had nearly identical potentials, with CV peaks centered around 125 mV versus NHE (top); the similarity is even more apparent by DPV (bottom). Small shifts within the measurement error (10-15 mV) may still occur in the three charge-altered mutants, but these nonetheless pale in comparison to the effect of DNA. All CVs shown were taken at 100 mV/s, while DPVs were taken at an amplitude of 0.05 V with a 0.5 s pulse period.

Discussion

In this study, we used thin film voltammetry on a PGE electrode to measure the potentials of EndoIII and MutY in the presence and absence of DNA, and, in the case of EndoIII, in the presence of poly-L glutamate and with point mutations altering the charged environment near the cluster. Table 1 summarizes all of the results. Notably, the potential shift observed upon DNA binding was smaller than previously reported (*8*), with the SQWV potentials of DNA-free EndoIII going from 250 ± 30 mV on HOPG to 130 ± 8 mV on PGE. This difference was certainly in large part the result of less than 100% DNA binding on PGE, but the distinct electrode environments on PGE and HOPG may also have played a role (*8*, *19-20*). Indeed, the presence of negatively charged surface oxides on PGE, but not HOPG, would be expected to lower the potential of adsorbed proteins to some extent (*29*). Even with a lower DNA-free potential, DNA-bound potentials still dropped by 65 mV on PGE, which supports the assertion that DNA binding has a prominent, although not necessarily additive, effect which cannot be duplicated by non-substrate polyanionic molecules such as poly-L glutamate.

The absence of a significant potential shift in the presence of poly-L glutamate does at first appear to be in conflict with the results of the Todorovic group, wherein the redox activation of EndoIII was reported on a mercaptoundecanoic acid monolayer assembled on a gold electrode in the absence of DNA (*16*). However, this effect was attributed to tight binding to the carboxylic acid film; in contrast, the solvated poly-L glutamate used here would present a very different environment than a thin film of small molecules, and it is unlikely that EndoIII was readily able to bind this non-substrate polymer. While it is reasonable that multiple negative charges could shift the potential, they can only do so if EndoIII is able to bind with some specificity.

The results obtained here for DNA demonstrate that large potential shifts are observable on PGE even under conditions where full DNA binding is unlikely. In contrast to the conditions required to study DNA-binding, the DNA-free EndoIII mutants were studied under conditions that generated very large and readily quantifiable signals; thus, the lack of any apparent change in the potential among these mutants indicates that individual charge alterations near the cluster do not have a significant effect. However, the error in these experiments was relatively high, leaving two interpretations open. First, there may genuinely be no shift. In this case, the conclusion would be that a single charge alteration is insufficient by itself to disrupt the local environment, even for residues ~5 Å from the cluster. Alternatively, because the measurement error was around 20 mV for WT, shifts on the order of 10 mV may have been present but undetectable. Such small shifts would still pale in comparison to those associated with DNA binding, which are larger as independently measured on a PGE electrode (-60 mV for the EndoIII $2^+/3^+$ couple) and on an HOPG electrode (-200 mV; *8*). Assuming that the charge mutants of EndoIII do have altered potentials and that the shift is at the upper limit of our error (~10 mV), the effect of DNA binding is at minimum 6 times greater than the effect of a single charge alteration. In either case, it should be noted that, despite the small signals and adverse conditions resulting from DNA in the film, a substantial shift in potential upon DNA binding was still detectable on PGE, while no obvious shift in potential relative to WT was observed for any of the mutants, even though their potentials were determined from very large DNA-free signals under ideal circumstances.

A similar result obtained with the same mutants on DNA-modified gold electrodes supports the observed lack of potential shifts. In those experiments, the similarity in potential was attributed to the presence of DNA being a dominant effect, with only an increase in current

correlated with protein folding stability occurring for K208E and Y205H (*18*). The smaller accessible window and poor to nonexistent DNA-free signal on gold electrodes prevented the experiments described here from being carried out, but taken along with our results, it appears that the mutants have similar potentials to WT both on and off of DNA.

In summary, these results reveal that, under the same experimental conditions, DNA binding is the dominant factor in tuning the redox potential of the [4Fe4S] cluster. This is supported by the similarity of the DNA-bound midpoint potentials of various repair enzymes (*6*), which all fall within a range of 60-85 mV vs NHE. In the context of DNA repair, we have previously proposed a model in which these proteins use single-electron transfers to signal to each other across the genome in the search for damage (*15*). In order for this model to work, the electron transfers must be reversible, necessitating that the proteins involved have approximately equal DNA-bound midpoint potentials. If this were not the case, proteins of lower reduction potential would remain bound at the expense of those of higher potential, and the search process would only be able to proceed by diffusion processes that are too slow to fully account for the time scale of DNA repair (*15*). This model depends on the large effect of DNA binding to bring potentials into a relevant regime, and our results verify that differences in the protein environment are unlikely to shift the potential sufficiently to affect the reversibility of inter-protein signaling on DNA. Overall, the similar DNA-bound potentials among diverse proteins facilitate not only the DNA damage search, but could also make long-range communication between diverse pathways possible.

References

1) Kim, Y.J.; Wilson, D.M. III. Overview of Base Excision Repair Biochemistry. *Curr. Mol. Pharmacol*. **2012**, *5*, 3 – 13.

2) Cunningham, R.P.; Asahara, H.; Bank, J.F.; Scholes, C.P.; Salerno, J.C.; Surerus, K.; Munck, E.; McCracken, J.; Peisach, J.; Emptage, M.H. Endonuclease III is an iron-sulfur protein. *Biochemistry* **1989**, *28*, 4450 – 4455.

3) Porello, S.L.; Cannon, M.J.; David, S.S. A Substrate Recognition Role for the $[4Fe4S]^{2+}$ Cluster of the DNA Repair Glycosylase MutY. *Biochemistry* **1998**, *37*, 6465 – 6475.

4) David, S.S.; O'Shea, V.L.; Kundu, S. Base-excision repair of oxidative DNA damage. *Nature* **2007**, *447*, 941 – 950.

5) Woods, R.D.; O'Shea, V.L.; Chu, A.; Cao, S.; Richards, J.L.; Horvath, M.P.; David, S.S. Structure and stereochemistry of the base excision repair glycosylase MutY reveal a mechanism similar to retaining glycosylases. *Nucleic Acids Research* **2016**, *44*, 801 – 810.

6) Boal, A.K.; Yavin, E.; Lukianova, O.A.; O'Shea, V.L.; David, S.S.; Barton, J.K. DNA-Bound Redox Activity of DNA Repair Glycosylases Containing [4Fe4S] Clusters. *Biochemistry* **2005**, *44*, 8397 – 8407.

7) Arnold, A.R.; Grodick, M.A.; Barton, J.K. DNA Charge Transport: from Chemical Principles to the Cell. *Cell Chemical Biology* **2016**, *23*, 183 – 197.

8) Gorodetsky, A.A.; Boal, A.K.; Barton, J.K. Direct Electrochemistry of Endonuclease III in the Presence and Absence of DNA. *J. Am. Chem. Soc.* **2006**, *128*, 12082 – 12083.

9) Thayer, M.M.; Ahern, H.; Xing, D.; Cunningham, R.P.; Tainer, J.A. Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *The EMBO J.* **1995**, *14*, 4108 – 4120.

10) Fromme, J.C.; Verdine, G.L. Structure of a trapped endonuclease III-DNA covalent intermediate. *The EMBO J.* **2003**, *22*, 3461 – 3471.

11) Guan, Y.; Manuel, R.C.; Arvai, A.S.; Parikh, S.S.; Mol, C.D.; Miller, J.H.; Lloyd, R.S.; Tainer, J.A MutY catalytic core, mutant, and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nat. Struct. Biol.* **1998**, *5*, 1058 – 1064.

12) Fromme, J.C.; Banerjee, A.; Huang, S.J.; Verdine, G.L. Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase. *Nature* **2004**, *427*, 652 – 656.

13) Stephens, P.J.; Jollie, D.R.; Warshel, A. Protein Control of Redox Potentials of Iron-Sulfur Proteins. *Chem. Rev*. **1996**, *96*, 2491 – 2513.

14) Ren, B.; Duan, X.; Ding, H. Redox control of the DNA damage-inducible protein DinG helicase activity via its iron-sulfur cluster. *J. Biol. Chem.* **2009**, *284*, 4829 – 4835.

15) Boal, A.K.; Genereux, J.C.; Sontz, P.A.; Gralnick, J.A.; Newman, D.K.; Barton, J.K. Redox signaling between DNA repair proteins for efficient lesion detection. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15237 – 15242.

16) Moe, E.; Sezer, M.; Hildebrandt, P.; Todorovic, S. Surface enhanced vibrational spectroscopic evidence for an alternative DNA-independent redox activation of endonuclease III. *ChemComm.* **2015**, *51*, 3255 – 3257.

17) Dey, A.; Jenney, F.E., Jr.; Adams, M.W.W; Babini, E.; Takahashi, Y.; Fukuyama, K.; Hodgson, K.O.; Hedman, B.; Solomon, E.I. Solvent Tuning of Electrochemical Potentials in the Active Sites of HiPIP Versus Ferredoxin. *Science.* **2007**, *318*, 1464 – 1468.

18) Pheeney, C.G.; Arnold, A.R; Grodick, M.A.; Barton, J.K. Multiplexed Electrochemistry of DNA-Bound Metalloproteins. *J. Amer. Chem. Soc.* **2013**, *135*, 11869 – 11878.

19) Banks, C.E.; Compton, R.G. New electrodes for old: from carbon nanotubes to edge plane pyrolytic graphite. *Analyst*. **2006,** *131*, 15 – 21.

20) Armstrong, F.A.; Bond, A.M.; Hill, A.O.; Oliver, B.N.; Psalti, I.S.M. Electrochemistry of Cytochrome c, Plastocyanin, and Ferredoxin at Edge- and Basal-Plane Graphite Electrodes Interpreted via a Model Based on Electron Transfer at Electroactive Sites of Microscopic Dimensions in Size. *J. Am. Chem. Soc.* **1989**, *111*, 9185 – 9189.

21) Yin, Y.; Lu, Y.; Wu P.; Cai, C. Direct Electrochemistry of Redox Proteins and Enzymes Promoted by Carbon Nanotubes. *Sensors* **2005**, *5*, 220 – 234.

22) Léger, C.; Elliott, S.J.; Hoke, K.R.; Jeuken, L.J.C.; Jones, A.K.; Armstrong, F.A. Enzyme Electrokinetics: Using Protein Film Voltammetry To Investigate Redox Enzymes and Their Mechanisms. *Biochemistry* **2003**, *42*, 8653 – 8662.

23) Blanford, C.F.; Armstrong, F.A. The pyrolytic graphite surface as an enzyme substrate: microscopic and spectroscopic studies. *J. Solid State Electrochemistry* **2006**, *10*, 826 – 832.

24) Baffert, C.; Sybirna, K.; Ezanno, P.; Lautier, T.; Hajj, V.; Meynial-Salles, I.; Soucaille, P.; Bottin, H.; Léger, C. Covalent Attachment of FeFe Hydrogenases to Carbon Electrodes for Direct Electron Transfer. *Anal. Chem.* **2012**, *84*, 7999 – 8005.

25) Chmiel, N.H.; Golinelli, M.P.; Francis, A.W.; David, S.S. Efficient recognition of substrates and substrate analogs by the adenine glycosylase MutY requires the C-terminal domain. *Nucleic Acids Res*. **2001**, *29*, 553 – 64.

26) Kissinger, P.T.; Heineman, W.R. (Eds.) Laboratory Techniques in Electroanalytical Chemistry (2nd ed.) **1996** Marcel Dekker

27) Wang, Y.; Limon-Peterson, J.G.; Compton, R.G. Measurement of the diffusion coefficients of $[Ru(NH_3)_6]^{3+}$ and $[Ru(NH_3)_6]^{2+}$ in aqueous solution using microelectrode double potential step potentiometry. *J. Electroanalytical Chem.* **2011**, *652*, 13 – 17.

28) Greeley, R.S.; Smith, W.T., Jr.; Stoughton, R.W.; Lietzke, M.H. Electromotive Forces in Aqueous Solutions at Elevated Temperatures. I. The Standard Potential of the Silver-Silver Chloride Electrode. *J. Phys. Chem.* **1960**, *64*, 652 – 657.

29) McCreery, R.L. Advanced Carbon Electrode Materials for Molecular Electrochemistry. *Chem. Rev.* **2008**, *108*, 2646 – 2687.

30) Blurton, K.F. An Electrochemical Investigation of Graphite Surfaces. *Electrochimica Acta* **1973**, *18*, 869 – 875.

31) Gooding, J.J. Nanostructuring electrodes with carbon nanotubes: A review on electrochemistry and applications for sensing. *Electrochimica Acta* **2005**, *50*, 3049 – 3060.

32) Breton, J.L.; Duff, J.L.C.; Butt, J.N.; Armstrong, F.A.; George, S.J.; Pétillot, Y.; Forest, E.; Schäfer, G.; Thomson, A.J. Identification of the iron-sulfur clusters in a ferredoxin from the archaeon Sulfolobus acidocaldarius. Evidence for a reduced [3Fe-4S] cluster with pH-dependent electronic properties. *Eur. J. Biochem.***1995**, *233*, 937 – 946.

33) Laviron, E. General expression of the linear potential sweep voltammogram in the case of diffusionless electrochemical systems. *J. Electroanaytical Chem.* **1979**, *101*, 19 – 28.

34) Lee, S.; Verdine, G.L. Atomic substitution reveals the structural basis for substrate adenine recognition and removal by adenine DNA glycosylase. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18497 – 18502.

35) Low, D.W.; Hill, M.G. Backbone-Engineered High-Potential Iron Proteins: Effects of Active-Site Hydrogen Bonding on Reduction Potential. *J. Amer. Chem. Soc.* **2000**, *122*, 11039 – 11040.

36) Shen, B.; Jollies, D.R.; Stout, C.D.; Diller, T.C.; Armstrong, F.A.; Gorst, C.M.; La Mar, G.N.; Stephens, P.J.; Burgess, B.K. *Azotobacter vinelandii* ferredoxin I. Alteration of individual surface charges and the $[4Fe\text{-}4S]^{2+/+}$ cluster reduction potential. *J. Biol. Chem.* **1994**, *269*, 8564 – 8575.

Chapter 3

S K-edge XAS Studies of the Effect of DNA Binding on the [4Fe4S] Site in Endo III and MutY

Adapted from: Ha, Y.; Arnold, A.R.; Nuñez, N.N.; Bartels, P.L.; Zhou, A.; David, S.S.; Barton, J.K.; Hedman, B.; Hodgson, K.O.; Solomon, E.I. Sulfur K-Edge XAS Studies of the Effect of DNA Binding on the [Fe_4S_4] Site in EndoIII and MutY. *J. Am. Chem. Soc.* **2017**, *139*, 11434 – 11442.

Y. Ha performed all XAS experiments and DFT analysis. A. Arnold, N. Nuñez, and A. Zhou purified DNA and proteins. P. Bartels performed UV-visible, CD, and EPR spectroscopy.

Introduction

Iron-sulfur clusters occur in a wide range of proteins with roles in electron transfer, catalysis, and regulation (*1*). Those involved in electron transfer are the rubredoxins and ferredoxins. Rubredoxin contains a mononuclear iron center with thiolate ligands, while ferredoxins contain multinuclear iron-sulfide clusters which includes [2Fe2S], [3Fe4S], and [4Fe4S] sites (*2*). The [Fe_4S_4] proteins can be further divided into two categories: the low-potential ferredoxins (Fd), which utilize the $[4Fe4S]^{2+/1+}$ redox couple with a potential as low as -600 mV, and high potential iron proteins (HiPIP), which have an $[4Fe4S]^{3+/2+}$ redox couple with a reduction potential as high as +350 mV (*1*). From X-ray crystallography (*3*), Extended X-Ray Absorption Fine Structure (EXAFS) (*4*) and resonance Raman (*5*) spectroscopic results, the $[4Fe4S]^{2+}$ sites in the Fds and HiPIPs are almost identical. They have similar bond distances and angles, as well as similar vibrational modes. The major difference is in their protein environments. The [4Fe4S] site in Fd is solvent exposed while this cluster in HiPIP is buried (*6*). Lyophilization of Fd and unfolding of HiPIP lead to significant changes in their S K-edge XAS spectra, showing that the H-bonds from solvent water change the Fe-S bond covalency (*7*). The higher covalency in HiPIP stabilizes the oxidized state over the reduced state, and this significantly contributes to their [4Fe4S] clusters utilizing different redox couples from Fd (*1*). The $[4Fe4S]^{2+}$ in Fd is less covalent and activated toward reduction, while the cluster in HiPIP is more covalent, and thus tuned toward oxidation.

Endonuclease III (EndoIII) and MutY are DNA glycosylases present in bacteria to humans that excise oxidized bases or their mispaired partners as initiating events in the Base Excision Repair (BER) pathway. Specifically, Endo III removes oxidized pyrimidines while MutY removes adenine from A:oxoG mispairs (where oxoG = 8-oxo-7,8-dihydroguanine) (*8, 9*).

The importance of these enzymes is highlighted by the correlation of inherited defects in human MutY (MUTYH) and early onset colorectal cancer, referred to as MUTYH-associated polyposis (MAP) (*10*). EndoIII and MutY both contain a $[4Fe4S]^{2+}$ cluster, which is redox inactive in solution in the absence of DNA (*11*). The $[4Fe4S]^{2+}$ cluster in these enzymes has been shown to be essential for substrate binding and catalysis, but, at least in the case of MutY, it is not required for global structural integrity of the enzyme (*9*). However, electrochemical experiments carried out with EndoIII and MutY on DNA-modified gold electrodes showed reversible redox signals at potentials ranging from 60 to 95 mV versus NHE, suggesting that DNA binding is able to activate these proteins for redox activity (*8*). A role for DNA in activating the cluster for redox activity was confirmed by electrochemistry on highly oriented pyrolytic graphite (HOPG) electrodes in the presence and absence of DNA; these experiments revealed that upon binding to DNA, the reduction potential of the $[4Fe4S]^{3+/2+}$ couple in EndoIII shifted ~ -200 mV into the physiological range and corresponding to a significantly increased DNA binding affinity of the oxidized form of the protein (*12, 13*). From electrochemical and *in vivo* experiments, a model has been proposed in which long-distance DNA-mediated electron transfer between two [4Fe4S] proteins with similar DNA-bound redox potentials facilitates the

search for damage across a vase genome (*8, 14-15*). If there is no DNA damage between the two binding sites, the charge transfer can proceed efficiently, and one of the DNA-bound proteins is reduced, thus its affinity for DNA is decreased. This protein can then dissociate and diffuse to another region of the genome. However, if there is a mismatch or lesion that disrupts the π-stacking of DNA, charge transfer will be attenuated and both proteins would remain bound to the DNA in the vicinity of the lesion. While the CT signaling model has been strongly supported in numerous studies, the cause of the large DNA-induced potential shift that makes such a damage search possible has remained puzzling. Electrostatic effects remain the most likely explanation, given that no significant structural changes are apparent between the free and DNA-bound forms of EndoIII and MutY. Unfortunately, the electrochemical techniques used in earlier work cannot reveal the fundamental molecular-scale changes involved, making an alternative approach necessary to elucidate these details.

S K-edge X-ray Absorption Spectroscopy (XAS) has been developed and applied to experimentally determine the covalency (α^2) of sulfur-metal bonds (i.e., the amount of S 3p character mixed into the metal 3d antibonding (Ψ^*) valence orbitals: $\Psi^* = \sqrt{1-\alpha^2}M_{3d} - \alpha S_{3p}$) (*16*). The S 1s orbital is localized on the S atom and the S 1s → S 3p transition is electric dipole allowed, thus the intensity of the 1s →Ψ^* pre-edge transition reflects the covalency of this bond. The energy of the pre-edge transition reflects the energy of the unoccupied or partially occupied Fe d-orbitals, which depends on the effective nuclear charge (Z_{eff}) of the metal and the nature of the ligand field (*16*). The pre-edge transition energy also depends on the charge of the S; sulfide

has a lower Z_{eff} than thiolate, thus the sulfide donor orbitals are at higher energy than those of the thiolate, and their associated S 1s $\rightarrow \Psi^*$ pre-edge transitions are at lower energy (*16*).

The intensity of a pre-edge peak D_0 is given by

$$D_0 = \sum c\alpha^2 \left| \langle S_{1s} | r | S_{3p} \rangle \right|^2 = \frac{\alpha^2 h}{3N} I_s \qquad (1)$$

where c is a constant, r is the dipole operator, α^2 is the bond covalency (i.e., sulfur p character mixed into the metal d orbitals), *N* is the total number of sulfurs bound to the metal, *h* is the total number of d electron holes, and I_s is the S 1s $\rightarrow$ 3p electric dipole integral, which is also dependent on the Z_{eff} of the S ligand and has been experimentally determined for thiolate and sulfide ligands in previous work (*17*). Our previous XAS studies on $[Fe_4S_4]$ sites showed that the contributions of thiolate and sulfide can be distinguished at the pre-edge, with the μ_3-sulfide pre-edge ~0.7 eV lower than the thiolate (*6, 18*). Quantitatively, for the [4(RS)4Fe4S] clusters, there are nine α holes and nine β holes (i.e., unoccupied valence orbitals), which have mainly Fe 3d and S 3p character (*6*); thus, the maximum covalency possible is 1800%. In particular, one unit of sulfide intensity corresponds to 30.6% S character per Fe-S bond, while 1 unit of thiolate corresponds to 70.8% S character in unoccupied valence orbitals. This difference reflects the different number of bonds to Fe (3 from sulfides and 1 from thiolate) and the difference in dipole integral due to the Z_{eff} of sulfide relative to thiolate S (6.54 vs 8.47) (*17*). Note that only the sulfur atoms bound to the iron contribute to the pre-edge feature in the XAS spectra, while all sulfur atoms including the free Cys and Met residues contribute to the edge. Thus, the pre-edge

intensity directly reflects the covalency of the thiolate-Fe and bridging sulfide-Fe bonds, but the edge normalization must be corrected for the total number of sulfurs in the protein.

The total pre-edge intensity, and therefore S covalency, in HiPIP is significantly higher than in Fd. The pre-edge intensity in a relevant alkyl thiolate [4Fe4S] model complex is even slightly higher than in HiPIP (*7*). Importantly, there is a direct correlation between the total S covalency and the redox potential of the [4Fe4S] clusters, with an increase of 1% of total S covalency corresponding to ~3.3 mV decrease in redox potential (*7*), which reflects the fact that higher covalency stabilizes the oxidized over the reduced state of the [4Fe4S] cluster. With respect to DNA repair proteins, the power of S K-edge XAS to monitor changes in [4Fe4S] cluster covalency under different conditions and to correlate these changes to redox potential results in an ideal method to elucidate the details of the DNA-induced potential shift at a molecular level.

In the present study, S K-edge XAS is applied to experimentally measure the Fe-S bond covalency of the [4Fe4S] clusters in EndoIII and MutY, both in the absence of and bound to DNA, and with and without solvent water. The effect of DNA binding on the S K-edge intensity and hence covalency is correlated to the reduction potential of the cluster according to a relationship defined in previous XAS studies on [4Fe4S] clusters (*7*) and is in agreement with the electrochemically observed reduction potential decrease upon DNA binding (*12*). The solvent effect is compared to those previously observed for Fd (*7*) and for the HiPIP proteins in this study. This work provides direct molecular evidence for the proposal that the negative charge of bound DNA tunes the potential of [4Fe4S] clusters and shows that this involves a change in the

covalency of the cluster that enables the $[4Fe4S]^{3+/2+}$ redox couple when EndoIII (and MutY) are bound to DNA.

Materials and Methods

Expression and Purification of EndoIII

WT *E. coli* EndoIII was overexpressed in BL21star-(DE3)pLysS cells containing a pET11-ubiquitin-His_6-*nth* construct and purified as detailed previously (*19*), with the exception that the final buffer contained 10% glycerol, rather than 20% glycerol (20 mM sodium phosphate, pH 7.5, 0.5 mM EDTA, 150 mM NaCl, 10% glycerol, Buffer A). EndoIII is less stable in the absence of glycerol; therefore glycerol was not removed until the day of sample preparation. Glycerol was removed from half the volume of protein solution using HiPrep 26/10 desalting column (GE Healthcare) equilibrated with Buffer A lacking glycerol. Next, the protein solutions either containing 10% glycerol or no glycerol were separately concentrated first with 10,000 MWCO (molecular weight cutoff) Amicon Ultra 15 mL centrifugation filter units (Millipore) and then with 10,000 MWCO Amicon Ultra 0.5 mL centrifugation filter units (Millipore) until the protein solutions were very dark colored, to approximately 300 μL each if using an entire protein preparation from 6 L of bacterial culture.

Expression and Purification of MutY

WT MutY was expressed as an N-terminal fusion with MBP (maltose binding protein) to increase the solubility at the concentrations needed for XAS experiments. Of note, the MBP tag has additional Met residues that can complicate XAS analysis. The MBP-MutY protein was overexpressed in BL21 DE3 competent cells and purified as detailed previously (*13*), with the following modification. Pelleted cells from overexpression were re-suspended in resuspension buffer (20 mM sodium phosphate, pH 7.5, 200 mM NaCl, 1mM EDTA, 10 % glycerol, 1 mM PMSF), sonicated on ice and centrifuged at 12,000 RPM for 15 minutes. The supernatant was

saved in a conical tube on ice, and the pellet was re-suspended and underwent a repeat of sonification and centrifugation. The supernatant was batch bound to amylose resin (New England BioLabs) for one hour, poured over a PD10 column and washed with amylose wash buffer (20 mM sodium phosphate, pH 7.5, 200 mM NaCl, 1mM EDTA) and eluted in amylose elutant buffer (20 mM sodium phosphate, pH 7.5, 200 mM NaCl, 1 mM EDTA, 10 mM maltose). The elutant was concentrated using an ultrafiltration cell with a 10,000 MWCO filter with stirring at 4oC. Protein was then diluted 10-fold in heparin buffer A (20 mM sodium phosphate, pH 7.5, 1 mM EDTA, 5% glycerol). The sample was applied to a Pharmacia Hi-trap heparin column on an AKTApurifier FPLC system, and eluted using a 10% linear gradient in heparin buffer A to 100% heparin buffer B (20 mM sodium phosphate, pH 7.5, 1 mM EDTA, 5 % glycerol, and 1 M NaCl). Fractions corresponding to MBP-MutY were combined and concentrated using an ultrafiltration cell with a 10,000 MWCO filter with stirring at 4 ^{o}C, to approximately 10mL. Purity of MBP-MutY samples was confirmed via 12% sodium dodecyl sulfate (SDS) PAGE stained with SYPRO orange and 7.5% acetic acid.

DNA Preparation

DNA strands for EndoIII studies were purchased from Integrated DNA Technologies (a 20-mer mixed sequence strand: 5′-GTGAGCTAACGTGTCAGTAC-3′ and its complement). DNA strands (5 μmol) were resuspended in MilliQ water (200 μL), and purified by ethanol precipitation. The purified strands were resuspended in Buffer A or Buffer A lacking glycerol and quantified based on calculated ε_{260} values for the strands (Integrated DNA Technologies) of 197,800 $M^{-1}cm^{-1}$ for the 20-mer strand and 190,200 $M^{-1}cm^{-1}$ for its complement. Annealing of the strands in either Buffer A or Buffer A without glycerol was accomplished by combining

equimolar amounts of the single-stranded DNAs, heating at 90°C for 5 minutes, and slowly cooling to room temperature.

DNA strands for MutY studies containing oxoG (8-oxo-guanine) or FA (2'-fluoro-adenine) were synthesized at the University of Utah DNA and Peptide Synthesis Core Facility and unmodified strands were ordered from Integrated DNA Technologies. The following DNA duplexes were used:

15-mer: 5′ -GGA GCC AXG AGC TCC-3

15-mer Complement: 3′ -CCT CGG TYC TCG AGG-5′

30-mer: 5′ -CGA TCA TGG AGC CAC XAG CTC CCG TTA CAG-3′

30-mer Complement: 3′ -GCT AGT ACC TCG GTG YTC GAG GGC AAT GTC-5'

X = G or oxoG and Y = C, FA, or A

Oligonucleotides containing the central oxoG or FA were deprotected and cleaved from the column by incubation in NH_4OH, with the addition of 2-mercaptoethanol to oxoG samples to prevent oxidation. The cleaved DNA substrates were dissolved in H_2O, filtered with a 0.2 μm filter, and HPLC purified using a Beckman Gold Nouveau system with a Waters AP1DEAE 8HR column with a 10-100% gradient of 90:10 H2O/acetonitrile with 2 M NH_4Ac. Isolated fractions were lyophilized down and de-salted with a SEP-PAK C18 column, and DNA integrity was confirmed using MALDI-MS. All DNA substrates were stored dried in the -20^{o}C freezer prior to annealing to the complement strand.

XAS Sample Preparation

To prepare the EndoIII XAS samples, concentrated solutions of EndoIII protein were mixed with the DNA duplex at a ratio of 1 mol EndoIII: 20 mol base-pairs DNA. An equivalent volume of buffer was added to EndoIII alone solutions so that identical EndoIII concentrations were obtained both with and without DNA. Mixtures were allowed to incubate on wet ice for 30 minutes to allow for binding before freezing in liquid nitrogen. No glycerol samples were lyophilized. Both samples were placed on dry ice and sent to Stanford University for measurement. To confirm that lyophilized protein was not irreversibly altered by treatment, UV-visible and circular dichroism (CD) spectra were taken of untreated and redissolved samples.

To prepare the MutY XAS samples, purified MBP-MUTY samples were buffer exchanged in concentration buffer, with the final buffer composed of 20 mM sodium phosphate, pH 7.5, 150 mM NaCl, 1mM EDTA, 10% glycerol for samples to remain in buffer, or MBP-MutY concentration buffer lacking glycerol for samples to be lyophilized for XAS. The sample was concentrated to a final concentration of 1.2 mM. 125 uL of the purified MBP-MutY is incubated with 25uL of the 6mM annealed DNA duplex (see DNA contexts below) at 25 ^{o}C for 30 minutes to afford a final concentration of 1mM protein and DNA. Following lyophilization or immediately for the buffer samples, they were snap frozen in liquid nitrogen, stored in the -80 ^{o}C or dry ice, and thawed on ice just prior to XAS experiments. Purity of MBP-MutY samples was confirmed via 12 % SDS page stained with SYPRO orange and 7.5 % acetic acid. The [4Fe4S] cluster loading was determined using the UV-visible absorbance at 410 nm (ε_{410} = 17,000 $M^{-1}cm^{-1}$) and at 280 nm (ε_{280} = 143,240 $M^{-1}cm^{-1}$); samples were typically 65-75% loaded with cluster.

S K-edge XAS

Sulfur K-edge XAS data were measured at the Stanford Synchrotron Radiation Lightsource on the unfocussed 20-pole, 2.0-Tesla wiggler Beam Line 4-3, under SPEAR3 storage ring parameters of 3 GeV and 500 mA. A Ni-coated, flat, bent pre-monochromator mirror was used for harmonic rejection and vertical collimation. A Si(111) double crystal monochromator was used for energy selection. The energy calibration, data reduction, and error analysis follow the methods described in reference (*20*). Solid samples were ground into a fine powder and dispersed as thinly as possible on Kapton tape to minimize potential self-absorption effects. The solution samples were loaded into 50μL teflon cells, with Kapton tape as the back window, and 6 μm-thick, sulfur-free polypropylene film as front window. A shutter was inserted automatically during each monochromator move to minimize photoreduction. The photon energy was calibrated to the maximum of the first pre-edge feature of $Na_2S_2O_3 \bullet 5H_2O$ at 2472.02 eV. At least three scans were measured for each sample to ensure reproducibility. Raw data were calibrated and averaged using MAVE in the EXAFSPAK package (*21*). Using the PySpline program (*22*), the background was removed from all spectra by fitting a second-order polynomial to the pre-edge region and subtracting this from the entire spectrum. Normalization of the data was accomplished by fitting a straight line to the post-edge region and normalizing the edge jump to 1.0 at 2490.0 eV. The error from background subtraction and normalization is less than 3%. Intensities of the pre-edge features were quantified by fitting the data with pseudo-Voigt line shapes with a fixed Lorentzian to Gaussian ratio of 1:1, using the EDG_FIT program (*21*). Pre-edge energies and widths of single peaks were locked based on the previously published results on $[Fe_4S_4]$ models and proteins (*6*). The error from the fitting procedure is less than 3%. The fitted sulfide and thiolate intensities were converted to %S 3p character according to ref *17*.

Note that all the perturbations (lyophilization, mixing with DNA, and the combination) on EndoIII and MutY were done on aliquots of the same protein sample. This means that all the data on each protein presented in this study have the same loading ratio; thus, the differences observed upon DNA binding and lyophilization are independent of the loading.

DFT calculations

DFT calculations with broken symmetry spin polarization were performed using Gaussian 09 (*23*), with the pure functional BP86, and with 6-311G(d) basis sets on Fe and S, and 6-31G(d) basis sets on C and H. This functional and basis set were chosen to be consistent with previous studies (*7*). The α-carbons of the 4 Cys ligands were fixed to their crystal structure positions (*24*) during geometry optimizations. To qualitatively evaluate the electrostatic effect of DNA binding, a point charge was placed at 5 Å from either a sulfide or a thiolate S atom in the [4Fe4S] cluster, and the electronic structure was reoptimized.

Iron Quantification

To measure the extinction coefficient at 410 nm for two EndoIII mutants, Y205H and K208E, we followed an iron quantification protocol described in ref. *25*. Briefly, 100-150 μM EndoIII in storage buffer (20 mM sodium phosphate, pH 7.5, 150 mM NaCl, 0.5 mM EDTA) was denatured by addition of 125 μL 2 M HCl and precipitated with 125 μL 20% w/v TCA in water. Samples were spun down at 16000 x g (20 minutes), and the supernatant was collected and diluted in 250 μL saturated NaAc. Free iron was then reduced to the 2+ oxidation state with 6 μL thiolglycolic acid immediately prior to chelation with 50 μL 2,2′-bipyridine (0.4% w/v in an aqueous solution of 5% v/v acetic acid) to form $[Fe(bpy)_3]^{2+}$. Finally, iron was quantified by measuring absorbance at 522 nm using an extinction coefficient of 8789 $M^{-1}cm^{-1}$.

UV-visible Spectroscopy with and without DNA

To see if changes to the cluster environment could be observed at lower energies, the effect of DNA on the cluster environment in WT EndoIII and the mutants E200K, Y205H, and K208E was assessed by UV-visible spectroscopy. 10 μM EndoIII (by cluster) was prepared either alone or with equimolar (by duplex) DNA of the same sequence used for XAS experiments, and incubated on ice for 30 minutes prior to measurement. Spectra were compared following subtraction of a sample containing only DNA.

EPR Spectroscopy

The stability of EndoIII mutants was assessed using continuous wave X-band EPR spectroscopy at 10 K. All samples were prepared anaerobically in a glove bag under a 95% nitrogen, 5% hydrogen atmosphere. EndoIII (WT, Y82A, E200K, Y205H, or K208E) was diluted to 20 μM in degassed storage buffer (20 mM potassium phosphate, pH 7.5, 150 mM KCl, 0.5 mM EDTA) and either added directly to an EPR tube or first treated with 0.5, 1, 3, 5, or 10-fold molar excess of $K_3Fe(CN)_6$ to oxidize the cluster. Oxidized samples were frozen in liquid nitrogen as soon as possible following treatment. Spectra were taken at 12.88 mW microwave power, 2 G modulation amplitude, and 5.02 x 103 receiver gain. All samples were scanned 3 times, and a buffer background was subtracted from each. The lack of $K_3Fe(CN)_6$ contribution to the signal was confirmed with a sample containing only oxidant and buffer.

Results

S K-Edge XAS of EndoIII in the Absence and Presence of DNA

Sulfur K-edge XAS spectra of EndoIII in the absence and presence of a 20-mer mixed sequence DNA duplex (5′-GTGAGCTAACGTGTCAGTAC-3′ and its complement) at a ratio of 1 mol protein to 20 mol base-pairs were measured in both solution and lyophilized forms. UV-visible and CD spectra were taken of EndoIII before lyophilization and once redissolved post-treatment (Figure 3.1). Both spectra were indistinguishable, confirming that lyophilization does not affect cluster integrity or global protein folding. Protein concentrations at 1 mM or greater were used for high quality XAS data. Although highly concentrated, the protein shows no evidence of aggregation, and no changes are observed in UV-visible spectra or enzymatic activity levels of protein diluted from similarly concentrated stocks. Equal concentrations of EndoIII were compared without and with the DNA 20-mer. The data were first normalized to one sulfur, then corrected for non-bonding sulfur atoms (six in EndoIII) and the iron-sulfur cluster loading ratio, calculated as described in Materials and Methods. Cluster loading ratios for EndoIIII were 70% or greater. The spectra were then multiplied by 8 to evaluate the total contribution of four sulfide and four thiolate sulfur atoms.

The normalized S-K edge XAS spectra of the pre-edge region and the background (rising edge) subtracted spectral fits are shown in Figure 3.2. The energies and intensities obtained from the fits included in Figure 3.2 are given in Table 3.1. The peak assignments and the pre-edge shapes used in these fits are based on our previous results on [4Fe4S] clusters (*6, 7*). All pre-edge

spectra have two major features. The lower energy feature at ~2470.2 eV is due to the μ-3 sulfide S 1s → Fe 3d transitions, and the higher energy feature at ~2470.9 eV is due to the thiolate S 1s → Fe 3d transitions. Both gain intensity through S 3p mixing into Fe 3d orbitals. For both unbound and DNA bound EndoIII, lyophilization decrease the pre-edge intensity. Notably, redissolving lyophilized EndoIII results in XAS spectra indistinguishable from samples that were never dried, consistent with UV-visible and CD spectra and providing further confirmation that lyophilization does not dramatically alter the protein (Figure 3.1c). Likewise, for both solution and lyophilized data sets, DNA binding increases the pre-edge intensity, indicating that the Fe-S bonds become more covalent. DNA binding in solution increases the total S covalency from 573% to 631% (Table 3.1). On the basis of our past studies, each % increase of total S covalency corresponds to ~3.3 mV decrease in redox potential (*7*). Thus, the covalency increase upon DNA binding in Figure 3.1 and Table 3.1 corresponds to a decrease of the reduction potential of the $[4Fe4S]^{3+/2+}$ couple in EndoIII by ~190 mV, which would activate this cluster for oxidation. Critically, the DNA-induced potential shift measured by XAS is nearly identical to that measured electrochemically (*12*).

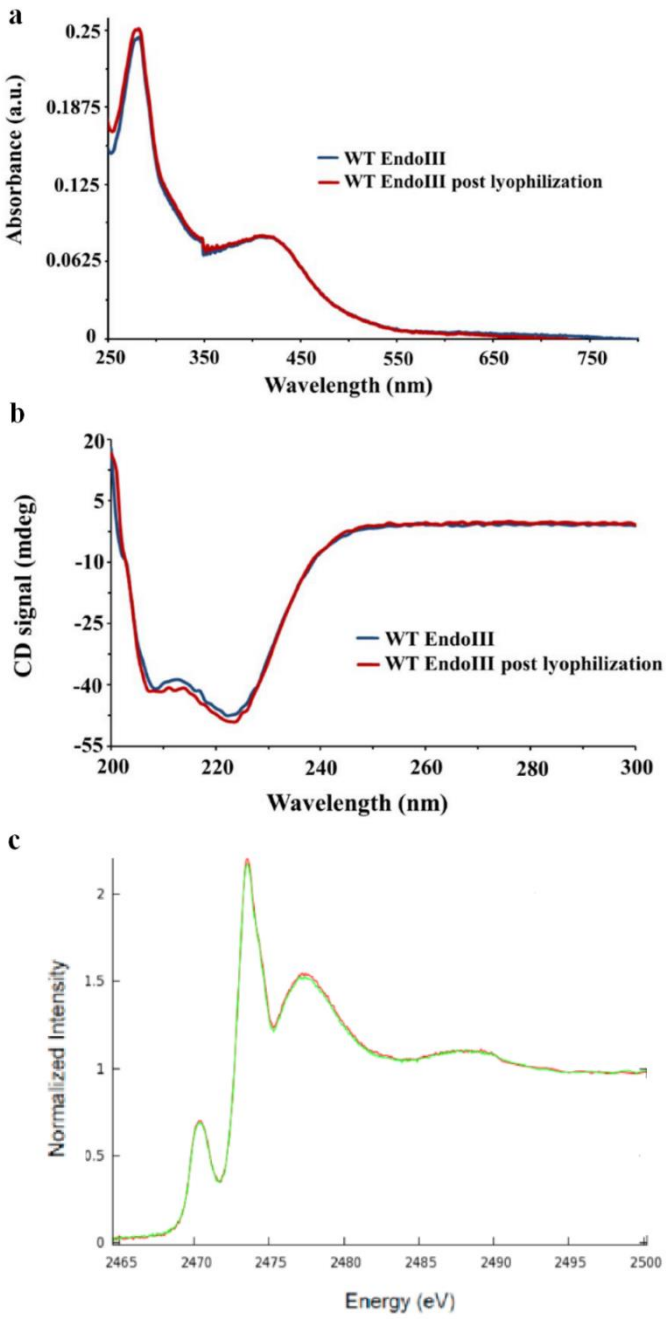


Figure 3.1 Spectroscopic data to confirm that lyophilization does not perturb the global structure of EndoIII or the cluster itself. UV-visible spectra (a) and circular dichroism (CD) spectra (b) taken with untreated sample and sample redissolved post-lyophilization are indistinguishable, confirming that cluster loading and local environment, as well as global protein folding, are unaltered by lyophilization. Consistent with this data, S K-edge XAS spectra taken of solvated EndoIII and a redissolved sample that had been previously lyophilized overlay almost perfectly (c).

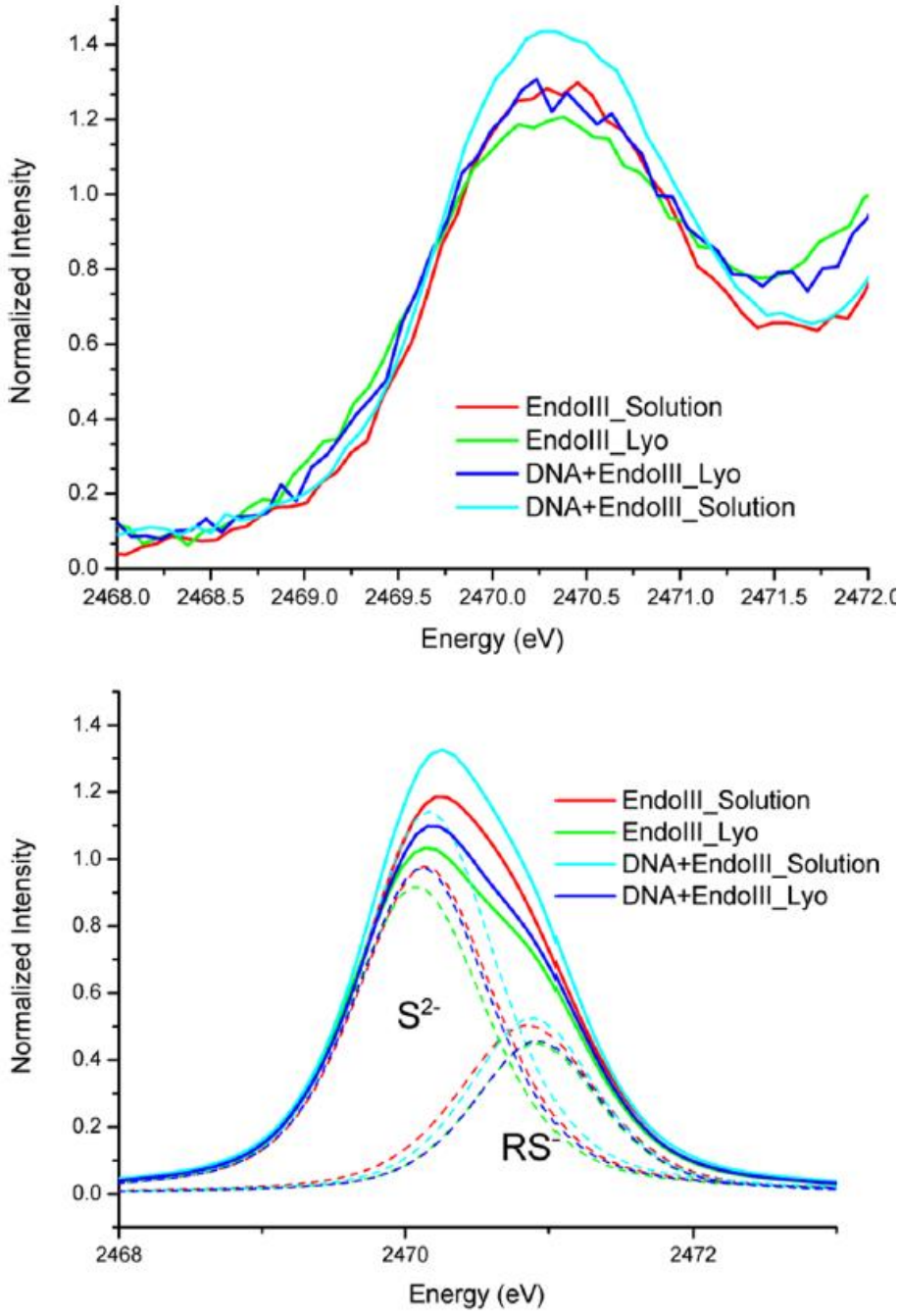


Figure 3.2 S K-edge XAS of EndoIII without and with DNA and upon lyophilization (top) and the fits of the pre-edge region (bottom) using two peaks, one for the sulfides at ~2470.1 eV and one for the thiolates at ~2470.9 eV.

Table 3.1 S K-edge XAS Data for $[4Fe4S]^{2+}$ Models and Proteins

Sample name		Energy (eV)	Area	Covalency (% per S)	Total Covalency
$[Fe_4S_4(SEt)_4]^{2-}$ model	Sulfide	2470.1	1.433	43.8	689.9±30
	Thiolate	2470.9	0.578	41.0	
Fd Solution	Sulfide	2470.1	1.145	35.0	555±23
	Thiolate	2470.9	0.481	34.0	
Fd Lyophilized	Sulfide	2470.1	1.276	39.0	620±27
	Thiolate	2470.9	0.523	37.0	
HiPIP Solution	Sulfide	2470.1	1.253	38.3	618.0±17
	Thiolate	2470.9	0.558	39.5	
HiPIP Lyophilized	Sulfide	2470.1	1.194	36.5	588.4±15
	Thiolate	2470.9	0.503	35.6	
EndoIII Solution	Sulfide	2470.2	1.185	36.2	572.8±28
	Thiolate	2470.9	0.488	34.5	
EndoIII Lyophilized	Sulfide	2470.1	1.056	32.3	522.6±25
	Thiolate	2470.9	0.477	33.8	
EndoIII+DNA Solution	Sulfide	2470.2	1.293	39.5	631.2±30
	Thiolate	2470.9	0.553	39.2	
EndoIII+DNA Lyophilized	Sulfide	2470.1	1.106	33.8	542.2±26
	Thiolate	2470.9	0.482	34.1	

S K-edge XAS of MutY in the Absence and Presence of DNA

S K-edge XAS of MutY without and with DNA are shown in Figure 3.2. Data were treated just as with EndoIII, except that MutY spectra required correction for eighteen rather than six nonbonding sulfur atoms. As with EndoIII, controls were performed to ensure that lyophilization pretreatment did not affect the protein. For MutY, these controls consisted of active site titrations using a ^{32}P-labeled 30-mer duplex containing a central oxoG:A mispair (Figure 3.3) (*26*). Under conditions of rate-limiting product release (burst kinetics), burst amplitudes, which correlate to active site concentration, exhibited no significant differences between samples.

XAS spectra reveal that, similar to EndoIII, the pre-edge intensity increases upon DNA binding, and decreases upon lyophilization (Figure 3.4). This is consistent with the fact that the N-terminal domain of MutY is structurally similar to EndoIIII in the presence of DNA (Figure 3.5). Notably, while MutY and EndoIII are structurally homologous, MutY shows specificity for adenine mispaired with 8-oxoguanine, whereas EndoIII targets a range of oxidized substrates. Incorporating 2′-deoxy-2′-fluoroadenosine (FA) in place of A across from an oxoG site generates a non-hydrolysable substrate mimic that enhances DNA binding affinity in MutY (*27*). Thus, to assess differences in non-specific and target DNA binding in MutY, the S K-edge XAS were measured both in the presence of specific (oxoG:FA) and non-specific DNA and with different DNA pair lengths under the same conditions as for EndoIII. One difference between MutY and Endo III is that MutY has an extra domain (Figure 3.5) that plays an important role in OG recognition and proper engagement of MutY on the oxoG:A substrate mispair (*28*). Figure 3.4

shows that in the case of MutY only the presence of specific DNA leads to the pre-edge intensity increase, and that the length of the DNA strand (15 and 30 base pairs) does not impact these intensity changes. This may be a consequence of affinity and MutY conformation with specific versus nonspecific DNA. MutY has much higher affinity for substrate-like DNA over nonspecific DNA, and MutY also induces dramatic remodeling of its substrate. However, there are more non-bound sulfur atoms in MutY than in EndoIII contributing to the background and precluding quantification of the changes in Figure 3.4.

Figure 3.5a shows that MutY and EndoIII not only are similar in conformation but share a set of conserved Arg residues that are responsible for DNA binding. Figure 3.5b and c show that these Arg locations are not perturbed by the DNA in EndoIII, although these two structures (both for *E. coli* protein homologues) share only 43% sequence similarity.

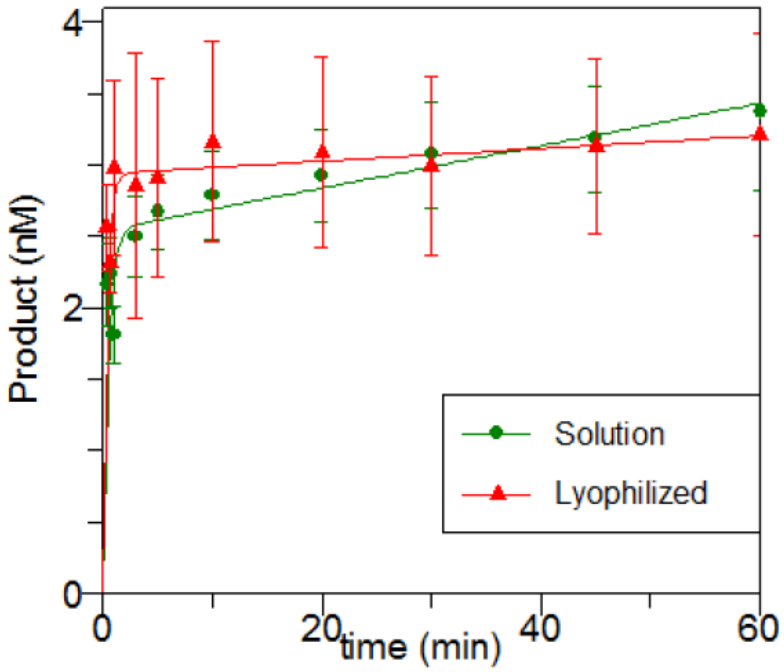


Figure 3.3 Average product curves and standard error bars of three separate aliquots each of solution (green) and lyophilized (red) MBP-MutY to determine the percent active fraction under maximum turnover conditions.

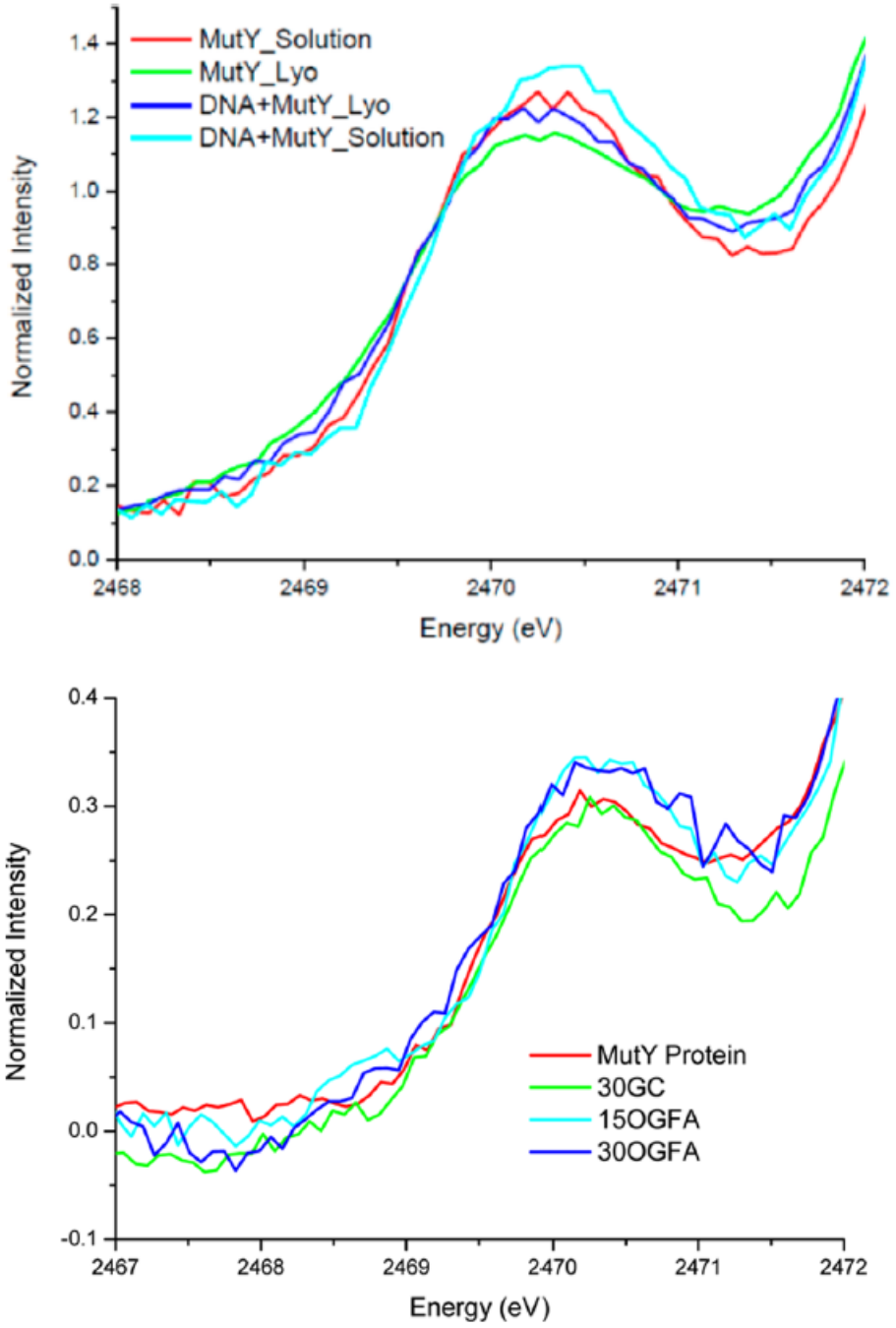


Figure 3.4 S K-edge XAS of MutY without and with nonspecific DNA and upon lyophilization (top), and S-K edge spectra showing impact of different DNA length (15 or 30) and of specific (OG:FA) as well as non-specific (GC) binding on MutY (bottom).

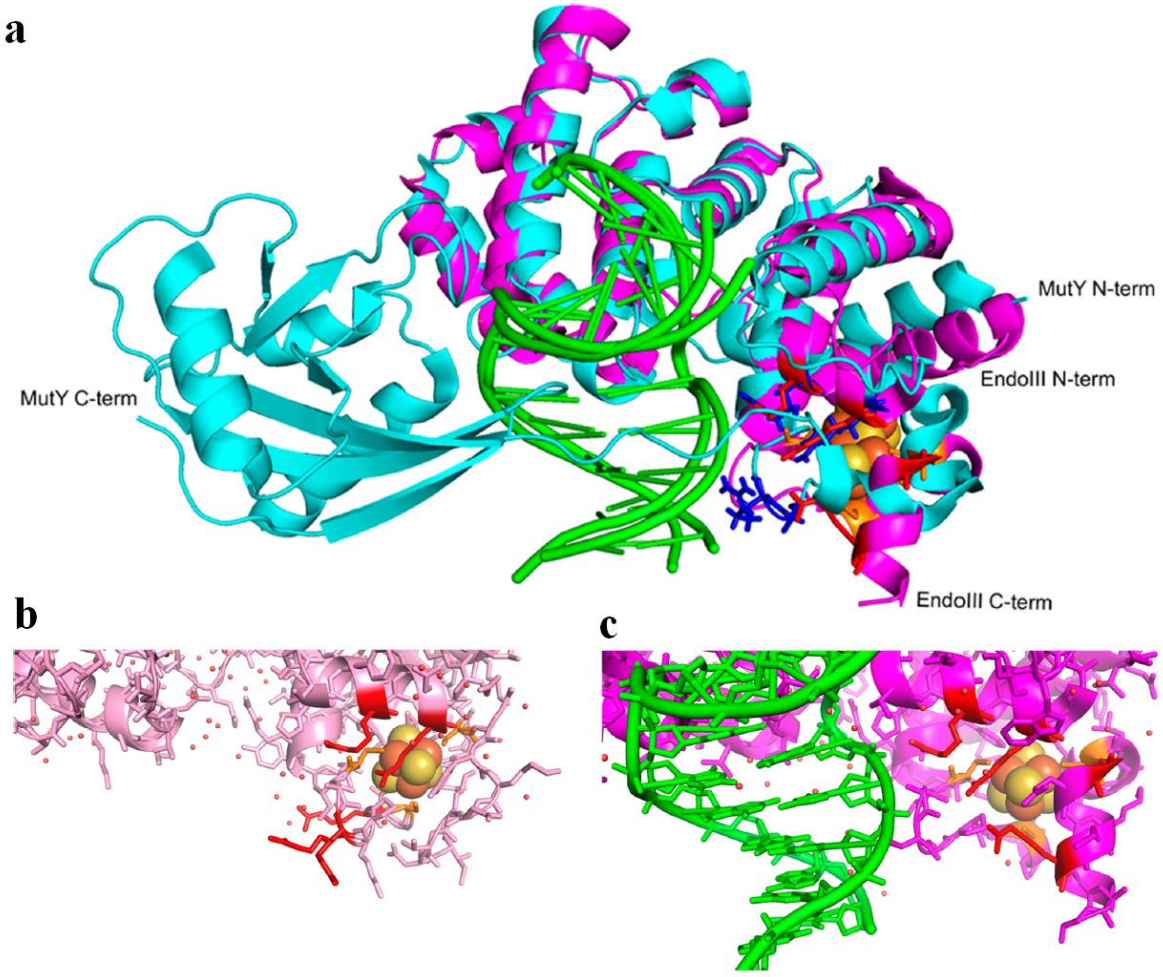


Figure 3.5 (**a**) Overlay of *G. stearothermophilus* (*Gs*) EndoIII (magenta, PDB 1ORN) and *Gs* MutY (cyan, PDB 5DPK), with [4Fe4S] cluster (orange for S and yellow for Fe atoms) and key Arg residues highlighted (red and blue, respectively). Bound DNA in both structures is green. (**b**, **c**) Comparison of *Ec* EndoIII without DNA (pink, PDB 2ABK) in (**b**) and *Gs* EndoIII (magenta, PDB 1ORN) bound to DNA (green) in (**c**) showing the molecular surroundings of the [4Fe4S] cluster, including critical Arg residues (red). This comparison highlights the overall structural similarity between *Gs* EndoIII in complex with DNA to that of the *Ec* homologue without DNA. Small structural differences ($RMSD_{100}$ = 1.4 Å) between the two homologues are attributed to their 43% sequence similarity.

Solvent Effect on HiPIP-Type Proteins

From our previous studies, lyophilization perturbs the H-bond environment around the [4Fe4S] cluster sites. For the studies described below, the spectroscopic features and protein activity were measured upon redissolving the lyophilized proteins to make sure that the lyophilization process did not lead to irreversible denaturation. Figure 3.2 and Table 3.1 show that upon lyophilization, the sulfur covalency of EndoIII decreases by ~50% total sulfur character out of ~570%, which is opposite to the solvent effect observed for [4Fe4S] *Bacillus thermoproteolyticus* ferredoxins (*Bt* Fd) (*7*). The pre-edge intensity of the [4Fe4S] cluster in *Bt* Fd increases by ~65% total sulfur character out of ~555% upon removal of solvent water. Unlike *Bt* Fd, where the [4Fe4S] cluster is exposed at the surface, EndoIII has a buried [4Fe4S] cluster, similar to MutY and *Chromatium vinosum* HiPIP. The XAS spectra of EndoIII (Figure 3.2), MutY (Figure 3.4) and *C. vinosum* HiPIP (Figure 3.6) all show a decrease in S pre-edge intensity with lyophilization (~30% out of a total of 618% sulfur character for HiPIP).

As shown in Figure 3.7a, the [4Fe4S] cluster in HiPIP (PDB code 1CKU) has five H-bonds from the protein environment to the thiolate sulfurs within 3.5 Å, all from the amide backbone. The higher S K-edge intensity in HiPIP relative to Fd and its decrease in S covalency upon lyophilization indicate that, relative to [4Fe4S] Fd, the H-bonds to [4Fe4S] cluster in HiPIP are weak. Removal of the solvent in HiPIP would lead to a more compact site, increasing the backbone H-bonds to the cluster. Also, three of the amides are solvent exposed, which could be directly affected by lyophilization (boxed). Furthermore, there are waters H-bonded to the

carbonyls that are conjugated to the amide N atoms, and removal of these water would enhance their N-H...S hydrogen bonds.

Figure 3.7b shows that in the DNA-free form of EndoIII (PDB code 4UNF), the [4Fe4S] cluster has six H-bonds to thiolate sulfurs within 3.5 Å: one from arginine, one from histidine, and four from the amide backbone. Figure 3.7c shows a similar H-bond pattern around the cluster for DNA-bound EndoIII (PDB code 1ORN). In EndoIII, the increased number of H-bonds and the stronger H-bond from the positively charged Arg are consistent with the somewhat lower S covalency of EndoIII relative to that of HiPIP (Table 3.1, solution: 573% vs 618%; lyophilized: 523% vs 588% for EndoIII and HiPIP, respectively). In addition to the more compact site that would result from lyophilization, the Arg, His, and amides are all at the surface of the protein and would also be directly impacted by loss of H_2O (red dots in Figure 3.5 and boxed region in Figure 3.7b and c). The larger number of surface exposed H-bonds is consistent with the larger decrease in S edge intensity in EndoIII relative to HiPIP upon lyophilization (50% vs 30% respectively). MutY has a similar DNA binding domain structure to EndoIII, and a surface Arg is shown in blue in Figure 3.5a.

Finally, as displayed in Figure 3.2 and Table 3.1, lyophilization of DNA bound EndoIII also leads to an intensity decrease (somewhat larger than for the unbound EndoIII). Thus even when bound to DNA, there is solvent access to the [4Fe4S] cluster in EndoIII.

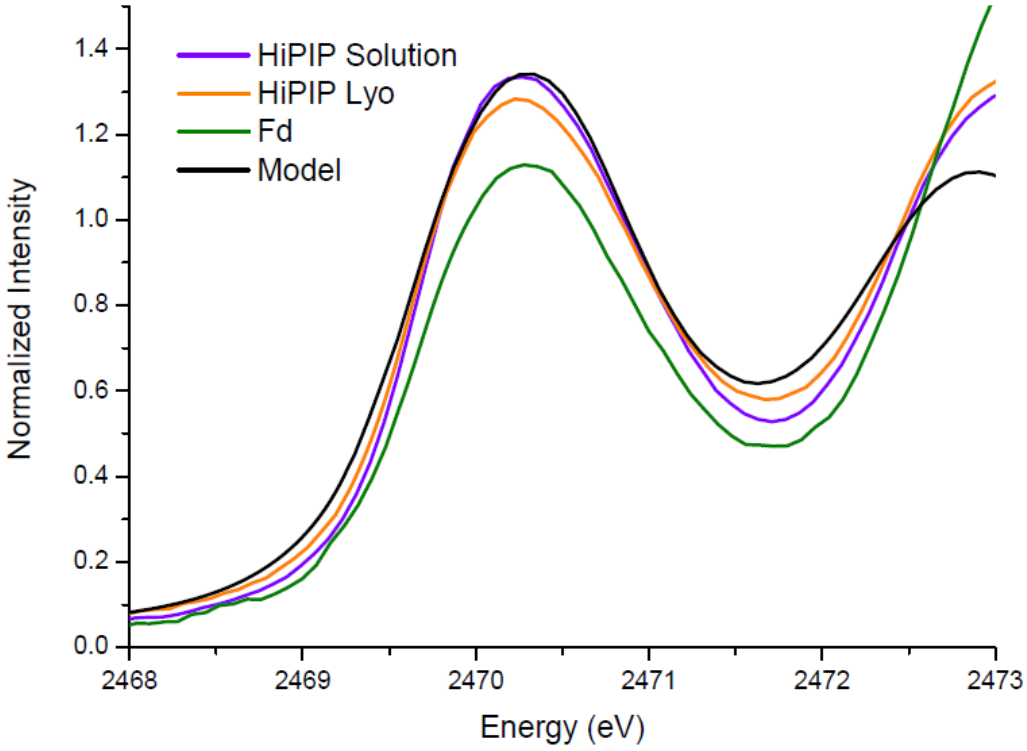


Figure 3.6 S K-edge XAS spectra of HiPIP compared with a model and Fd. As with EndoIII, HiPIP pre-edge absorbance is highest for solvated sample. In all cases, however, HiPIP absorbance is much higher than that of Fd, consistent with lower solvent exposure of the cluster in HiPIP.

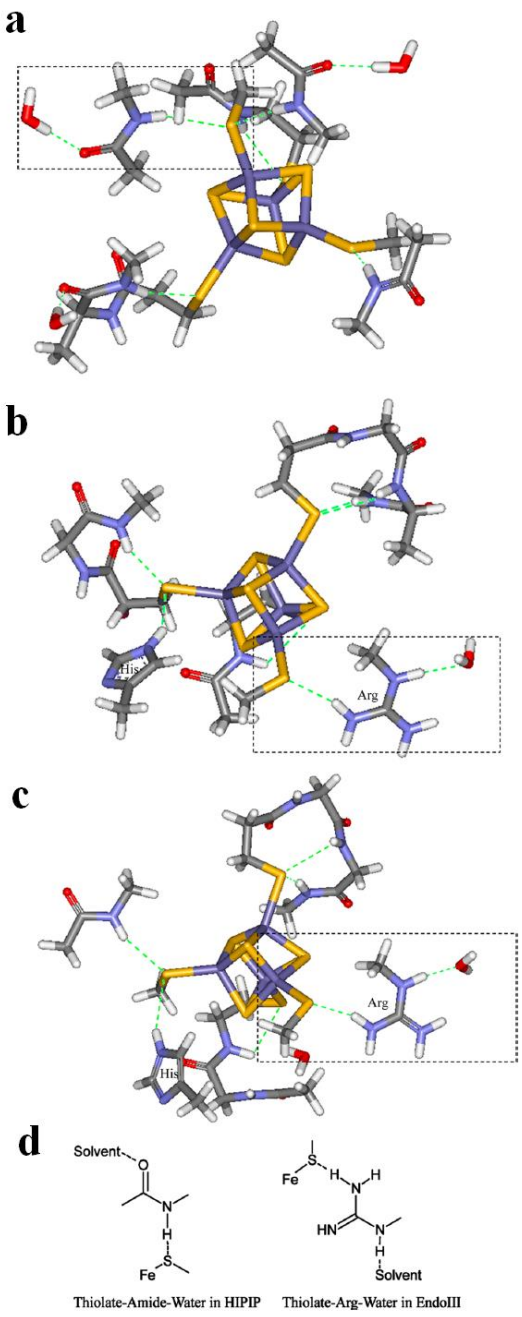


Figure 3.7 Local H-bond network of the Fe_4S_4 cluster for (**a**) HiPIP (PDB code 1CKU) and (**b**) EndoIII without DNA (PDB 4UNF) and (**c**) with DNA (PDB 1ORN), illustrating some of the H-bonds that could be affected by solvent water. The H-bonds are shown in dashed green, and Arg and His are labeled. Schemes of the boxed region are shown in (**d**). In (**a**), an amide backbone H-bond to the thiolate S bound to Fe and accessible solvent is shown in the box, and in (**b**) and (**c**), an arginine residue H- bond to the thiolate bound to Fe and accessible to solvent is shown in the box.

DFT Calculations

DNA is rich in negative charge. To qualitatively test whether the presence of negative charge would increase total [4Fe4S] sulfur covalency and understand the nature of this effect, simple models were evaluated. The results are shown in Table 3.2.

Placing a negative point charge 5 Å away from one thiolate increases the total sulfide and thiolate contributions from 429.0% to 439.1% and from 139.3% to 144.8% respectively, which correspond to a total sulfur contribution increase from 568.4% to 583.9%. This is qualitatively similar to experiment but the effect is less than observed experimentally due to the simple model. Placing a negative point charge 5 Å from one sulfide gave similar results (Table 3.2); the total sulfide and thiolate contributions increased to 438.3% and 143.9% respectively, and total sulfur contribution increased to 582.2%.

The DFT calculations in Table 3.2 show that the addition of the negative charge near the [4Fe4S] cluster can qualitatively model the increase in covalency observed with DNA binding and thus provide insight into the origin of this effect. Figure 3.8 gives the change in Mulliken charge distribution (a), Mayer bond orders (MBO) (b) and C^2 contributions (c, which sums the contributions of the valence atomic orbitals over the unoccupied valence molecular orbitals, for each Fe and S atom of the cluster) due to the addition of the negative point charge. The change in C^2 would directly relate to the change in S edge intensity observed experimentally. The changes calculated for placing the point charge next to a thiolate are given in blue, and next to a sulfide are given in red.

Adding a point charge near the thiolate (Figure 3.8, blue) results in a shift of the electron density toward the opposite side of the cluster (a). The thiolate next to the point charge (denoted S*) loses electron density (change in Mulliken charge +0.13), while the three thiolates and one sulfide on the opposite side of the cluster gain some negative charge (~ -0.03, A). Thus S* donates more charge density to the Fe it coordinates (denoted Fe*); this is reflected in its increase in MBO (+0.12 for Fe*-S* in B). The three sulfides that coordinate to Fe* then donate less to it to compensate (MBO ~-0.01 each in b), but donate more to the remaining Fe's (MBO ~+0.01). The other three thiolates have decreased MBO (-0.02 each in b) as they now donate less charge density to their coordinated Fe atoms, which is consistent with their increase in negative Mulliken charge (a). Although there are some compensation effects, the total MBO increases (+0.06). In terms of C^2 (Figure 3.8c), because the Fe*-S* bond becomes stronger (from the MBO), more S 3p character is mixed into the unoccupied Fe 3d orbitals, leading to the increase in S* C^2 contribution (+3.4) and decrease of the Fe* contribution (-4.9). Although the other three sulfide bonds to Fe* get slightly weaker (from the MBO in b), these sulfides have net higher C^2 (+4.6, -0.4, +3.4 in C). The negative point charge at S* destabilizes the occupied S p orbital energies (Scheme 1), which results in more S character mixed into the unoccupied valence molecular orbitals. The Fe d characters all decrease in response to the increase of S p, and both sulfide and thiolate C^2 increase as observed experimentally.

Similar effects were observed when placing the point charge 5Å away from the sulfide at the corner of the cube (Figure 3.8, red). Note that except for the S atom and Fe-S bond(s) closest to

the charge, there is only a limited change in Mulliken charge, MBO and C^2 contributions for the rest of the cluster. The net change is an increase in C^2 character, again qualitatively consistent with the S K-edge XAS experimental data.

Table 3.2 Results of DFT Calculation

Model	Total Fe %	Total Sulfide %	Total Thiolate %	Total S %
$[Fe_4S_4(SMe)_4]^{2-}$ model	958.4	429.0	139.3	568.4
$[Fe_4S_4(SMe)_4]^{2-}$ with $1e^-$ at 5Å from sulfide	944.1	438.3	143. 9	582.2
$[Fe_4S_4(SMe)_4]^{2-}$ with $1e^-$ at 5Å from thiolate	949.3	439.1	144.8	583.9

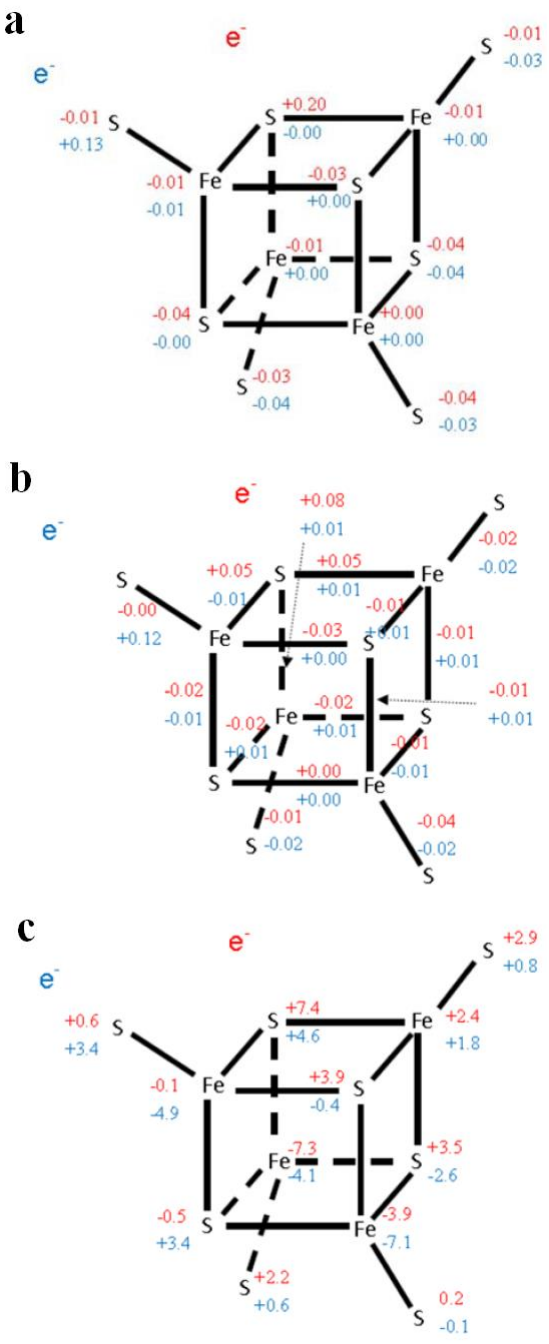
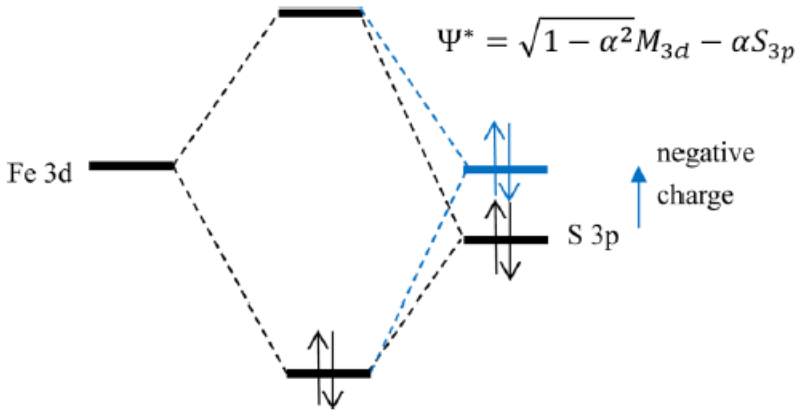


Figure 3.8 Difference in Mulliken charge distribution (**a**), Mayer bond orders (MBP) (**b**) and C2 contributions (**c**) of each Fe and S atom of the cluster without and with a point charge at 5 Å distance from the cluster. Blue represents values obtained for a point charge near a thiolate S, and red represents those for a point charge near a sulfide S.



Scheme 3.1 Adding negative charges such as DNA in the proximity of the [4Fe4S] cluster destabilizes the S 3p orbital energy and increases the S 3p character (α^2) in Ψ*. This increases the Fe-S covalency and stabilizes the oxidized form than the reduced form of the cluster and decreases the reduction potential.

Spectroscopic Characterization of EndoIII Point Mutants

To see if shifting the charges near the [4Fe4S] cluster in EndoIII might alter the redox potential, we prepared a series of point mutants (E200K, Y205H, and K208E) in which the altered residues are all within 5 Å of the cluster (*19*). By changing a single charge from negative to positive (E200K) or vice versa (Y205H, K208E), we reasoned that it might be possible to shift the redox potential both negative, like DNA, and in the opposite direction. However, no evidence for a shift was seen in these mutants when bound to DNA on a gold electrode, nor was a difference observed by direct electrochemistry of DNA-free protein on edge plane graphite (*19*, *29*). The error in the latter experiments, however, did not preclude a small shift on the order of 10 mV. Overall, we wished to use S K-edge XAS as an independent means of assessing the potential in these mutants.

Solvated and lyophilized samples of each mutant were prepared in the absence and presence of DNA exactly as with WT EndoIII, and spectra were taken alongside a WT sample. Consistent with previous electrochemical results that show no potential shift in these mutants, the pre-edge absorbance of E200K overlay WT EndoIII almost perfectly; results with DNA present are inconclusive due to noise (Figure 3.9) (*7*). The pre-edge features of both Y205H and K208E are much lower than expected for HiPIP-type proteins; indeed, K208E is well in the range of Fd, and Y205H is lower still, suggesting a redox potential below -500 mV vs NHE (Figure 3.9). Such a large decrease in Fe-S covalency was wholly unexpected, and thus required further experimental follow-up. Other than a genuine potential shift in these mutants, possible

explanations included a different cluster absorbance at 410 nm (ε_{410}) than WT leading to inaccurate loading measurements and thus incorrect XAS corrections or cluster degradation before or during XAS measurements.

To address a possible difference in cluster absorbance at 410 nm, we used a combination of UV-visible spectroscopy, Bradford assays, and iron loading experiments to measure ε_{410} and [4Fe4S] loading in the same WT, Y205H, and K208E stocks used for XAS. In each case, total protein concentration and [4Fe4S] cluster absorbance were measured for three replicates each of WT EndoIII and each mutant, and an equal number of replicates were subjected to iron loading analysis to obtain total iron concentration in the sample. By this analysis, ε_{410} values were determined to be 17618 ± 229 $M^{-1}cm^{-1}$, 16661 ± 221 $M^{-1}cm^{-1}$, and 18008 ± 135 $M^{-1}cm^{-1}$ for WT, K208E, and Y205H, respectively. Notably, none of these differs appreciably from the assumed value of 17000 $M^{-1}cm^{-1}$, and cluster loading determined by Bradford assays using these updated values was not greatly altered, indicating that the unexpected XAS results did not occur as a miscalculation of cluster loading.

Having confirmed that cluster loading calculations could not have caused the discrepancies in pre-edge absorbance, we sought an independent technique to probe the cluster environment of the EndoIII mutants. Previous electrochemical results with DNA present suggested that the cluster environment of the mutants was dominated by DNA binding, in contrast to the XAS results. In an effort to resolve this issue, we compared UV-visible spectra of WT and mutant EndoIII in the presence and absence of DNA, reasoning that any changes in

cluster electronic environment induced upon DNA binding should be similar between all EndoIII variants if the differences seen by XAS were due to sample degradation. In three replicates of WT EndoIII, the broad absorbance of the [4Fe4S] cluster did not change markedly in the presence of DNA, although a slight increase in absorbance was consistently present (Figure 3.10). These data suggest that DNA binding does induce a very minor, but still detectable, change in cluster electronics that could be used to compare WT and mutants. These experiments were repeated with E200K, Y205H, and K208E, all of which showed a very similar small increase in absorbance upon DNA binding (Figure 3.10b). The similarity of absorbance changes in WT and all mutants confirms that DNA binding has the same effect on cluster electronics as WT, supporting previous electrochemical results and suggestive of sample degradation in the XAS experiments.

At this point, our experiments indicated that the iron loading in XAS samples was correct, and that neither Y205H nor K208E was noticeably different by UV-visible absorbance at 410 nm in the presence of DNA. Previous electrochemical, melting temperature, and CD analysis suggested that the cluster is more solvent-exposed in Y205H and K208E, lowering their stability relative to WT, while E200K is more shielded and of comparable or greater stability than WT (*19*). To see if low stability might explain the XAS results, we carried out X-band EPR spectroscopy at 10 K with WT, E200K, Y205H, K208E, and an additional mutant, Y82A (Figure 3.11). Y82A is known to be CT-deficient in the presence of DNA, but is expected to have an identical cluster environment to WT (*19*); thus, this mutant served as an additional control for

these experiments. To probe cluster stability, we treated protein samples in the absence of DAN with $K_3Fe(CN_6)$ in varied molar equivalents. $K_3Fe(CN)_6$ does not oxidize EndoIII very effectively in the absence of DNA, but a quantifiable signal centered at g = 2.02 attributable to a $[Fe_3S_4]^+$ cluster can nonetheless be observed at a sufficient excess of oxidant (*8, 11*). If some mutants were unstable, they would be expected to be more sensitive to such treatment than WT.

As expected, 20 μM WT EndoIII was EPR silent when untreated and showed a clear $[3Fe4S]^+$ signal at g = 2.03 when treated with a molar excess of oxidant (Figure 3.11a). Signal size increases linearly with increasing oxidant concentration, although some signal loss occurrs at 10-fold excess $K_3Fe(CN)_6$ when samples were not frozen immediately after treatment. The latter effect is consistent with oxidative degradation. To minimize this risk, EPR spectra of EndoIII mutants were initially taken at a 5-fold molar excess of $K_3Fe(CN)_6$. Like WT, all mutants are EPR silent when untreated. Y82A signals are as large or larger than those of WT, confirming that the cluster environment in this mutant is, as expected, unaltered. K208E spectra are slightly smaller than WT in the presence of 5-fold molar excess $K_3Fe(CN)_6$, while E200K and Y205 are both much smaller (Figure 3.11b). E200K signals could be increased by including a 10-fold excess of oxidant, but still remained much smaller than those of WT; these results are in agreement with previous work suggesting a more solvent-shielded cluster in this mutant. In contrast, Y205H signals decreased with 10-fold excess oxidant, and were much more sensitive to incubation time than any other EndoIII variant. To see if better signals could be obtained, we tried treating Y205H with 0.5, 1, and 3-fold excess oxidant; signals increased linearly up to 3-

fold excess, and were similar to those obtained at 5-fold excess. These data support the previously mentioned instability of Y205H; furthermore, UV-visible spectra of some (but not all) stocks sent for XAS showed increased absorbance associated with oxidative degradation after one additional freeze-thaw cycle. In summary, our EPR results are in agreement with previous studies, indicating greater instability in the Y205H cluster. EPR spectra did not show any significant differences between K208E and WT, but the bulk of previous evidence supports cluster instability in this mutant as well, although to a lesser degree than Y205H.

Overall, the XAS results for EndoIII mutants are equivocal. E200K spectra fit the expected pattern and match previous results, but K208E and, especially, Y205H appear to be too unstable to carry out reliable experiments. Thus, the unexpectedly low pre-edge absorbance seen in the Y205H and K208E XAS spectra most likely reflects cluster degradation rather than an actual shift to a Fd-like environment.

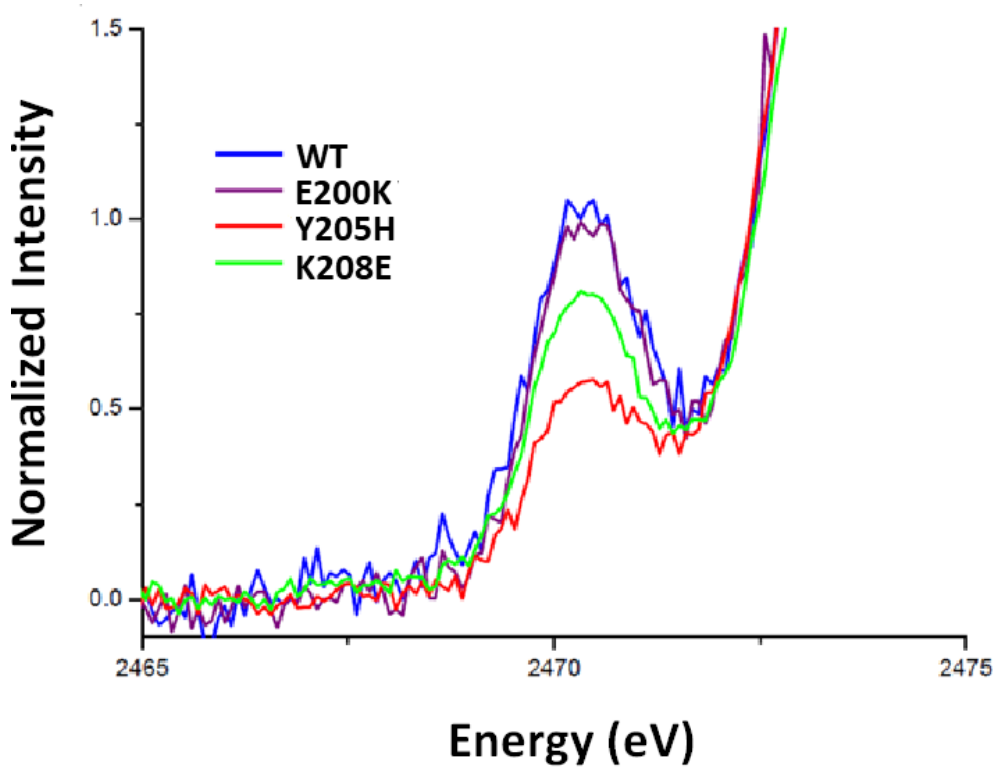


Figure 3.9 S K-edge XAS spectra of WT EndoIII and the mutants E200K, Y205H, and K208E. These mutations all reside within 5 Å of the cluster, and were intended to alter the local environment sufficiently to induce a shift in potential. E200K overlaid well with WT, consistent with electrochemical experiments, while K208E and Y205H showed pre-edge absorbance more similar to Fd than HiPIP or the other EndoIII variants. The latter results were likely due to cluster degradation in these less stable mutants.

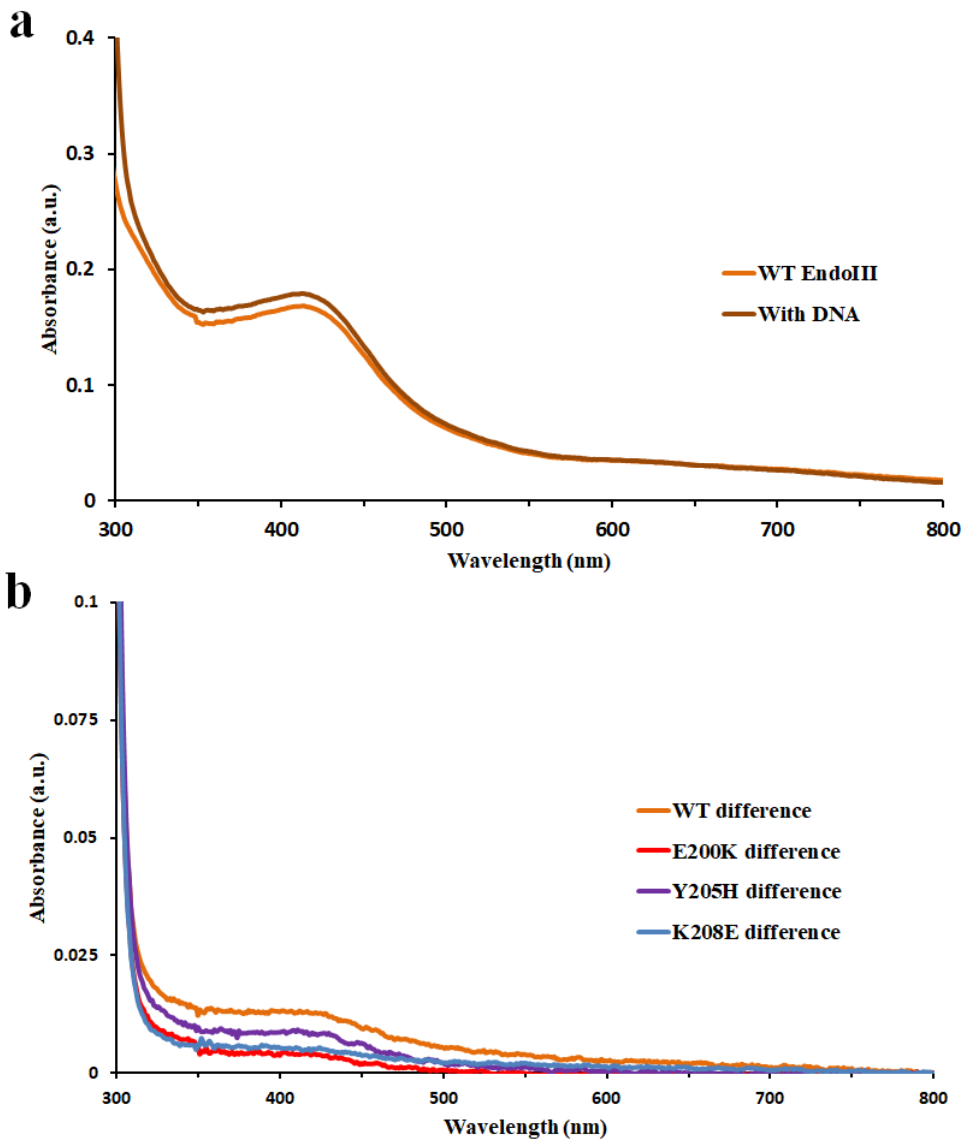


Figure 3.10 UV-visible spectra of WT EndoIII and the mutants E200K, Y205H, and K208E with and without DNA present. WT EndoIII shows very little difference in absorbance at these wavelengths (**a**), with only a slight increase most easily seen in subtraction spectra (**b**). The other mutants were similar, indicating no significant differences in cluster environment upon DNA binding.

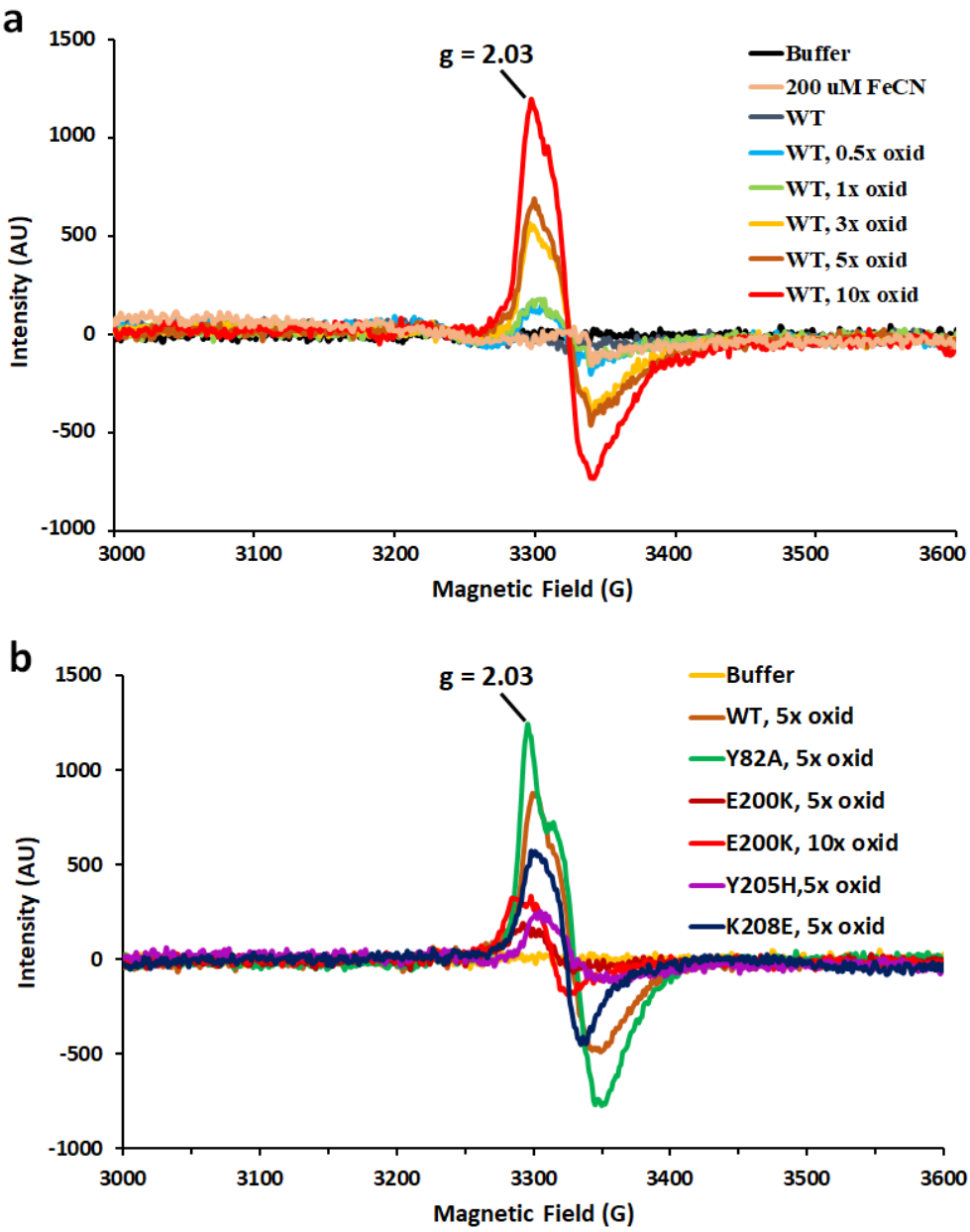


Figure 3.11 Continuous wave X-band EPR spectra of 20 μM WT EndoIII (**a**) and the mutants Y82A, E200K, Y205H, and K208E treated with $K_3Fe(CN_6)$ in the indicated molar excess (**b**). All variants are EPR silent when untreated and show formation of the $[3Fe4S]^+$ oxidative degradation product in the presence of $K_3Fe(CN)_6$. WT and Y82A yield substantial signals, while E200K is markedly less sensitive to oxidation. Y205H is highly susceptible to oxidation and degradation, giving the largest signals with 3-fold excess oxidant. K208E was similar to WT, although signals were slightly smaller. All spectra were taken at 10 K, 12.88 mW microwave power, 5.03 x 10^3 receiver gain, and are the average of 3 scans.

Discussion

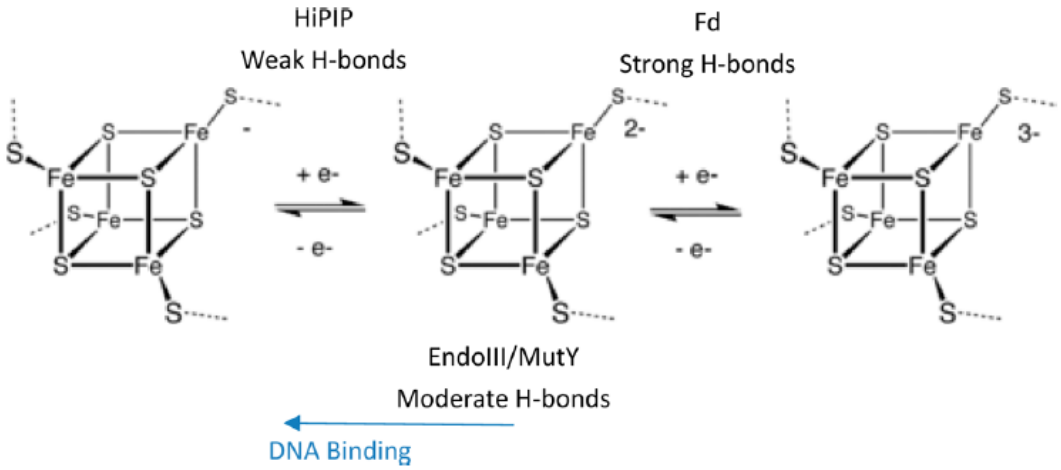
Both EndoIII and MutY contain [4Fe4S] clusters that are largely redox inert in solution, but their reduction potentials decrease by ~200 mV upon DNA binding to turn on their function in DNA damage recognition. In nature, a large number of proteins also have [4Fe4S] clusters, with a wide range of reduction potentials (Fd: -700 to -300 mV; HiPIP: 100 to 400 mV) (*1*) (Scheme 2). S K-edge XAS has been found to be a powerful technique to evaluate the contributions of H-bonds and electrostatics to the reduction potentials of these [4Fe4S] clusters (*30*). From our past studies on Fd and HiPIP, the redox properties of these [4Fe4S] clusters are highly related to their local protein environment (*7*). In Fd, the [4Fe4S] cluster is at the surface of the protein. There are strong H-bonds from solvent water, and the Fe-S bonds are less covalent. This raises the reduction potential of the +2 state and results in the $[4Fe4S]^{2+/1+}$ redox couple in the physiological potential range. Upon lyophilization, the S covalency significantly increases due to the loss of H-bonds from solvent to the [4Fe4S] cluster and the covalency becomes similar to that of HiPIP. In contrast, HiPIP, which is not solvent exposed, has only weak H-bonds from the backbone amides to the iron-sulfur cluster, and thus more covalent Fe-S bonds. In HiPIP, the $[4Fe4S]^{3+/2+}$ redox couple is activated in the physiological potential range (Scheme 2). In the present study, lyophilization of HiPIP leads to an inverse effect on the Fe-S bond covalency relative to Fd (i.e. the covalency decreases). Loss of solvent would lead to a more compact site and increase the H-bonding to the [4Fe4S] cluster from backbone. Loss of water H-bonds to the

surface amides would further strengthen these H-bonds to the thiolates. Lyophilization results in a similar effect on EndoIII and MutY.

EndoIII and MutY both have a strong Arg H-bond to their $[4Fe4S]^{2+}$ cluster, which is consistent with the lower covalency relative to HiPIP and making the [4Fe4S]3+/2+ couple less accessible (Table 3.1). Overall, this behavior correlates with their lack of redox activity in the absence of DNA (*31*, *32*). S K-edge XAS experiments show that DNA binding to EndoIII and MutY significantly increases their Fe-S bond covalency, while in EndoIII DNA binding lowers its $[Fe_4S_4]^{3+/2+}$ potential by ~200 mV into the physiologically accessible range. An alignment of the unbound and DNA bound structures of EndoIII (Figure 3.12) shows no significant distortion of the protein ($RMSD_{100}$ = 1.4 Å; note that these EndoIII proteins are homologues with 43% sequence similarity); thus, DNA binding does not appear to structurally affect its $[Fe_4S_4]$ site. DNA binding, although at a distance of ~15 Å from the [4Fe4S] cluster, does introduce significant negative charge into the system. For the iron-sulfur cluster, the negative charge would destabilize the occupied S p orbital energies, and result in more S character donated into the unoccupied valence Fe 3d-based molecular orbitals (Scheme 2). This increases the Fe-S bond covalency of the cluster, stabilizing the oxidized over the reduced state, and thus decreasing the reduction potential. Simple DFT models using a negative point charge to represent the effect of the DNA qualitatively reproduce this covalency increase and show that it dominantly involves an

increase in electron donation from the S oriented toward the negative charge with some limited compensation by the remote sulfurs.

S K-edge XAS studies on iron sulfur proteins both alone and bound to DNA, and with and without solvent water show that the local environments of the [4Fe4S] clusters affect the covalencies of their Fe-S bonds and thus tune the potentials of the clusters. In Fd, the strong H-bonds from solvent reduces the covalency and stabilizes the reduced form of the cluster, while in HiPIP, the cluster is buried and with only weak amide H-bonds, stabilizing the oxidized form. In EndoIII and MutY the cluster is also buried but with an additional Arg H-bond that reduces the covalency relative to HiPIP and makes the HiPIP couple less accessible. However, binding to DNA introduces negative charge that increases the [4Fe4S] covalency for both EndoIII and MutY. Previous correlations between [4Fe4S] covalency and redox potential indicate that this decrease~~s~~ corresponds to an approximately 200 mV negative shift in reduction potential and would enable the $[4Fe4S]^{2+/3+}$ redox couple that is observed experimentally. Overall, the results obtained by XAS are in remarkably close agreement with electrochemical studies carried out with EndoIII in the presence and absence of DNA, indicating that the physical mechanism of the potential shift is a charge-induced increase in thiolate-iron bond covalency within the [4Fe4S] cluster. It is these molecular-level changes that are responsible for the remarkable ability of otherwise redox-inert repair proteins to communicate with each other over vast distances using reversible DNA-mediated charge transfer, making the daunting task of searching an entire genome for damage possible on a physiologically relevant time scale.



Scheme 3.2 Protein environment of the [4Fe4S] cluster tunes its redox couple. Fds have strong H-bonds to S (5 amide-thiolate H-bonds, 3 amide-sulfide bonds, and, more importantly, H-bonds to surface-exposed thiolate from solvent water) and thus the S donate less electron density to Fe, stabilizing the reduced state (redox potential range -300 to -700 mV). HiPIPs have only weak H-bonds to S (5 amide-thiolate H-bonds), thus these S donate more electron density to Fe, stabilizing the oxidized state (redox potential range 100 to 400 mV). EndoIII/MutY have moderate H-bonding (a few amide-thiolate H-bonds, and an Arg-thiolate H-bond), thus are largely redox inert. However, binding to DNA introduces negative charge, thus stabilizing the oxidized state and facilitating redox activity at potentials of ~80 mV vs NHE.

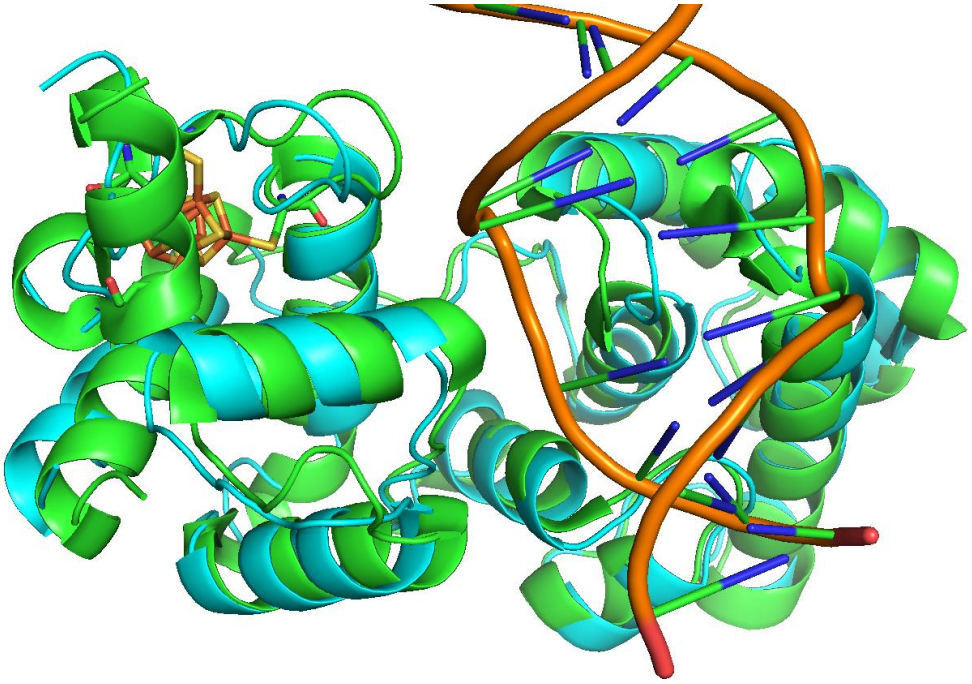


Figure 3.12 Structural alignment comparison of free (blue, PDB code 2ABK) and DNA bound (green, PDB code 1ORN) forms of EndoIII.

References

1) Holm, R. H.; Kennepohl, P.; Solomon, E. I. Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev.* **1996,** *96*, 2239 – 2314.

2) Stiefel, E. I.; George, G. N. Ferredoxins, Hydrogenases, and Nitrogenases: Metal-Sulfide Proteins. In *Bioinorganic Chemistry*, Bertini, I., Ed. University Science Books: Mill Valley, **1994**; pp 365 – 453.

3) Backes, G.; Mino, Y.; Loehr, T. M.; Meyer, T. E.; Cusanovich, M. A.; Sweeney, W. V.; Adman, E. T.; Sanders-Loehr, J. The environment of Fe4S4 clusters in ferredoxins and high-potential iron proteins. New information from x-ray crystallography and resonance Raman spectroscopy. *J. Am. Chem. Soc.* **1991,** *113*, 2055 – 2064.

4) Teo, B.-K.; Shulman, R. G.; Brown, G. S.; Meixner, A. E. EXAFS Studies of Proteins and Model Compounds Containing Dimeric and Tetrameric Iron-Sulfur Clusters *J. Am. Chem. Soc.* **1979,** *101*, 5624.

5) Czernuszewicz, R. S.; Macor, K. A.; Johnson, M. K.; Gewirth, A.; Spiro, T. G. Vibrational mode structure and symmetry in proteins and analogs containing Fe4S4 clusters: resonance Raman evidence that HiPIP is tetrahedral while ferredoxin undergoes a D2d distortion. *J. Am. Chem. Soc.* **1987,** *109*, 7178 – 7187.

6) Dey, A.; Glaser, T.; Couture, M. M.-J.; Eltis, L. D.; Holm, R. H.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. Ligand K-Edge X-ray Absorption Spectroscopy of $[Fe_4S_4]^{1+,2+,3+}$ Clusters: Changes in Bonding and Electronic Relaxation upon Redox. *J. Am. Chem. Soc.* **2004,** *126*, 8320 – 8328.

7) Dey, A.; .JenneyJr, F. E.; Adams, M. W. W.; Babini, E.; Takahashi, Y.; Fukuyama, K.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. Solvent Tuning of Electrochemical Potentials in the Active Sites of HiPIP Versus Ferredoxin *Science* **2007,** *318*, 1464 – 1468.

8) Boal, A. K.; Yavin, E.; Lukianova, O. A.; O'Shea, V. L.; David, S. S.; Barton, J. K. DNA-Bound Redox Activity of DNA Repair Glycosylases Containing [4Fe-4S] Clusters. *Biochemistry*. **2005,** *44*, 8397 – 8407.

9) Lukianova, O. A.; David, S. S. A role for iron–sulfur clusters in DNA repair. *Curr. Op. Chem. Biol.* **2005,** *9*, 145 – 151.

10) Manlove, A. H.; Nuñez, N. N.; David, S. S. The GO Repair Pathway: OGG1 and MUTYH. In *The Base Excision Repair Pathway*, III, D. M. W., Ed. World Scientific: **2016**; pp 63 – 115.

11) Cunningham, R. P.; Asahara, H.; Bank, J. F.; Scholes, C. P.; Salerno, J. C.; Surerus, K.; Munck, E.; McCracken, J.; Peisach, J.; Emptage, M. H. Endonuclease III is an iron-sulfur protein. *Biochemistry* **1989,** *28*, 4450 – 4455.

12) Gorodetsky, A. A.; Boal, A. K.; Barton, J. K. Direct Electrochemistry of Endonuclease III in the Presence and Absence of DNA. *J. Am. Chem. Soc.* **2006,** *128*, 12082 – 12083.

13) Boon, E. M.; Livingston, A. L.; Chmiel, N. H.; David, S. S.; Barton, J. K. DNA-mediated charge transport for DNA repair. *Proceedings of the National Academy of Sciences* **2003,** *100*, 12543 – 12547.

14) Boal, A. K.; Yavin, E.; Barton, J. K. DNA repair glycosylases with a [4Fe–4S] cluster: A redox cofactor for DNA-mediated charge transport? *Journal of Inorganic Biochemistry* **2007,** *101*, 1913 – 1921.

15) Boon, E. M.; Pope, M. A.; Williams, S. D.; David, S. S.; Barton, J. K. DNA-Mediated Charge Transport as a Probe of MutY/DNA Interaction. *Biochemistry* **2002,** *41*, 8464 – 8470.

16) Solomon, E. I.; Hedman, B.; Hodgson, K. O.; Dey, A.; Szilagyi, R. K. Ligand K-edge X-ray absorption spectroscopy: covalency of ligand–metal bonds *Coord. Chem. Rev.* **2005,** *249*, 97 – 129.

17) Sarangi, R.; George, S. D.; Rudd, D. J.; Szilagyi, R. K.; Ribas, X.; Rovira, C.; Almeida, M.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. Sulfur K-Edge X-ray Absorption Spectroscopy as a Probe of Ligand−Metal Bond Covalency: Metal vs Ligand Oxidation in Copper and Nickel Dithiolene Complexes. *J. Am. Chem. Soc.* **2007,** *129*, 2316 – 2326.

18) Glaser, T.; Rose, K.; Shadle, S. E.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. S K-edge X-ray Absorption Studies of Tetranuclear Iron-Sulfur Clusters: μ-Sulfide Bonding and Its Contribution to Electron Delocalization. *J. Am. Chem. Soc.* **2001,** *123*, 442 – 454.

19) Pheeney, C.G.; Arnold, A.R.; Grodick, M.A.; Barton, J.K. Multiplexed Electrochemistry of DNA-Bound Metalloproteins. *J. Am. Chem. Soc.* **2013**, *135*, 11869 – 11879.

20) Shadle, S. E.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. Ligand K-Edge X-ray Absorption Spectroscopy as a Probe of Ligand-Metal Bonding: Charge Donation and Covalency in Copper-Chloride Systems. *Inorg. Chem.* **1994,** *33*, 4235 – 4244.

21) George, G. N. *EXAFSPAK*, Stanford Synchrotron Radiaton Laboratory: Menlo Park, CA, **1990**.

22) Tenderholt, A. L.; Hedman, B.; Hodgson, K. O. In PySpline: A Modern, Cross-Platform Program for the Processing of Raw Averaged XAS Edge and EXAFS Data, AIP conference Proceedings, **2006**; pp 105 – 107.

23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas;

Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision B.01. Wallingford CT, **2009**.

24) Fromme, J. C.; Verdine, G. L. Structure of a trapped endonuclease III–DNA covalent intermediate. *EMBO J.* **2003,** *22*, 3461 – 3471.

25) Arnold, A.R.; Barton, J.K. DNA Protection by the Bacterial Ferritin Dps via DNA Charge Transport. *J. Am. Chem. Soc.* **2013**, *135*, 15726 – 15729.

26) Porello, S.L.; Cannon, M.J.; David, S.S. A Substrate Recognition Role for the $[4Fe4S]^{2+}$ Cluster of the DNA Repair Glycosylase MutY. *Biochemistry* **1998**, *37*, 6465 – 6475.

27) Chepanoske, C.L.; Porello, S.L.; Fujiwara, T.; Sugiyama, H.; David, S.S. Substrate Recognition by *Escherichia coli* MutY using Substrate Analogs. *Nucleic Acids Res.* **1999**, *27*, 3197 – 3204.

28) Chmiel, N.H.; Golinelli, M.-P.; Francis, A.W.; David, S.S. Efficient Recognition of Substrates and Substrate Analogs by the Adenine Glycosylase MutY Requires the C-Terminal Domain. *Nucleic Acids Res.* **2001**, *29*, 553 – 564.

29) Bartels, P.L.; Zhou, A.; Arnold, A.R.; Nuñez, N.N.; Crespilho, F.N.; David, S.S.; Barton, J.K. Electrochemistry of the [4Fe4S] Cluster in Base Excision Repair Proteins: Tuning the Redox Potential with DNA. *Langmuir* **2017**, *33*, 2523 – 2530.

30) Hadt, R. G.; Sun, N.; Marshall, N. M.; Hodgson, K. O.; Hedman, B.; Lu, Y.; Solomon, E. I. Spectroscopic and DFT Studies of Second-Sphere Variants of the Type 1 Copper Site in Azurin: Covalent and Nonlocal Electrostatic Contributions to Reduction Potentials. *J. Am. Chem. Soc.* **2012,** *134*, 16701 – 16716.

31) Chepanoske, C. L.; Golinell, M.-P.; Williams, S. D.; David, S. S. Positively Charged Residues within the Iron–Sulfur Cluster Loop of E. coli MutY Participate in Damage Recognition and Removal. *Arch. Biochem. Biophys.* **2000,** *380*, 11 – 19.

32) Kuo, C.-F.; McRee, D. E.; Fisher, C. L.; O'Handley, S. F.; Cunningham, R. P.; Tainer, J. A. Atomic structure of the DNA repair [4Fe-4S] enzyme endonuclease III. *Science* **1992,** *258*, 434 – 440.

Chapter 4

A Human MUTYH Variant Linking Colonic Polyposis to Redox Degradation of the $[4Fe4S]^{2+}$ Cluster

Adapted from: McDonnell, K.J.; Chemler, J.A.; Bartels, P.L.; O'Brien, E.; Marvin, M.L.; Ortega, J.; Stern, R.H.; Raskin, L.; Li, G.; Sherman, D.H.; Barton, J.K.; Gruber, S.B. A Human MUTYH Variant Linking Colonic Polyposis to Redox Degradation of the $[4Fe4S]^{2+}$ Cluster. **2018**, *Submitted*.

P. Bartels carried out all electrochemistry and UV-visible, circular dichroism, and EPR spectroscopy. E. O'Brien assisted with electrochemistry. K. McDonnell and J. Chemler identified the C306W mutation, purified proteins, and carried out characterization of the latter (activity assays, biolayer interferometry, and iron quantification).

Introduction

In cells sustaining oxidative damage, genomic guanine residues may be oxidized to 8-oxo-7,8 dihydroguanine (8-oxoG). Unlike guanine, 8-oxoG can pair effectively with either cytosine or adenine bases, with potentially serious mutagenic consequences (*1, 2*). A DNA glycosylase conserved among species from bacteria to humans, known in humans as MUTYH, removes adenine from these 8-oxoG:A mispairs as part of the base excision repair pathway. In humans, germline *MUTYH* mutations that impair enzymatic activity lead to an increase in G:C to T:A transversions that have been shown to result in missense mutations in the *APC* tumor suppressor gene. Mutations in *APC* are the first recognizable genetic events that initiate malignant transformation of normal colonic epithelia into polyps, specifically adenomas, prior to the acquisition of other mutations that complete the neoplastic conversion sequence from normal tissue through adenoma to carcinoma (*3, 4*). Bi-allelic mutations of *MUTYH* give rise to the autosomal recessive cancer genetic syndrome, *MUTYH*-associated polyposis, MAP (*5-7*). Typically, by their fifth decade, MAP patients develop 10-100 colonic polyps (*8-10*). *MUTYH* variants are common with a prevalence of at least 1-2% among Western Europeans (*11*), and the colorectal cancer risk increases nearly 2- and 100-fold for mono- and bi-allelic *MUTYH* mutations, respectively (*12, 13*).

The MUTYH protein comprises three major regions (*14*): the N-terminus, which contains the Endonuclease III 6-helix barrel catalytic domain, the interdomain connector (IDC), and the C-terminus, corresponding to protein residues 1-306, 315-366, and 368-500, respectively. The MUTYH N-terminus contains a $[4Fe4S]^{2+}$ cluster ligated by four cysteine residues; though metabolically expensive, the cluster is conserved in MutY homologues across all domains of life (*15*). A rare exception is the yeast endonuclease III homologue Ntg1, which has no cluster, but

one is present in a second yeast homologue (Ntg2) (*16*). Indeed, the only organisms lacking a cluster in any MutY homologue are specialized anaerobes subjected to lower levels of oxidative stress (*17*). Studies performed on *Escherichia coli* MutY and its homologue endonuclease III (EndoIII) have demonstrated that the cluster is unnecessary for structural integrity and is largely redox inert in solution (*18, 19*). However, when *E. coli* MutY and EndoIII were incubated on duplex DNA-modified gold electrodes, a reversible redox signal centered near 80 mV versus NHE was observed for both proteins and identified as the $[4Fe4S]^{3+/2+}$ couple, an assignment supported by EPR spectroscopy (*20*). Subsequent experiments with EndoIII on a graphite electrode in the presence and absence of DNA revealed that binding to the negatively charged DNA backbone shifts the redox potential of these proteins by about -200 mV, activating the cluster toward oxidation and resulting in a significant increase in binding affinity of the oxidized $[4Fe4S]^{3+}$ form of the enzyme relative to the native $[4Fe4S]^{2+}$ form (*21*).

These studies have led to a model in which [4Fe4S] BER proteins with similar DNA-bound redox potentials use reversible redox exchanges to signal to one another across the genome, taking advantage of the unique ability of DNA to conduct charge across the π-stacked base pairs (bps) in a process known as DNA-mediated charge transport (DNA CT) (*22*, *23*). DNA CT has both a very shallow distance dependence and an exquisite sensitivity to even slight disruptions in base pair stacking, making it an ideal lesion reporter. In our CT signaling model, oxidative stress generates highly reactive species, such as guanine radicals, which can then oxidize proteins including MutY (*24*). If another [4Fe4S] protein is bound at a distal site and the intervening DNA is undamaged, it can send an electron through the DNA to reduce the first protein. Upon reduction, the protein's affinity for DNA is decreased and the protein dissociates to another region of the genome, while the oxidized protein remains bound. In the presence of a

lesion, DNA CT is impaired, and the oxidized protein will remain bound and diffuse toward the site of damage. Thus, DNA CT constitutes a means for [4Fe4S] proteins to scan a vast genome on a relevant time scale and redistribute in the vicinity of lesions. In the case of *E. coli*, long-range signaling by DNA CT has been estimated to reduce the damage search time from 45 minutes to 10 minutes or less (*25*).

In the present study, we describe a novel MUTYH variant, p.C306W, discovered in a patient exhibiting colonic polyposis. This mutant was isolated using an *E. coli* overexpression system along with WT and the well-characterized mutants Y179C and G396D, and electrochemistry, UV-visible and EPR spectroscopy were used to compare the redox properties of these four MUTYH variants. Enzymatic activity and DNA binding parameters were compared using glycosylase assays and biolayer interferometry (BLI), respectively. Together, these results provide strong evidence for a primary function of the [4Fe4S] cluster in DNA-mediated redox signaling and establish MUTYH C306W as a pathogenic variant, enhancing our understanding of the role of the [4Fe4S] cluster in human disease.

Materials and Methods

Determination of Trans Chromosomal Configuration of MUTYH Gene Variants

Germline DNA was amplified using the polymerase chain reaction (PCR) to generate a 935 base pair amplicon that includes the open reading frame positions c.918C>G (p.C306W) and c.1187G>A (p.G396D). The PCR reaction used the forward primer 5′-CCA GGA GAT TTC AAC CAA GC-3′ and the reverse primer 5′ -AAG GGT CAA GGG GTT CAA AT-3′. The c.1187G>A mutation creates a unique BglII restriction endonuclease site which allowed generation of a 719 base pair fragment from the parent 935 base pair amplicon. The shorter 719 base pair fragment was resolved using agarose gel electrophoresis, purified, and its DNA sequence determined (University of Michigan (U-M) Sequencing Core) to establish the identity of the c.918 position.

Identification of APC gene G:C → T:A Transversions

Tumor DNA was extracted from a formalin-fixed, paraffin-embedded colonic adenoma originating from the proband using the RecoverAll Total Nucleic Acid isolation kit (Ambion). A portion of the mutation cluster region of the *APC* gene (*26*) was amplified using PCR with the forward primer 5′-TGC CAC AGA TAT TCC TTC ATC A-3′ and the reverse primer 5′-CAT GGT TTG TCC AGG GCT AT-3′. The PCR product was subsequently sequenced (U-M Sequencing Core).

Cloning of wild type and Mutant MUTYH Expression Plasmids

A plasmid containing the open reading frame for the beta3 isoform of *MUTYH* (NM_001048174.1) was obtained from OriGene (Catalog #: RC201376-OR, Rockville, MD.) and used as a template for the cloning of derivative constructs. Mutant *MUTYH* construct synthesis was accomplished using PCR-based site-directed mutagenesis of wild type MUTYH.

The wild type and mutants were then cloned as maltose binding protein (MBP) fusions into pMCSG19 (*27*) between the KpnI and XbaI restriction sites to increase protein solubility (*28*). The plasmids were further modified by removing the first fourteen codons encoding the mitochondrial recognition sequence, which alleviated heterologous protein toxicity. Furthermore, to decrease the capture of truncated heterologous protein, the N-terminus His_6 tag located between the MBP sequence and the TEV cleavage site was removed and a C-terminus His_{10} tag was attached with a flexible $(SG)_7$ linker to increase solvent exposure.

Preparation of Protein

Heterologous MUTYH proteins in *Escherichia coli* strain BL21(DE3) were initially purified in accordance with previously published protocols using nickel affinity chromatography with 1 mM DTT (*29*). To improve yields and purity, the expression plasmids were transformed into the *E. coli* expression strain for toxic proteins, BL-AI (Invitrogen) also harboring the rare codon plasmid pRARE2-CDF (*30*). One liter of fresh Terrific Broth modified with 4% glycerol and 50 µg/mL of antibiotics (ampicillin and streptomycin) in a three L baffled flask was inoculated with 25 mL of overnight cultures in the same medium. Cultures were grown at 37 ºC in a horizontal shaker at 175 rpm until the OD_{600} reached approximately 2.5. The temperature was adjusted to 15 ºC, and after 90 minutes, 0.25 mM IPTG and 0.2% arabinose were added. After 12-16 hours, cells were harvested by centrifugation, flash frozen in liquid nitrogen, and stored at -80 °C until processing. Cell pellets were thawed in an ice bath and re-suspended in 80 mL of ice cold 10% glycerol before the addition of 53 mg/mL of CelLytic Express (Sigma), one tablet of Protease Inhibitor Cocktail (Sigma) and 20 mM imidazole. The samples were clarified on a nutator for 30 minutes at 4ºC before the addition of 20 mM of β-mercaptoethanol. The crude cell lysate was passed through a 0.45 µm filter in preparation for nickel affinity chromatography.

His-tagged proteins were loaded onto a 5 mL HisTrap HP column (GE Healthcare) at 2 mL/min using an AKTA Explorer FPLC instrument (GE Healthcare) at 4 ºC. The columns were first washed with 20 column volumes of 93% Buffer A (20 mM Tris-HCl, pH 7.4, 1 M NaCl, 20 mM β-mercaptoethanol, 10% glycerol) and 7% Buffer B (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 500 mM imidazole, 10% glycerol) and the proteins were eluted using a 7-100% Buffer B gradient over 10 column volumes. Fractions containing MBP-MUTYH protein (as determined by SDS-PAGE), were pooled. Typical yields of purified protein for MBP-MUTYH wild type, Y179C and G396D were between 10-20 mg from one liter cultures and 0.5-1.5 mg of soluble protein was obtained for MUTYH p.C306W. Monomeric protein was obtained by size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare) in Buffer C (20 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, 10% glycerol). Fractions eluting near the expected molecular weight (104 kD) were collected, partitioned into aliquots, flash frozen in liquid nitrogen, and stored at -80 ºC until further use.

DNA Substrates for Glycosylase and Binding Assays

All oligonucleotides (Integrated DNA Technologies, Coralville, Iowa) were purchased PAGE purified. Duplexes were obtained by heating 50 µl 50 µM complementary strands at 85 ºC then decreasing the temperature by 0.5 ºC every 30 seconds until attaining room temperature. The FAM labelled 8-oxoG:A duplex used in the DNA glycosylase assay consisted of 5′-ACA AAG AAC TTA TAG CTC CTC CTT GAG CAC ACA GAG GTG TTC GAT GTA GTT G/A/C GCA GGA CGG GTT CAG T/6-FAM/-3′ and 3′-TGT TTC TTG AAT ATC GAG GAG GAA CTC GTG TGT CTC CAC AAG CAT GAT CAA C/8oxoG/G CGT CCT GCC CAA GTC A-5′. The biotin labeled 8-oxoG:A duplex used in the binding assay consisted of 5′-

/BiotinTEG/AC AAA GAA CTT ATA GCT CCT CCT TGA GCA CAC AGA GGT GTT CAT GTA GTT G/A/C GCA GGA CGG GTT CAG T-3′ and the 8-oxoG oligomer.

DNA Glycosylase Assay

The DNA glycosylase assay was adapted as previously reported (*29, 31-32*). The activity was evaluated by providing 10 nM of DNA substrate containing a single 8-oxoG:A mismatch to wild type or mutant MBP-MUTYH proteins (0-1000 nM) at 37ºC in a buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, and 0.1 mg/mL BSA). Reactions were quenched after 1 hour with 80 mM NaOH followed by heating the samples to 90 ºC for 4 minutes, cooled, and then diluted with three volumes of formamide spiked with GeneScan™ 500 LIZ™ Size Standard (Invitrogen). DNA fragmentation was determined employing capillary electrophoresis (U-M DNA Sequencing Core, ABI 3730 DNA Analyzer). Traces were analyzed using Peak Scanner™ Software (version 1.0, Applied Biosystems). The percent of excised DNA was calculated as the ratio of the 6-FAM peak area migrating at 44 oligonucleotides to the total peak area (at 44 and 66 oligonucleotides).

Multiple Turnover Assay: Active Site Titration

The multiple turnover assay was adapted as previously reported (*29, 31-32*). Reactions were analyzed for scission of 10 nM FAM-labeled 8-oxoG duplex DNA after the addition of MUTYH protein. The total protein concentrations, selected to give a burst amplitude in a detectible range, were 25, 2670, 500, and 25 nM of MUTYH wild type, Y179C, C306W, and G396D, respectively. Samples were drawn over a 20-minute time course and processed as described above. The cleaved product concentration, [P], was fitted with Equation 1 to determine the amplitude of the burst (A_0), k_B (rate constant during the burst phase) and k_L (rate constant for the linear phase).

Equation 1: $[P] = A_0[1-\exp(-k_B t)]+k_L t$

The percent of active protein was calculated as a ratio of the A_0 to total protein concentration.

Binding Kinetics (Biolayer Interferometry)

All biolayer interferometry measurements were made on an Octet RED instrument (Pall ForteBio, Menlo Park, CA) using streptavidin (SA) biosensors (*33*). Assays were performed in 96-well black microplates at 25 ºC and 1000 rpm. All volumes were 200 µL. All proteins were buffer exchanged using PD-10 columns (GE Healthcare) pre-equilibrated with PBS then serial diluted (25, 12.5, 6.25, 3.125, 0.78125 nM) into working volumes with 1X Kinetics Buffer (Pall FortBio;10 mM Phosphate, pH 7.4, 150 mM NaCl, 0.02% Tween-20, 1 mg/mL BSA). The biotinylated duplex DNA was first immobilized onto the SA biosensors for 300 seconds and then equilibrated in 1X Kinetics Buffer for 300 seconds. Protein association was performed for 150-300 seconds followed by dissociation into 1X Kinetics Buffer for 900 seconds. A reference sensor with immobilized dsDNA was subtracted from each data set. Shift data was analyzed with ForteBio's Analysis software (version 7.1). Kinetic parameters k_{on} and k_{off} and affinity (K_D) were determined from a global non-linear regression of association and dissociation binding kinetics using a 1:1 Langmuir binding model.

Fe Elemental Analysis

The presence of elemental Fe within MUTYH protein samples was determined using a Thermo Scientific Element2 ICP-HRMS (*34*). Purified proteins were dialyzed overnight to remove glycerol and to allow equilibration with blank buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl and 1 mM DTT) using a 10K MWCO Slide-A-Lyzer dialysis cassettes (Thermo Scientific).

DNA Synthesis and Purification for Electrochemistry

Thiol, FA, and OG modified DNA strands were prepared on an automated DNA synthesizer (Applied Biosystems) and purified by HPLC on a PLRPS column (Agilent) as described in previously published protocols (*35*); unmodified strands were ordered from IDT and purified by HPLC. For electrochemistry, 50 μL 50 μM complementary DNA strands were degassed and annealed in storage buffer (5 mM sodium phosphate, pH 7.0, 50 mM NaCl) in equimolar amounts by a 5-minute incubation at 95ºC followed by slow cooling (1.5 hours) to RT on a thermocycler. Well-matched (WM) duplex DNA, DNA with an FA:OG lesion, and substrates containing an abasic site were all prepared in this way. Duplex sequences were as follows:

WM DNA

5′ – ACT GAA CCC GTC CTG CGT CAA CTA CAT GAA CAC CTC – 3′
3′ – TGA CTT GGG CAG GAC GCA GTT GAT GTA CTT GTG GAG – 5′ – C6 Thiol

FA:OG DNA

5′ – ACT GAA CCC GTC CTG CG**OG** CAA CTA CAT GAA CAC CTC – 3′
3′ – TGA CTT GGG CAG GAC GC**FA** GTT GAT GTA CTT GTG GAG – 5′ – C6 Thiol

Abasic DNA

5′ – ACT GAA CCC GTC CTG CGT CAA CTA CAT GAA C**Ab**C CTC – 3′
3′ – TGA CTT GGG CAG GAC GCA GTT GAT GTA CTT GTG GAG – 5′ – C6 Thiol

OG = 8-oxoguanine, FA = 2′-fluoroadenine, Ab = abasic site

Electrochemistry on DNA Self-Assembled Monolayers (SAMs)

Electrochemical characterization of MUTYH was carried out on a multiplexed chip platform consisting of 16 individually-addressable gold electrodes separable into four quadrants (*35*). Self-assembled DNA monolayers were formed by adding 25 μL 25 μM duplexed DNA in phosphate buffer (5 mM sodium phosphate, pH 7.0, 50 mM NaCl) to each quadrant of the chip

and incubating overnight. After monolayer formation, gaps in the film were eliminated by backfilling for 45 minutes at RT with 1 mM 6-mercapto-1-hexanol in phosphate buffer with 5% glycerol. The surface was then extensively rinsed in phosphate buffer, followed by protein storage buffer (described below). To compare different monolayer morphologies, DNA was incubated in phosphate buffer as described to generate low-density DNA monolayers (surface coverage of ~15 pmol/cm^2) or in the presence of 100 mM $MgCl_2$ to form high-density monolayers (surface coverage of ~40 pmol/cm^2) (*36, 37*). Bulk electrolysis experiments were performed using gold rod electrodes in a custom-made electrochemical cell. Experiments were carried out in air unless otherwise noted; anaerobic experiments were performed in a glove bag (Coy) under a 95% N_2/5% H_2 atmosphere.

MUTYH concentration was determined by UV-vis, using an extinction coefficient of 17000 $M^{-1}cm^{-1}$ at 410 nm to determine [4Fe4S] cluster concentration and 102330 $M^{-1}cm^{-1}$ at 280 nm to determine total protein concentration; cluster loading was determined by dividing [4Fe4S] cluster concentration by total protein concentration, and was typically around 15%. Initial characterization was carried out in Tris storage buffer (20 mM Tris, pH 7.4, 100 mM NaCl, 1 mM DTT, 10% glycerol v/v), while later electrochemical and spectroscopic experiments used a HEPES buffer (20 mM HEPES, pH 7.4, 100 mM KCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v). MUTYH was transferred into HEPES using Amicon 10 kDa MW cutoff spin tubes (Millipore Biomedicals) at 4^{o}C.

Once in an appropriate buffer, MUTYH was added to a multiplexed chip and incubated for several hours with cyclic and square wave voltammetry (CV and SQWV, respectively) scans taken once per hour. In typical experiments, CV scans were taken in a potential window of -0.188 to 0.412 V vs NHE at a scan rate of 100 mV/s, while SQWV scans were taken at a

frequency of 15 Hz with 0.025 V amplitude. To plot the scan rate dependence of CV current, scans in the same window were carried out at 20, 50, 80, 100, 200, 500, 750, and 1000 mV/s. All experiments were performed on a CH Instruments potentiostat with a Ag/AgCl reference in 3 M NaCl and Pt wire counter electrode. Potentials were converted to NHE by adding 212 mV to the measured potentials, accounting for both the salt concentration (209 mV according to BASi®) and ambient temperature (*38*). Bulk electrolysis was carried out at 0.412 mV versus NHE, and yields were estimated by subtracting the total charge passed with only buffer present from that passed when MUTYH was included. All buffer components were purchased from Sigma-Aldrich, the Ag/AgCl reference electrode was purchased from BASi®, and the Pt wire counter electrode was purchased from the Kurt J. Lesker Company.

EPR Spectroscopy

Continuous wave X-band EPR was carried out at 10 K on a Bruker EMX instrument. Samples were prepared aerobically, using 150 μL 5 – 15 μM MUTYH in parallel with a storage buffer blank. Spectra were taken from the summation of 9 sweeps at 12.88 mW microwave power, 2 G modulation amplitude, and a receiver gain of 5.02 $x10^3$.

Results

Identification and Functional Deficiencies of a Novel MUTYH Variant

A novel germline *MUTYH* variant, c.918C>G (p.C306W), together with the previously well-described *MUTYH* mutation, c.1187G>A (p.G396D), were identified in a patient with colonic polyposis and a family history significant for early age colon cancer. The cysteine at position 306 represents one of the four cysteine residues that mediate integration of the conserved MUTYH [4Fe4S] cluster. In bacterial MutY cluster loss is associated with decreased protein function (*19*) which suggests that in MUTYH the C306W variant may affect the integrity of the [4Fe4S] cluster and represent a pathologic mutant (Figure 4.1 illustrates the structure of bacterial MutY and identifies corresponding residues in MUTYH).

Supporting the potential pathogenicity of the novel c.918C>G variant, we established that this variant was situated in a trans chromosomal configuration relative to the c.1187G>A *MUTYH* mutation (Figure 4.2). Further evidence of the pathogenic nature of the c.918C>G variant was apparent in the sequencing of the *APC* gene in somatic DNA originating from a colonic adenoma from the patient. This sequencing revealed the presence of a C:G to T:A transversion in the APC gene, which is the hallmark genetic lesion distinguishing deficient *MUTYH*-mediated DNA repair (*4*) (Figure 4.3).

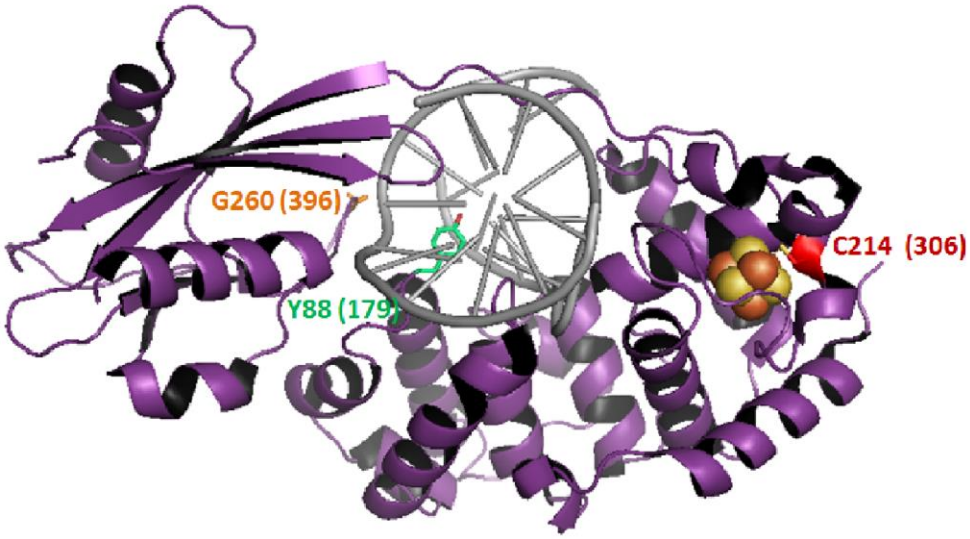


Figure 4.1 Structure of *Geobacillus stearothermophilus* MutY with residues homologous to disease-relevant variants in human MUTYH isoform β3 highlighted. Y179C and G396D have been previously shown to be pathological mutations involved in MUTYH-associated polyposis (MAP); C306W was identified in this study. Notably, C306 is one of four cysteine residues involved in [4Fe4S] cluster ligation. Structure from PDB ID 1RRQ.

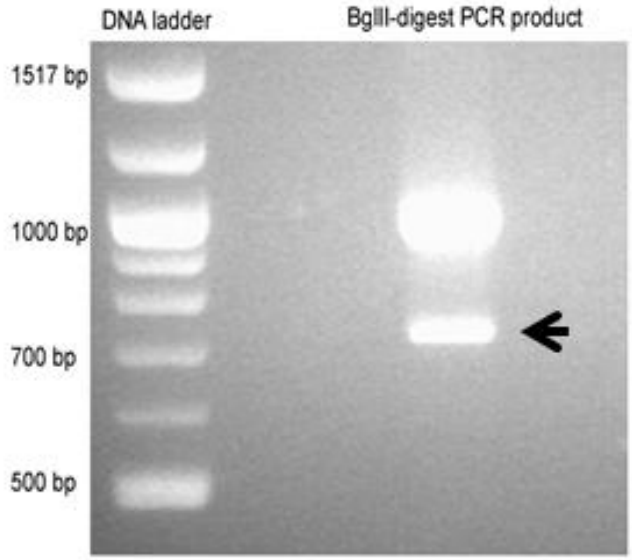


Figure 4.2 The c.918C>G MUTYH variant is situated trans relative to the c.1187G>A MUTYH mutation. The c.1187G>A (p.G396D) *MUTYH* mutation creates a unique BglII restriction enzyme site which we employed to isolate the chromosomal DNA strand containing the c.1187G>A *MUTYH* mutation. A 957 bp fragment encompassing the *MUTYH* open reading frame nucleotide positions 918 and 1187 was generated with PCR and then restriction enzyme digested with BglII. The digestion products were resolved on 1.5% agarose gel and the lower molecular weight DNA band isolated (arrow). Sequencing of the lower molecular weight band demonstrated the wildtype cytosine nucleotide at position 918 consistent with a trans configuration of the c.1187G>A and c.918C>G *MUTYH* alterations.

Normal Colonic Tissue

Colonic Adenoma

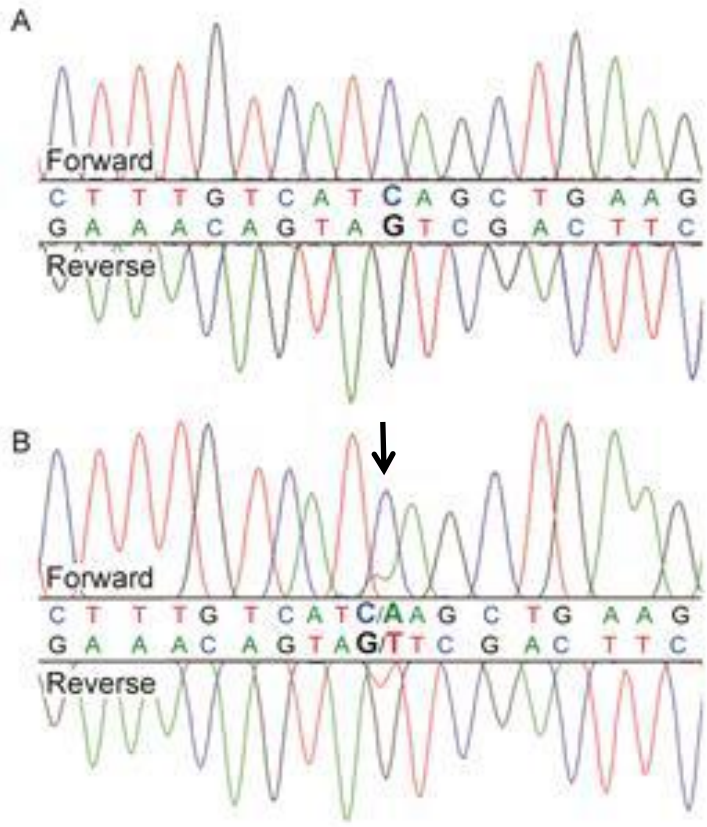


Figure 4.3 Trans c.1187G>A (p.G396D) and c.918C>G (p.C306W) MUTYH mutations are associated with the signature G:C to T:A transversion in the APC gene distinctive of deficient MUTYH DNA repair. From one of the patient's adenomatous polyps we extracted DNA to assess for the presence of the signature G:C to T:A nucleotide transversion which results from deficient MUTYH DNA repair. The figure depicts a portion of the germline nucleotide sequence from the mutation cluster region of the APC gene (upper tracing); this sequence reveals the normal wildtype *APC* sequence. In comparison, the lower tracing depicts the forward and reverse nucleotide sequence from DNA extracted from one of the patient's colonic polyps for this same mutation cluster region of the *APC* gene; there is identified a G:C to T:A transversion (arrow) resulting in a premature stop codon in the *APC* open reading frame and pathologic APC protein product.

Preparation of Monomeric MUTYH Protein

The four MUTYH proteins, WT, G396D, Y179C, and C306W were overexpressed in *E. coli* and purified by nickel affinity chromatography followed by size exclusion FPLC (Figure 4.4). All of the MUTYH proteins eluted as soluble aggregates in the void fraction; evidence of aggregation in these samples was also clear from the UV-visible spectra. Aggregated WT, G396D and Y179C MUTYH proteins demonstrated glycosylase activity consistent with previous reports, with attenuated activity observed in the G396D and Y179C MUTYH proteins (*29-31*); aggregated C306W protein lacked glycosylase activity (Figure 4.5a).

Aggregated protein is unlikely to exhibit native activity levels; thus, 20 mM β-mercaptoethanol was added prior to purification in an effort to disperse the aggregates and yield monomeric protein. This additional measure was effective, and monomeric proteins were of expected molecular weight (104kD) and eluted around 12 minutes during the size exclusion FPLC step (Figure 4.4). Relative to aggregated protein, monomeric WT demonstrated enhanced glycosylase activity. The glycosylase activities of the monomeric G396D and Y179C MUTYH proteins were also enhanced relative to the aggregated species, however they displayed less activity compared with WT protein. Monomeric C306W MUTYH protein remained severely deficient in glycosylase activity (Figure 4.5b). Based on the higher catalytic function of native, monomeric MUTYH protein was employed for all subsequent experiments.

For a more complete functional comparison of the four MUTYH variants, time course glycosylase assays were conducted under multiple turnover conditions to quantitatively determine the proportion of active enzyme in each sample (*29, 31-32*). These assays were performed using dsDNA containing an 8-oxoG:A mispair together with varying concentrations of MUTYH proteins. The experimental results demonstrate an initial burst of adenine excision

activity proportional to the active fraction, A_o, of the protein sample (Figure 4.5c, Table 4.1). The rate constants, k_B and k_L, were determined for the exponential and linear phases of the reaction, respectively (Table 4.1). Both WT MUTYH and G396D proteins had comparable linear rates for turnover and the highest fraction of active protein. In contrast, the C306W MUTYH mutant was essentially devoid of adenine excision activity and Y179C had no detectable turnover. The fractions of active MUTYH were then employed to correct for the total amount of protein used in the glycosylase assay (Figure 4.5d), confirming that WT MUTYH and G396D mutant had comparable activities, while C306W and Y179C mutants displayed poor activity. Poor activity observed in MUTYH C306W could have two possible explanations: that either this mutant was catalytically inactive or it was unable to bind specifically to DNA (as is the case with low activity in the weakly-bound Y179C).

To help distinguish between these possibilities, we employed biolayer interferometry (BLI) (*33*) to measure the binding parameters of the MUTYH proteins. After preliminary experiments demonstrated that there was no significant difference in binding of WT MUTYH to DNA that contained an 8-oxoG:A mispair relative to DNA without the mispair, we compared the binding of WT MUTYH and the G396D, Y179C, and C306W mutants to DNA containing an 8-oxoG:A mispaired duplex. The binding kinetics data are summarized in Table 4.2. Relative to WT MUTYH, the G396D and Y179C variants demonstrated increasing values of K_D primarily due to decreased association rates. There was no detectable binding for the C306W mutant within the protein concentration range tested, suggesting that the low activity levels observed in this mutant were due primarily to ineffective DNA binding.

Together these data demonstrate the functional deficiency of the C306W MUTYH mutant. The cysteine at position 306 represents one of the four cysteine residues that ligate the

MUTYH [4Fe4s] cluster. In bacterial MutY, [4Fe4S] cluster loss is associated with decreased protein function (*19*), which suggests that the C306W variant of MUTYH may affect the integrity of the [4Fe4S] cluster accounting for this mutant's pathogenicity.

To assess the integrity of the [4Fe4S] cluster, iron loading of the clusters of WT MUTYH and the mutants Y179C, G396D and C306W were compared by quantifying the iron present in each sample using ICP-HRMS for elemental analysis (*34*). Consistent with disruption of the Fe-S cluster loop in the C306W variant, this protein exhibited substantially lower iron content relative to the other MUTYH proteins tested (Table 4.3). However, UV-visible spectra taken from disrupted aggregates of all four variants distinctly showed the broad peak centered at 410 nm that is characteristic of a [4Fe4S] cluster, indicating that MUTYH C306W is still capable of binding an intact cluster and further suggesting that loading by cellular machinery remains effective (Figure 4.6). In addition, circular dichroism (CD) spectra of WT MUTYH and the C306W mutant were indistinguishable, confirming that no global conformational changes were induced by the mutation (Figure 4.7). Thus, the low cluster content as measured by ICP-HRMS was instead tentatively associated with decreased protein stability in this mutant, and subsequent electrochemical and EPR experiments were used to more reliably examine the [4Fe4S] cluster properties in detail.

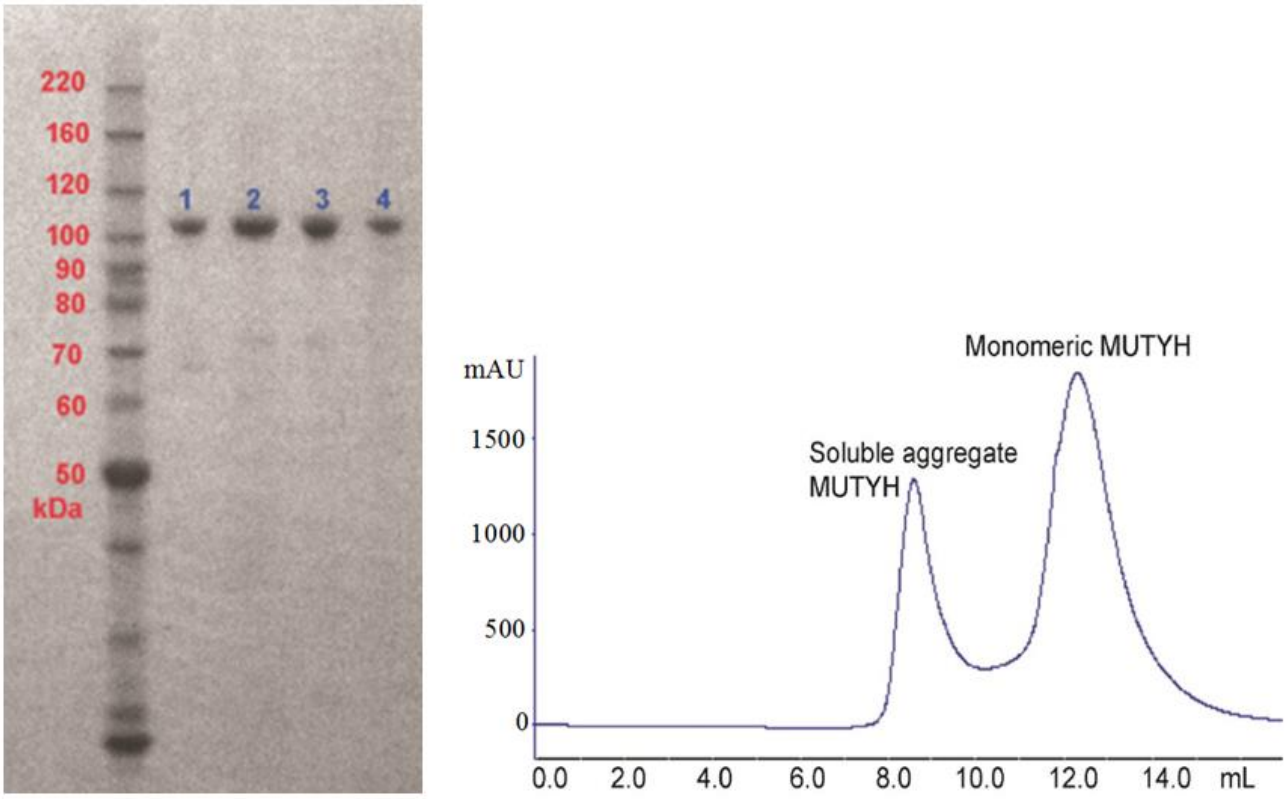


Figure 4.4 Purification of monomeric MUTYH proteins. (**Left**) SDS-PAGE gel of purified monomeric MUTYH proteins. The left lane contains the protein molecular weight ladder. The lanes 1-4 contain 1ug of MUTYH wild type, Y179C, C306W, and G396D, respectively. (**Right**) SEC-FPLC UV_{260} trace of WT MUTYH following treatment with 20 mM β-mercaptoethanol. While some soluble aggregates still occur in the void volume, the majority of the purified protein is monomeric.

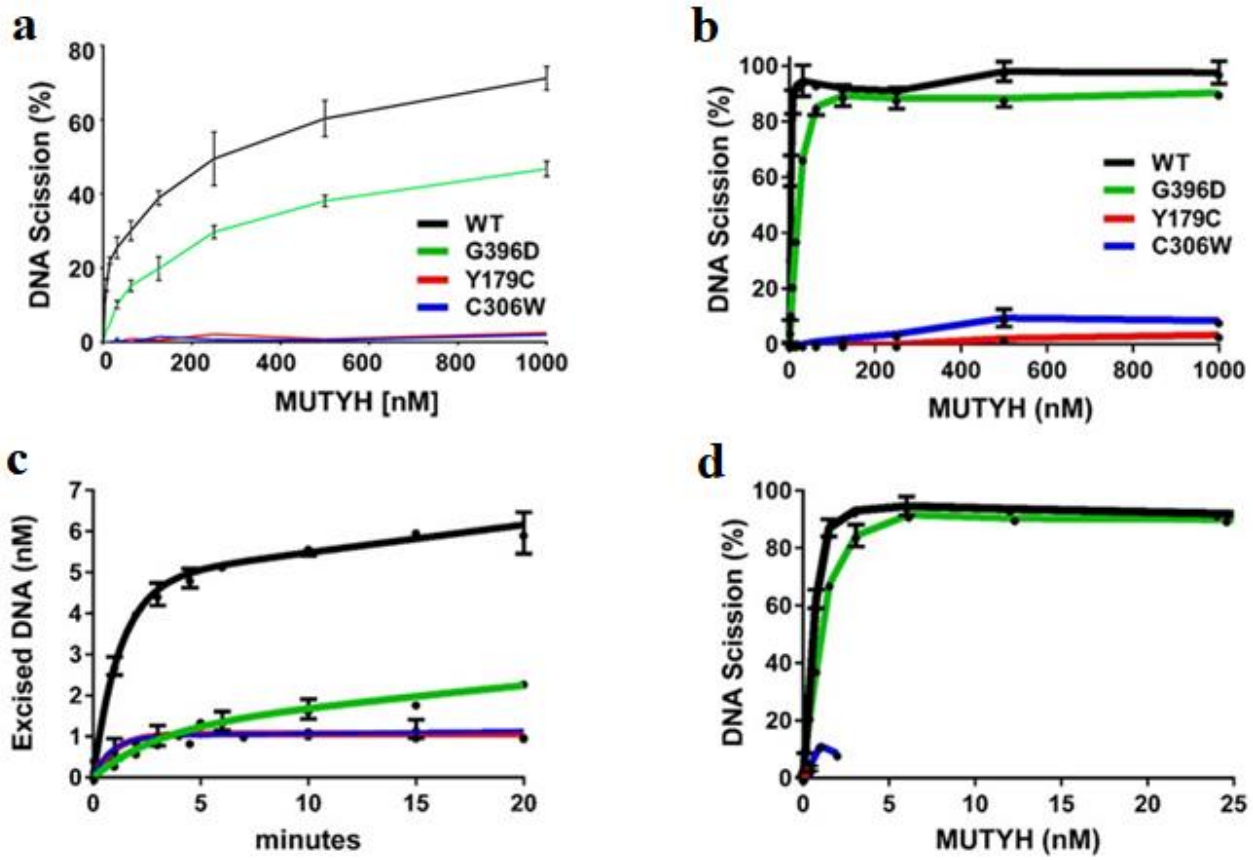


Figure 4.5 A novel human MUTYH variant, C306W, lacks glycosylase activity. Black lines, wild type MUTYH; Green lines, G396D; Red lines, Y179C; Blue lines, C306W
All data are presented as mean $\pm$ s.d., $n = 3$.
(a) Glycosylase assays with soluble MUTYH aggregates. Activity levels for WT MUTYH and the variants Y179C and G396D are comparable to previously reported values. MUTYH C306W shows only minimal activity, comparable to Y179C.
(b) Glycosylase assay using soluble, monomeric MUTYH. The pattern matches that of aggregated protein, and confirms the observed defect in MUTYH C306W. Relative to WT and G396D MUTYH, Y179 and C306 proteins demonstrate severely attenuated DNA scission activity.
(c) Multiple turnover reaction conditions define the concentration of active protein within a purified protein sample. The glycosylase assay was performed with sufficient MUTYH protein to generate reaction burst amplitudes (A_o) within the detectable range. MUTYH active fractions, A_o, and k_B and k_L rate constants of the excision reaction during the exponential and linear phase, respectively, were determined by fitting the curves from the reactions at top to the equation $[P]= A_o\,[1-\exp(-k_B)t] + k_L t$.
(d) Adenine excision activity of wild type and mutant MUTYH proteins after correcting for active fractions, A_o. The correction for active MUTYH C306W and Y179C proteins shifted their assay concentrations below 2.5 nM.

Table 4.1 Determination of Rate Constants from Multiple Turnover Assay

Enzyme ID	[protein] (nM)	A_0	A_0/[protein]	k_B (min^{-1})	k_L (min^{-1} nM)
Wild type (WT)	25	4.8 ± 0.2	19.2%	0.8 ± 0.1	0.07 ± 0.01
Y179C	2670	1.1 ± 0.1	0.04%	1.0 ± 0.2	0
C306W	500	1.0 ± 0.1	0.2%	1.2 ± 0.2	0.01 ± 0.01
G396D	25	1.2 ± 0.2	4.9%	0.3 ± 0.1	0.05 ± 0.01

Data show mean ± s.d., $n = 3$

Table 4.2 Enzyme/DNA Kinetic Binding Data Obtained from Biolayer Interferometry.

Enzyme ID	k_{on} (1/Ms) x 10^4	k_{dis} (10^{-4} 1/s)	K_D (10^{-9} M)
WT	43 ± 0.4	2.7 ± 0.03	0.6 ± 0.01
Y179C	2.9 ± 0.4	5.0 ± 0.21	17 ± 2.57
C306W	No binding	-	-
G396D	10 ± 0.5	5.3 ± 0.12	5.2 ± 0.27

Data show mean ± s.d., *n* = 3

Table 4.3 Elemental iron analysis of MUTYH proteins by ICP-HRMS.

Enzyme ID	[Fe] (μM)	[MUTYH] (μM)	% Ratio 4Fe/ Enzyme
WT	6.42 ± 0.32	1.36	115%
Y179C	6.40 ± 0.32	1.65	95%
C306W	0.78 ± 0.03	1.63	9%
G396D	6.59 ± 0.33	1.62	99%
Buffer	0.17 ± 0.03	-	-

Data show mean ± s.d., $n = 3$

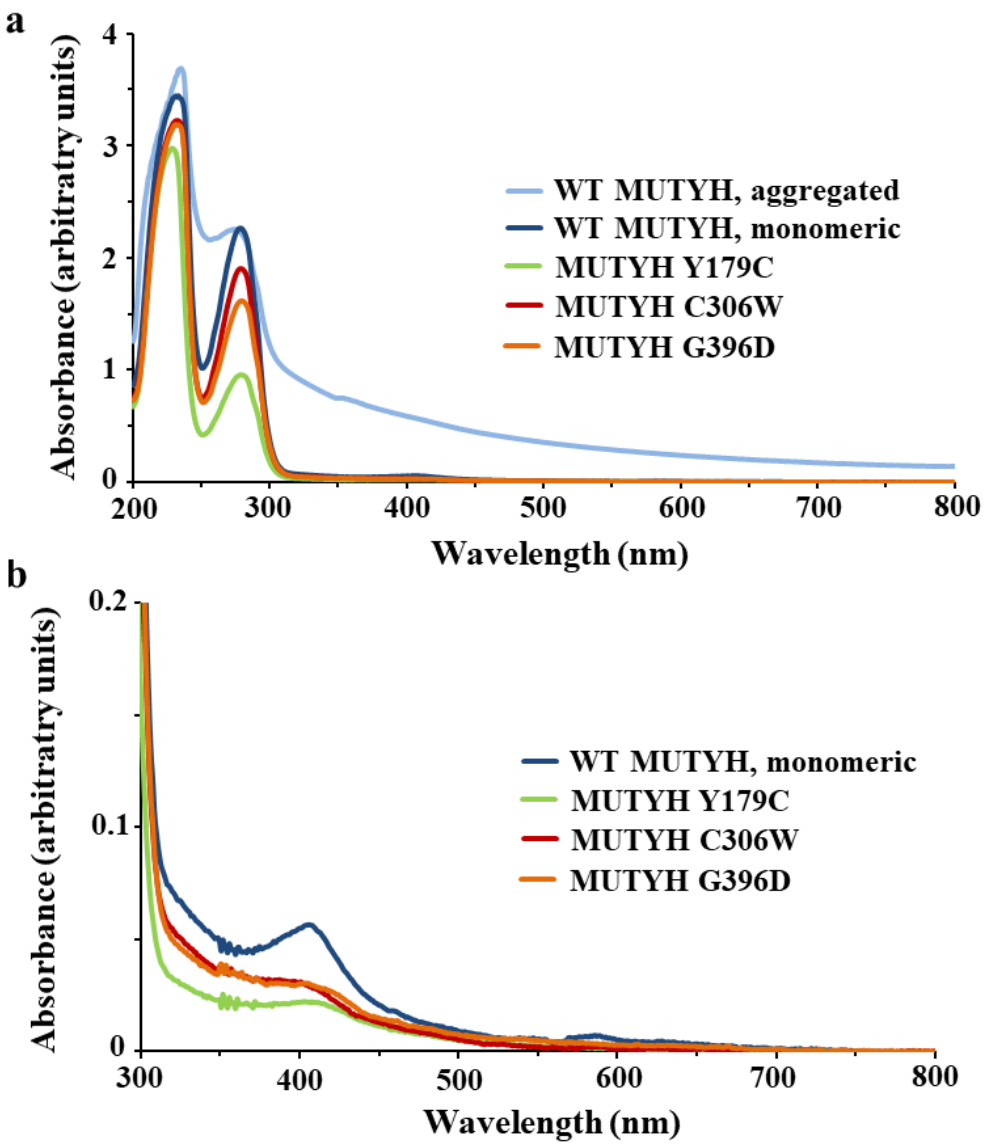


Figure 4.6 MUTYH UV-visible spectra. (**a**) Complete UV-visible absorption spectra of monomeric WT MUTYH and the mutants Y179C, C306W, and G396D. Aggregated MUTYH (light blue) can be distinguished by a U-shaped spectrum with elevated absorbance at 800 nm. (**b**) Zoomed in version of spectra in (a) highlighting the broad peak centered at 410 nm, which is characteristic of $[4Fe4S]^{2+}$ clusters. The presence of this peak confirms that the C306W mutation does not prevent cluster binding, and indeed cluster loading is comparable to that of WT and the other variants studied. All monomeric proteins were scanned in standard storage buffer (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1mM DTT, 0.5 mM EDTA, 10% glycerol v/v), while the aggregated WT spectrum is in a phosphate buffer (20 mM sodium phosphate, pH 7.4, 100 mM NaCl, 0.5 mM EDTA, 10% glycerol v/v).

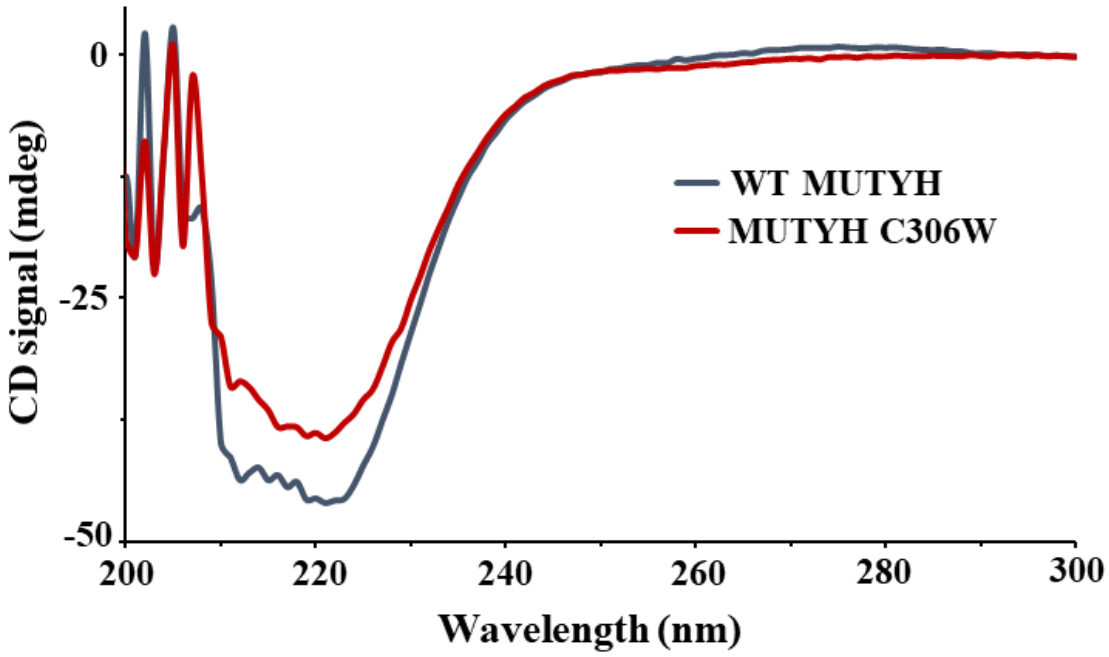


Figure 4.7 Circular dichroism (CD) spectra of WT MUTYH and MUTYH C306W. The similarity in spectra suggest that the C306W mutation does not result in significant global structural perturbations. These experiments were performed in HEPES storage buffer (30 mM HEPES, 100 mM KCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, pH 7.4).

DNA-Bound Electrochemistry of WT and Mutant MUTYH

Having observed that MUTYH C306W appeared to incorporate less iron despite its capacity to bind an intact cluster, we next assessed its redox properties on DNA-modified gold electrodes alongside WT, G396D, and Y179C MUTYH (Figure 4.8). Electrochemical analysis was reasoned to be an insightful approach in these studies for two reasons: first, access to highly purified WT MUTYH allowed us to determine if the human protein behaved like its bacterial counterpart, and, second, electrochemical monitoring would provide an effective way to assess the stability of MUTYH C306W over time. Specifically, the predicted instability of MUTYH C306W was expected to lead to electrochemical signals that were either smaller than WT or less stable over time.

In these experiments, MUTYH was incubated in storage buffer (20 mM Tris, pH 7.4, 100 mM NaCl, 1 mM DTT, 10% glycerol v/v) and periodically scanned by CV and SQWV. Aggregated protein preparations were treated with 20 mM β-mercaptoethanol and exchanged into storage buffer with fresh DTT immediately prior to electrochemical analysis. Cluster loading as determined by UV-visible spectroscopy was low in monomeric WT MUTYH prepared from disassembled aggregates (around 15% by the A_{410}:A_{280} ratio), although all deaggregated MUTYH variants, including C306W, showed similar loading percentages. It is likely that apoproteins can compete to some extent with loaded holoenzyme for DNA binding on electrode surfaces, making low loading potentially problematic for electrochemistry. However, when making direct comparisons between different MUTYH variants, similar [4Fe4S] loading between the proteins studied is a more important factor than absolute loading as long as quantifiable signals can be obtained. Thus, we proceeded to electrochemical experiments despite the low loading values.

To test multiple conditions simultaneously, DNA monolayers were prepared on multiplexed gold electrodes, which enabled up to four experiments to be conducted in parallel (*35*). In this case, half of the available quadrants consisted of unmodified (WM) DNA and half of substrate trap FA:OG DNA included in an effort to enhance signal amplitude by increasing the DNA binding affinity. In addition, both high density (formed in the presence of 100 mM $MgCl_2$) and low density monolayers (formed without Mg^{2+}) were compared on a single chip. DNA surface density is an important parameter for protein experiments (*35*): high density films have more DNA on the surface (30-50 pmol/cm^2) that can improve DNA-mediated signaling by sterically hindering large proteins, while low density films contain less DNA (15-20 pmol/cm^2) that can be more readily accessible to proteins (*36, 37*). Overall, the effect of each type of film is likely to depend strongly on the particular protein being studied, and both have been employed in previous studies.

In initial experiments, WT MUTYH (~2.5 μM [4Fe4S] cluster) yielded a reversible redox signal with a midpoint potential of 106 ± 1 mV versus NHE (Figure 4.8b), which is similar to the 65-95 mV (versus NHE) range reported for other DNA-bound [4Fe4S] proteins. That the MUTYH potential was slightly higher can most likely be attributed to disparate buffer conditions. The signal appeared almost immediately after addition to the electrode and increased in magnitude over the course of a 3-hour experiment. Although small, the signals were readily quantifiable: on low density films containing WM DNA, CV peak areas were 31 ± 1 x 10^{-2} nC for the reductive peak and -33 ± 2 x 10^{-2} nC for the oxidative peak, while the equivalent values on high density films were 25 ± 2 x 10^{-2} nC and -27 ± 3 x 10^{-2} nC, respectively. Interestingly, no significant differences in signal intensity were observed between WM and FA:OG DNA, with CV reductive and oxidative peak charges of 28 ± 1 x 10^{-2} nC and -34 ± 4 x 10^{-2} nC on low

density FA:OG DNA films and 24 ± 2 x 10^{-2} nC and -24 ± 1 x 10^{-2} nC on high density films. The FA:OG substrate trap is known to increase the binding affinity of the very similar murine MUTYH on a 30-mer duplex by an order of magnitude (*39*), and the absence of any clear change in signal intensity suggests that our system is not sufficiently sensitive to detect this difference. Several possible explanations exist for this insensitivity. First, it may be that the absolute amount of accessible DNA on either surface is too low to detect a difference between WM and FA:OG substrates. Alternatively, the significant amount of apoprotein present may have blocked some of the accessible DNA from fully loaded protein. Finally, the fact that DNA is tethered to a surface in these experiments rather than in solution might lower the chances of a successful binding event.

The potential and maximum signal size of MUTYH C306W (~2.5 μM [4Fe4S] cluster) were comparable to WT, although peak area was considerably more variable between experiments. Contrary to expectations, the C306W mutant was not obviously CT-deficient, but the signal did steadily decrease in size after 1-2 hours of incubation. The observed signal loss is consistent with the cluster lability suggested by ICP-HRMS. In addition, a second, irreversible peak centered around -50 mV versus NHE appeared and increased in size as the reversible $[4Fe4S]^{3+/2+}$ signal decayed (Figure 4.8b). This pattern of reversible signal loss and secondary peak growth was observed in several experiments with C306W and did not occur to any appreciable extent in WT MUTYH, confirming that it was not a contaminant. A secondary peak at this potential was unprecedented among previously studied base excision repair proteins, but its growth along with the parallel loss of the reversible signal suggested that it was some form of degradation product.

While the DNA processing enzymes studied thus far have generally shown stabilization of the $[4Fe4S]^{3+}$ form upon DNA binding to yield a reversible $[4Fe4S]^{3+/2+}$ signal on an electrode, loss of iron by the oxidized $[4Fe4S]^{3+}$ species to form the $[3Fe4S]^{+}$ cluster has been reported in bacterial MutY and EndoIII when the samples were frozen for EPR under aerobic conditions (*20*). As the loss of one iron atom is the first step in cluster degradation, we considered the $[3Fe4S]^{+}$ cluster to be a likely candidate for the identity of the unexpected redox-active MUTYH C306W species. At ~ -50 mV versus NHE, the MUTYH C306W secondary peak fell within the range of reported $[3Fe4S]^{+}$ cluster potentials (*40*), supporting this assignment. The irreversible nature of the signal was unusual, given that [3Fe4S] clusters can access a reversible $1^{+}/0$ redox couple, but, given the significant impact of even a single unit of cluster charge on DNA binding affinity (*21*), irreversibility in our experiments can be rationalized as protein dissociation from DNA following reduction to the neutral $[3Fe4S]^{0}$ state.

Confirming that the [4Fe4S] cluster degradation observed in MUTYH C306W was unique to this mutant, electrochemical analysis of DNA bound Y179C and G396D (both 2.5 μM [4Fe4S] cluster) yielded reversible signals at nearly the same potential as WT with no secondary peak present (Figure 4.8b, Table 4.4). Like WT, the signals from both of these variants increased over time and remained stable for several hours. Notably, the Y179C signal was only about half as large as WT, consistent with the lower DNA binding affinity of this mutant relative to both WT and G396D (K_D of 7.5 nM for Y179C versus 2.2 and 4.9 nM for WT and G396D, respectively) (*29*). Overall, both mutants were more similar to WT than C306W in their redox properties, an unsurprising outcome given that the cluster in these variants is unaltered. The propensity for the MUTYH C306W cluster to degrade by redox activity provides a possible explanation for the low DNA binding affinity observed with BLI binding studies, as previous

work with *E.* coli MutY has demonstrated that apoprotein lacking cluster is defective in DNA binding despite remaining structurally intact (*19*).

Characterization of the C306W Degradation Product

In an effort to characterize the MUTYH C306W cluster degradation product more fully, we proceeded to assess its dependence on oxygen, which is often an important factor in $[3Fe4S]^{+}$ cluster formation (*20*). Specifically, electrochemical and UV-visible spectroscopic analysis were employed to compare aerobically oxidized proteins with those maintained in an anaerobic environment. For an effective comparison, a single C306W sample was concentrated and one half was diluted to 2.5 μM (by [4Fe4S] cluster) in degassed buffer and placed on a chip containing low density WM DNA in an anaerobic glove bag (95% N_2/5% H_2 atmosphere), while the other half was maintained in aerobic conditions and oxidized on a DNA-modified gold rod electrode held at 0.412 V versus NHE. The anaerobic sample was scanned periodically by CV and SQWV, and, following electrolysis, the oxidized sample was transferred to the glove bag and added to a separate quadrant on the same chip. In addition to electrochemistry, UV-visible spectra were recorded before and after electrolysis both to observe changes in the 410 nm peak and to ensure that the oxidized sample did not aggregate.

Quantification of the total charge passed during electrolysis indicated near-complete oxidation of the aerobic protein by ~60 minutes. Prior to electrolysis, the UV-visible spectrum showed the broad peak centered at 410 nm characteristic of a [4Fe4S] cluster, but after oxidation, the absorbance increased over a broad range from 700 – 300 nm, with a poorly-defined peak around 410 nm and a substantial shoulder between 400 and 300 nm (Figure 4.8c). Such absorbance features are a general characteristic of cluster oxidation, although UV-vis spectra

alone are insufficient to precisely identify the oxidized species generated (*41*). Importantly, the 280 nm peak remained sharp and distinct even after oxidation, and the spectrum was not elevated at 800 nm, demonstrating that the protein had not aggregated and confirming that all changes were due solely to cluster oxidation. This result stands in stark contrast to the soluble MUTYH aggregates observed in both WT and C306W in the absence of DTT, which were visibly cloudy with a U-shaped UV-visible spectrum characteristic of aggregation, with highly elevated absorbance at 800 nm, no distinct [4Fe4S] peak, and a very slight 280 nm peak visible only as a shoulder.

CV and SQWV of the aerobically oxidized C306W MUTYH variant revealed an irreversible peak comparable in size to the main reversible peaks; in contrast, the equivalent peak in the anaerobic sample was much smaller than the main peak and had not changed from initial levels (Figure 4.8d). Furthermore, the reversible signal of the anaerobic sample increased over time and exceeded even the strongest signals observed for aerobically incubated WT MUTYH, which was all the more intriguing given that the anaerobic sample had been incubating on the electrode for several hours. Supporting greater instability of the C306W [4Fe4S] cluster, aerobic oxidation of WT MUTYH gave low bulk electrolysis yields, and no readily apparent irreversible peak was present by CV. In addition, the UV-vis spectra of WT before and after oxidation were indistinguishable. Taken together, the apparent sensitivity of MUTYH C306W to oxidation and degradation in air along with the absence of any observable degradation in aerobically oxidized WT MUTYH all supported the assignment of the secondary peak to a $[3Fe4S]^{+}$ cluster. However, UV-visible spectroscopy and electrochemistry alone are insufficient to distinguish between possible oxidized Fe-S species, which requires EPR spectroscopy.

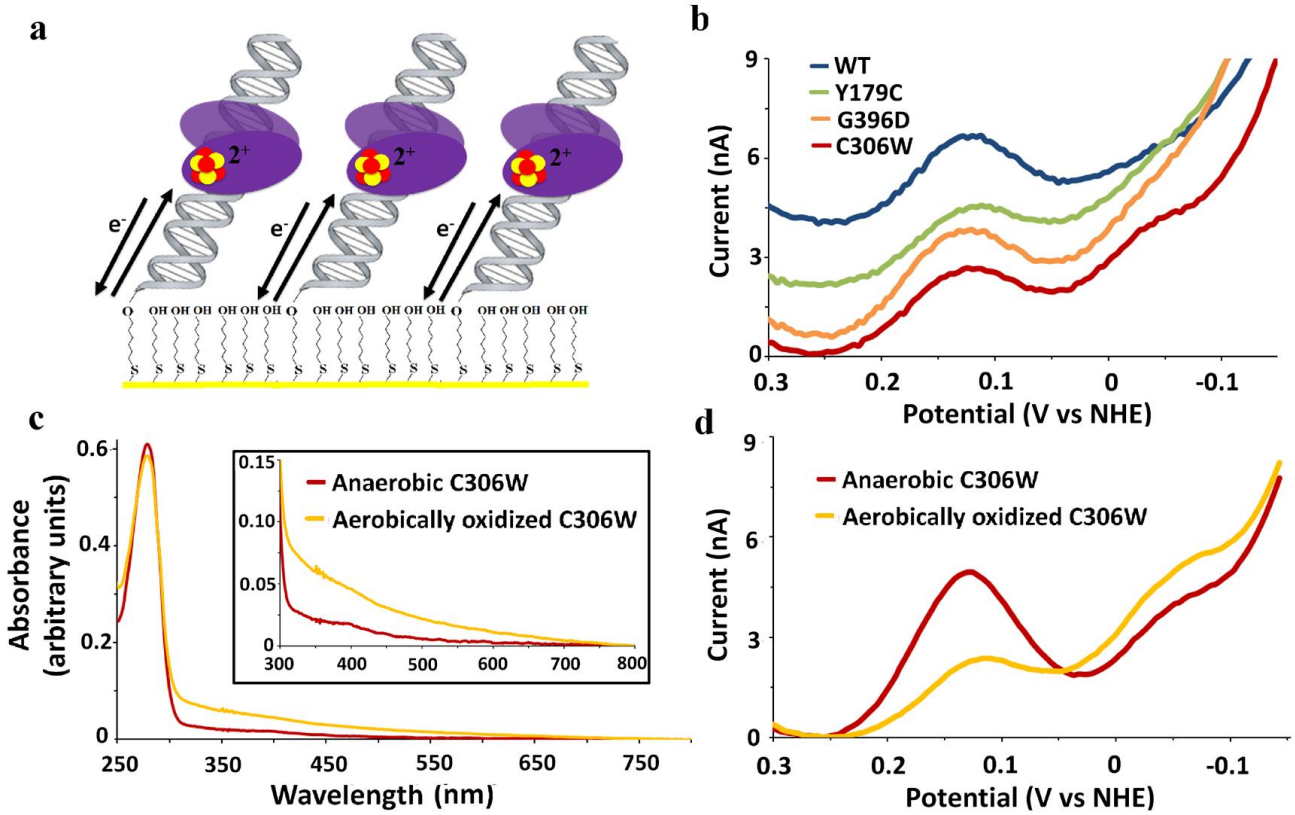


Figure 4.8 Initial electrochemical and spectroscopic characterization of MUTYH variants. (**a**) Electrochemistry is carried out on DNA-modified gold electrodes, which allow controlled reduction or oxidation of the [4Fe4S] cluster. (**b**) Incubation of 2-2.5 μM WT, Y179C, G396D, or C306W MUTYH on a DNA-modified electrode in storage buffer (20 mM Tris, 100 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, pH 7.4) results in a reversible signal with a midpoint potential of approximately 105 mV versus NHE. C306W uniquely exhibited an irreversible reductive peak around -50 mV versus NHE, which was presumed to be some form of oxidative degradation product. (**c**) After aerobic oxidation by bulk electrolysis, the UV-vis spectrum of C306W showed no evidence of aggregation but did display a broad increase in absorption from 700 to 300 nm, suggestive of cluster oxidation. (**d**) Consistent with cluster degradation in the presence of oxygen, aerobic bulk electrolysis of C306W enhanced the size of the peak relative to the reversible signal, while anaerobic incubation provided a protective effect. All SQWV measurements were obtained at a frequency of 15 Hz and 0.025 V amplitude, and the signals shown are an average from at least four separate electrodes on a multiplexed chip. The SQWV background current levels in (**b**) were all comparable and have been adjusted for ease of visualization. Bulk electrolysis was performed for 1 hour at 0.412 V vs NHE on a DNA-modified gold rod electrode in a glass cell.

EPR Spectroscopy of MUTYH

Having confirmed that the secondary C306W reductive peak was an oxidation product formed under aerobic conditions, we turned to EPR spectroscopy as a final step toward its identification. EPR provides a means of distinguishing among different paramagnetic species and is commonly employed to study iron-sulfur proteins (*40* - *42*). While EPR analysis can be very informative, there were two general concerns with respect to MUTYH. First, EPR experiments are generally performed with significantly higher levels of concentrated protein than those used in our electrochemical experiments with MUTYH: signals have been reported for 10 μM *E. coli* EndoIII (*20*) and ~9 μM DNA polymerase δ (*43*), but even these were still 3-4 times more concentrated than the MUTYH samples. Second, the low temperatures necessary to resolve signals from [4Fe4S] clusters (10-35 K) require the samples to be frozen prior to analysis, which can impact protein stability if the buffer pH changes with temperature, as is the case for the Tris buffers employed in our MUTYH studies (*44*).

Therefore, prior to attempting EPR, all MUTYH variants were concentrated and exchanged into a HEPES buffer (20 mM HEPES, pH 7.4, 100 mM KCl, 1 mM DTT, 10% glycerol v/v, pH 7.4); potassium was used in place of sodium as a further precaution to stabilize the pH at low temperature (*45*). UV-visible spectroscopy and electrochemistry (Figure 4.9a, 10) were used to confirm protein stability in HEPES and to verify that redox properties remained comparable. As observed in the UV-vis spectra (Figure 4.9a), all of the MUTYH variants maintained their monomeric form in this buffer, although they could not be concentrated beyond 15 μM without forming soluble aggregates. When such aggregates did form, they were readily resolved by simple dilution, and the UV-visible spectra of WT MUTYH and the mutants Y179C and G396D all retained a sharp 410 nm peak with ~15% cluster loading even after an additional

freeze-thaw cycle one week after buffer exchange (Figure 4.9a). In contrast, the spectrum of the MUTYH C306W protein closely resembled the aerobically oxidized sample described previously (Figure 4.9a). Oxidation may have occurred over the extended aerobic buffer exchange process or during the freeze-thaw cycle on the day of EPR experiments. Because the extinction coefficient at 410 nm was unknown for the MUTYH C306W product, cluster loading in this mutant was estimated by comparing the magnitude of absorbance with the earlier aerobically oxidized sample, yielding a concentration comparable to the other variants.

All EPR spectra of MUTYH proteins were obtained with 15 μM WT and C306W and 5 μM G396D and Y179C (the latter mutants were not available in such large amounts). From the broadened UV-vis absorption, we predicted that the corresponding EPR spectrum would show evidence of either the $[3Fe4S]^+$ cluster or a more advanced degradation product, while WT, Y179C, and G396D were expected to be diamagnetic and thus EPR silent. Unexpectedly, all of the samples showed a small, broad signal with a shoulder at g = 2.04 (Figure 4.9b) most likely attributable to oxidation during the aerobic freezing process, as reported previously for *E. coli* EndoIII and MutY[20]. Nonetheless, MUTYH C306W protein displayed a much sharper signal with a clear peak centered at g = 2.018 (Figure 4.9b), which is characteristic of $[3Fe4S]^+$ clusters (*40*, *42*). Importantly, the C306W EPR signal, but not the broad signals of WT and the other MUTYH variants, closely resembled spectra from chemically oxidized *E. coli* EndoIII and MutY that were also assigned to the $[3Fe4S]^+$ cluster (*20*). While some of the C306W [4Fe4S] degradation may have occurred during sample freezing, the significantly larger and sharper EPR signal relative to the other variants and the UV-visible spectrum indicating previous oxidation suggest that the cluster was largely in this state prior to freezing.

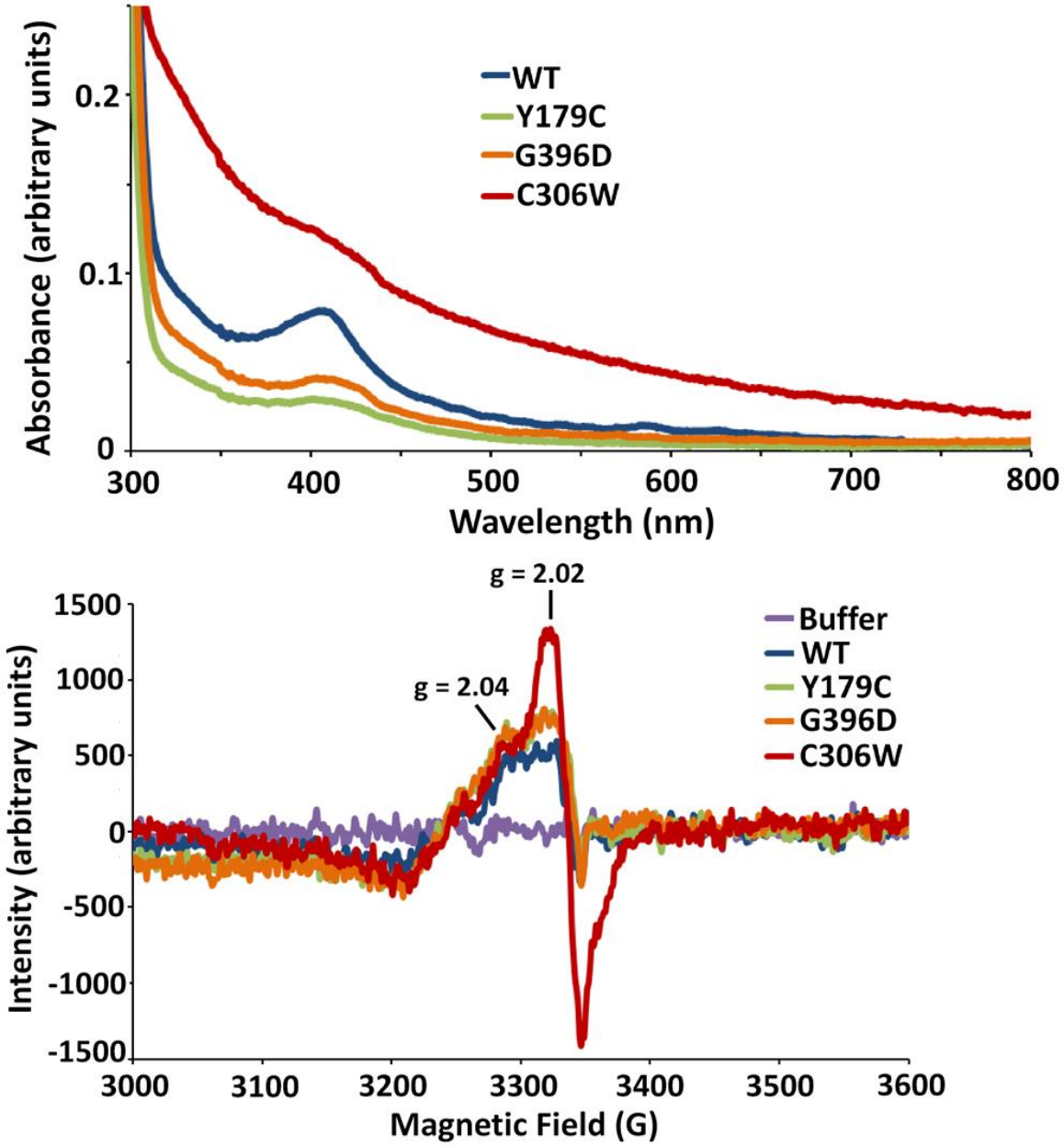


Figure 4.9 Characterization of MUTYH in HEPES and analysis of the C306W decay product. (**top**) UV-visible spectra of concentrated WT, Y179C, and G396D exchanged into HEPES buffer (30 mM HEPES, 100 mM KCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, pH 7.4) all displayed the characteristic [4Fe4S] cluster absorption band centered at 410 nm. In contrast, the C306W spectrum showed broadly increased absorbance associated with oxidation. (**bottom**) EPR spectra of 15 μM (WT and C306W) or 5 μM (Y179C and G396D) MUTYH variants. The C306W EPR spectrum shows a sharp peak at g = 2.018 with a shoulder at g = 2.04 EPR, supporting identification of the degradation product as a $[3Fe4S]^{+}$ cluster. Continuous wave X-band EPR spectra were measured at 10 K with 12.85 mW microwave power, 2 G modulation amplitude, and $5.02*10^{3}$ receiver gain.

Electrochemical Characterization of MUTYH in EPR Buffer

In parallel with EPR experiments, we performed electrochemistry with concentrated (5 μM) MUTYH in HEPES buffer to verify that the redox properties were comparable to those observed in Tris buffer (Figure 4.10). In HEPES buffer, the midpoint potential of WT MUTYH decreased from 105 ± 1 mV vs NHE in Tris to 93 ± 1 mV vs NHE (Table 4.4), but the overall signal shape and properties remained unaltered. Unsurprisingly, the addition of more concentrated protein to the electrode (~5 μM for all electrochemical experiments in HEPES) gave substantially larger signals, with WT MUTYH yielding a reductive peak area of 7.6 ± 3 nC and an oxidative peak area of -5.6 ± 3 nC on low density monolayers. Remarkably, although the protein was only about twice as concentrated compared with previous experiments, the signal was over 10 times larger, suggesting more efficient redox coupling in HEPES.

Like WT, all of the mutants showed ~10-fold larger signals when concentrated in HEPES, and their midpoint potentials decreased by a similar margin, placing all variants in the same potential window (Table 4.4). The signal from the MUTYH Y179C protein was still noticeably smaller than WT (reductive and oxidative peaks at 48 ± 18% and 33 ± 17% of WT), most likely as a result of lower DNA binding affinity, while G396D was not significantly different from WT (reductive and oxidative peaks at 77 ± 32% and 76 ± 41% of WT). To ensure accurate comparison with EPR spectra, electrochemistry for the less stable C306W mutant was performed on the same day as EPR, and using the same sample stock. Consistent with the oxidation indicated in the UV-visible spectrum and the presence of a characteristic $[3Fe4S]^{+}$ cluster EPR signal, the electrochemical signal from this sample was significantly smaller than any other mutant (reductive and oxidative peaks 35 ± 15% and 26 ± 14% of WT) and possessed a prominent secondary reductive peak centered at -88 ± 6 mV vs NHE (Figure 4.10). The larger

signals made this peak much more readily quantifiable compared with Tris buffer, and the total charge of this irreversible peak was found to represent ~20% of either peak of the reversible $[4Fe4S]^{3+/2+}$ couple. The area in this case was lower than estimated for the C306W sample aerobically oxidized in Tris buffer, where the secondary peak was roughly equal in magnitude to the main peak, which suggested that the concentrated sample in HEPES was not 100% oxidized. Overall, UV-visible and EPR spectra obtained from the same MUYH C306W protein strongly supported the identification of the irreversible peak as a $[3Fe4S]^{+/0}$ redox couple. When compared to WT MUTYH, or even the Y179C and G396D mutants, the susceptibility of the C306W [4Fe4S] cluster to degradation upon oxidation during redox signaling was apparent, and the irreversibility of the degradation product signal indicates that the resultant $[3Fe4S]^{+/0}$ species does not bind DNA effectively.

Based on the measured midpoint potentials in Tris buffer (Table 4.4), electrochemical signals from dilute MUTYH samples were most likely DNA-bound (*35*); however, the small signal sizes made it impossible to determine if the signal was DNA-mediated. The substantially larger signals recorded in HEPES buffer with concentrated samples allowed this issue to be addressed, and chips containing half well-matched DNA and half DNA containing an abasic site were prepared for this purpose. Earlier work with DNA-mediated EndoIII signals showed a dependence on monolayer morphology (*35*); thus, to see if this was also the case for MUTYH, the abasic site discrimination studies were carried out on both high and low density monolayers.

On abasic DNA, a maximum charge attenuation of 38% for the reductive peak and 46% for the oxidative peak was obtained on low density monolayers (Figure 4.11); in contrast, no appreciable discrimination was observed on high density monolayers. The observed abasic site discrimination confirmed that MUTYH can take part in DNA-mediated signaling, and the

differences in high and low density DNA monolayers emphasize the importance of surface accessibility to large proteins. The sterically hindered high density films clearly do not provide sufficient access to DNA in an appropriate conformation to observe such a signal, and the peaks seen in this case are likely attributable to DNA-bound proteins signaling directly through the monolayer surface (*35*).

Interestingly, the pattern of abasic site discrimination observed here is opposite to previously published data for *E. coli* EndoIII (*35*); this difference can be rationalized by considering two important factors. First, the DNA used in the EndoIII experiments was only 15 bp, while EndoIII has a binding footprint of 12 bp (*46*). Thus, the protein would have taken up much of the available space and the cluster could readily bypass the mismatch on fully accessible low density DNA. In contrast, the DNA substrate chosen for MUTYH was 36 bp while the binding footprint is likely close to the 16 bp reported for mouse MUTYH (*39*), making bypass of the abasic site on low density monolayers less likely. Second, a considerable size discrepancy exists between the two proteins: unmodified MUTYH is 61 kDa, while EndoIII is only 24 kDa. Thus, steric hindrance in high density films would be expected to have a greater impact on MUTYH, a notion supported by the observation that the largest signals obtained on high density monolayers containing WM DNA were ~30% smaller than the equivalent signals on low density films.

Table 4.4 CV Midpoint Potentials of MUTYH Variants in Tris and HEPES Buffers.

MUTYH variant	E_{mdpt} in Tris (mV)	E_{mdpt} in HEPES (mV)
WT	106 ± 1	93 ± 1
C306W	114 ± 3	97 ± 1
Y179C	105 ± 0.8	100 ± 2
G396D	115 ± 0.1	99 ± 4

Error is the standard deviation of the mean from at least 3 separate experiments.

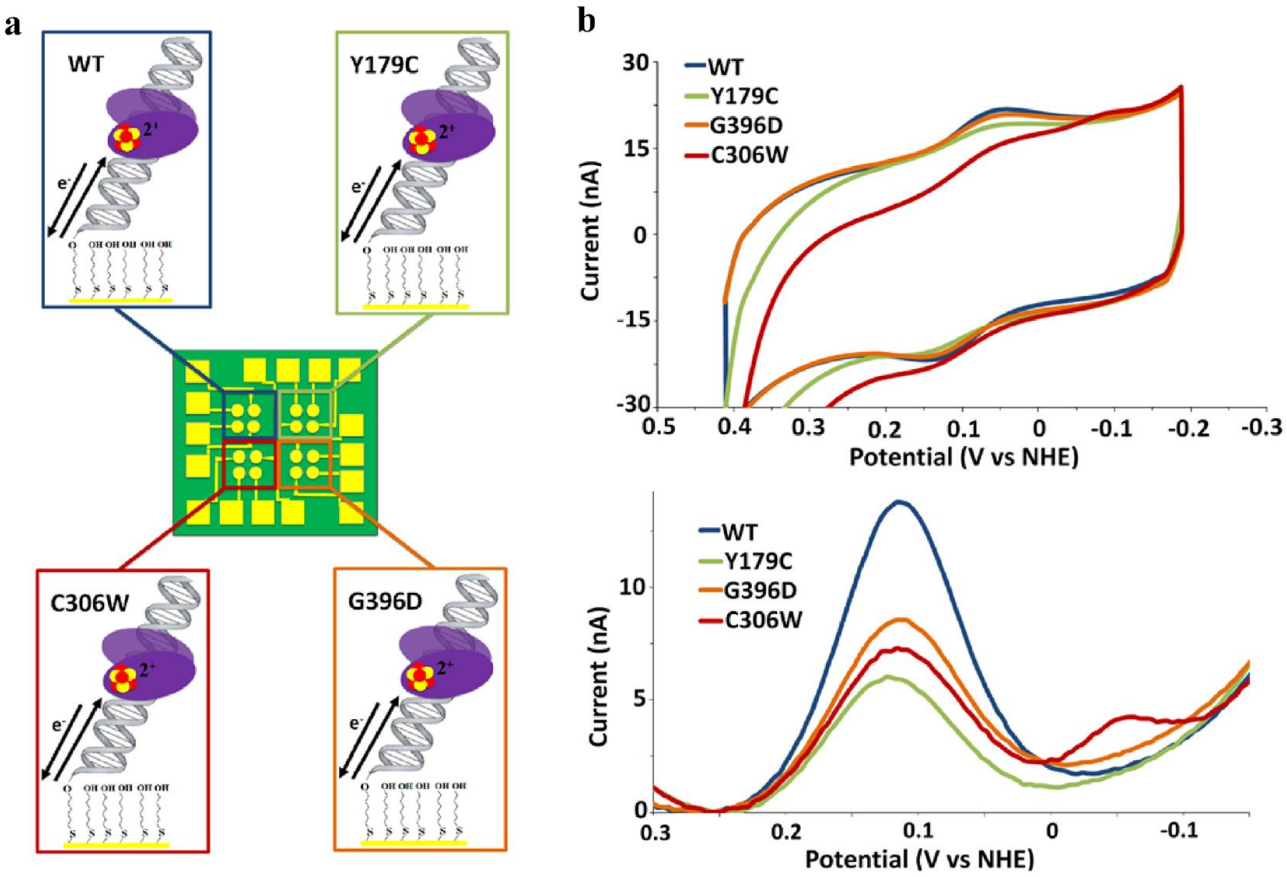


Figure 4.10 Electrochemical characterization of MUTYH variants in HEPES buffer. (**a**) Arrangement of MUTYH variants on a multiplexed chip. (**b**) CV (top) and SQWV (bottom) scans of 5 μM WT, Y179C, G396D, and C306W variants in HEPES storage buffer (30 mM HEPES, 100 mM KCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, pH 7.4). Under these conditions, substantial signals occur even by CV. The MUTYH C306W secondary peak is readily apparent in both CV and SQWV. CV scans were taken at a scan rate of 100 mV/s, while SQWV scans were taken at a frequency of 15 Hz with 0.025 V amplitude. All scans shown are an average obtained from at least seven separate electrodes.

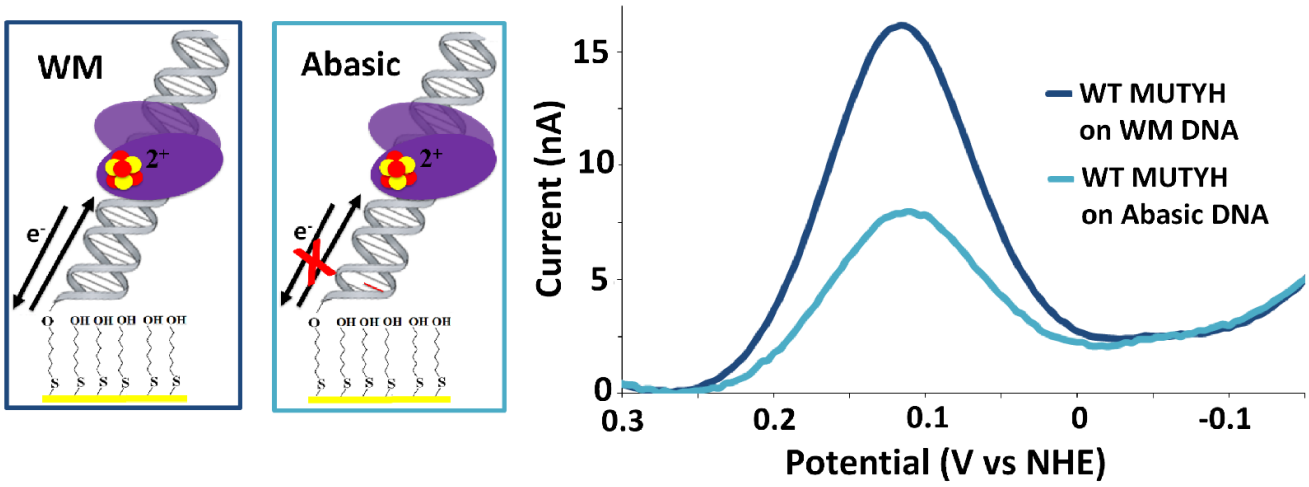


Figure 4.11 Abasic site discrimination by WT MUTYH. 5.0 μM WT MUTYH was incubated in HEPES buffer (30 mM HEPES, 100 mM KCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, pH 7.4) on a single multiplexed chip with half of the available quadrants containing DNA with an abasic site (light blue) near the electrode and the other half containing well matched (WM) DNA (dark blue). As is apparent from the SQWV voltammetry, the signal on WM DNA was comparable in size to previously observed signals (Figure 4.10), but the total peak area on abasic DNA was decreased by ~40%, indicating that the signal was DNA-mediated. The SQWV voltammograms were obtained at 15 Hz and 0.025 mV amplitude, and signal shown is the average of at least seven separate electrodes on a single multiplexed chip.

Discussion

In the present study we describe a novel MUTYH variant, C306W, and its association with the development of colonic polyposis and a family history of colon cancer. We determined that the C306W variant lacks DNA abscission activity and has decreased ability to bind target DNA, establishing the pathogenicity of this variant. In C306W MUTYH there occurs loss of a cysteine residue that ligates the MUTYH Fe-S cluster. This finding raised the possibility that the loss of the cysteine might disrupt the integrity of the Fe-S cluster and be causative of the pathogenicity associated MUTYH C306W. The observation by ICP-HRMS that the Fe-S cluster of C306W exhibits significantly lower iron content provided further bolsters this hypothesis and prompted an electrochemical mechanistic investigation of the MUTYH Fe-S cluster and the effect of cysteine loss in the C306W variant.

Toward this end, in our investigation, we present the first direct evidence of redox signaling in eukaryotic MUTYH, which until now had only been assumed based on studies using *E. coli* MutY. DNA-modified electrochemical analysis revealed the redox potentials of all MUTYH variants studied to be in general agreement with earlier work (Table 4.4). Notably, the potentials as measured in HEPES buffer are almost identical to those obtained for *E. coli* MutY in phosphate buffer (*20*). The similarity of the WT MUTYH electrochemical signals to those of the *E. coli* protein strongly supports the notion that the primary function of the conserved Fe-S cluster is redox activity in all organisms. Furthermore, the DNA-mediated nature of this signal in MUTYH suggests that a process akin to the DNA-mediated redox-based damage search observed in bacteria may also be present and operating in humans.

Unlike WT MUTYH, the C306W mutant showed an unexpected, and irreversible, reduction between -50 and -100 mV versus NHE in combination with loss of the reversible

signal; EPR spectroscopy confirmed this additional signal to be the $[3Fe4S]^{+/0}$ couple. The observed degradation and poor DNA binding in MUTYH C306W are consistent with the higher DNA binding affinity associated with increasing charge in [4Fe4S] clusters, in which Coulombic effects cause the $[4Fe4S]^{3+}$ cluster to bind the DNA polyanion significantly more tightly than the $[4Fe4S]^{2+}$ form (*21*). The relationship of cluster charge to binding affinity predicts that the $[3Fe4S]^{+}$ and $[3Fe4S]^{0}$ degradation products, with one and zero net charges, would bind much more weakly to DNA than the $[4Fe4S]^{2+}$ form, and the irreversible reduction observed in degraded MUTYH C306W was experimentally verified. Overall, our results suggest that, for MUTYH C306W, ordinary redox activity on DNA would lead to oxidation to the $[4Fe4S]^{3+}$ state, as is typical in these proteins, but the lower stability of the cluster would promote the loss of an iron atom and irreversible dissociation following a second redox signaling cycle. Ultimately, this process could result in the low iron content measured by ICP-HRMS, an effect that might well be exacerbated if the dissociated $[3Fe4S]^{0}$ form degraded further when removed from the protective environment of DNA. Cluster degradation in MUTYH C306W is also consistent with the low levels of glycosylase activity and poor DNA binding affinity as measured by BLI (Table 4.2), which are attributes of bacterial MutY following cluster removal (*19*). This inherent instability of the C306W Fe-S cluster and consequent loss of function, we propose to be causative of pathogenicity this MUTYH variant.

With regard to other potential causes of pathogenicity, we recognize that MUTYH is also regulated by post-translational modifications, including phosphorylation and ubiquitination, that could be altered by this mutation (*47, 48*). However, these sites are in different regions of the protein relative the [4Fe4S] domain, and are thus unlikely to be affected by this particular

mutation. Thus, redox-stimulated cluster degradation is most likely the primary cause of pathogenicity in MUTYH C306W.

The irreversible [3Fe4S] cluster signal seen in MUTYH C306W has not been previously observed in electrochemical studies of DNA-processing [4Fe4S] proteins, but the signal was within the same redox potential range reported for the $[3Fe4S]^{+/0}$ couple of bacterial nickel-iron hydrogenase and fumarate reductase enzymes (*49-51*). In *E. coli* MutY, [4Fe4S] cluster ligand substitution of the corresponding cysteine residue has been shown to result in defective DNA binding, similar to the situation with MUTYH C306W (*52*). It should be noted, however, that none of the substitutions involving *E. coli* MutY were to a residue as bulky or destabilizing as tryptophan (*52*).

Given the results obtained for MUTYH C306W, it appears probable that mutations in other residues that alter the region around the [4Fe4S] cluster will be similarly deficient in their ability to mediate repair of oxidatively damaged DNA *in vivo* (*53-57*). Indeed, four arginines participate in H-bonding to the four cysteines that coordinate with the [4Fe4S] cluster (*56*) (Table 4.5) and mutations to each one are associated with colorectal cancer (*1*). It is probable that these mutations also result in instability, degradation and dysfunction of the Fe-S cluster secondary to the same mechanisms detailed above. The effects of these lesions as well as the C306W variant underscore the importance of the [4Fe-4S] co-factor in establishing competent MUTYH-mediated DNA repair.

The current study advances our basic electrochemical understanding of the redox chemistry, function, and integrity of the [4Fe4S] cluster. Concomitantly, we are acquiring a better appreciation for the pathologic sequelae resulting from disruption of the Fe-S cluster. Specifically, within the context of the present investigation we have documented and provided

explanation for a novel mechanism of colonic polyposis and cancer predisposition caused by electrochemical compromise of the MUTYH [4Fe4S] cluster. Future studies, we anticipate, will provide further clarification of the central role of the cluster in MUTYH-mediated DNA repair and its underlying electrochemistry.

Table 4.5 The four coordinating cysteines and surrounding arginines predicted or reported[14] to be associated with MAP.

Residue	Reported Variant
C290	C290W[55]
C297	not reported
C300	not reported
C306	C306W (this study)
R241	R241W[53,55], R241G[14]
R245	R245L[56], R245C[14], R245H[14,54]
R247	R247G[14]
R309	R309C[9,57]

References

1) Nghiem, Y.; Cabrera, M.; Cupples, C. G.; Miller, J. H. The mutY gene: a mutator locus in Escherichia coli that generates G.C----T.A transversions. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 2709 – 2713.

2) Markkanen, E.; Dorn, J.; Hubscher, U. MUTYH DNA glycosylase: the rationale for removing undamaged bases from the DNA. *Front. Genet.* **2013**, *4*, 18.

3) Wallace, S. S.; Murphy, D. L.; Sweasy, J. B. Base excision repair and cancer. *Cancer Lett.* **2012**, *327*, 73 – 89.

4) Al-Tassan, N. *et al.* Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. *Nature Genetics* **2002**, *30*, 227 – 232.

5) Nielsen, M.; Morreau, H.; Vasen, H. F.; Hes, F. J. MUTYH-associated polyposis (MAP). *Crit. Rev. Oncol. Hematol.* **2011**, *79*, 1 – 16.

6) Jasperson, K. W.; Tuohy, T. M.; Neklason, D. W.; Burt, R. W. Hereditary and familial colon cancer. *Gastroenterology* **2010**, *138*, 2044 – 2058.

7) Sampson, J. R. *et al.* Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. *Lancet* **2003**, *362*, 39 – 41.

8) Wang, L. *et al.* MYH mutations in patients with attenuated and classic polyposis and with young-onset colorectal cancer without polyps. *Gastroenterology* **2004**, *127*, 9 – 16.

9) Sieber, O. M. *et al.* Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. *N. Engl. J. Med.* **2003**, *348*, 791 – 799.

10) Isidro, G. *et al.* Germline MUTYH (MYH) mutations in Portuguese individuals with multiple colorectal adenomas. *Hum. Mutat.* **2004**, *24*, 353 – 354.

11) Cleary, S. P. *et al.* Germline MutY human homologue mutations and colorectal cancer: a multisite case-control study. *Gastroenterology* **2009**, *136*, 1251 – 1260.

12) Jones, N. *et al.* Increased colorectal cancer incidence in obligate carriers of heterozygous mutations in MUTYH. *Gastroenterology* **2009**, *137*, 489 – 494.

13) Farrington, S. M. *et al.* Germline susceptibility to colorectal cancer due to base-excision repair gene defects. *Am. J. Hum. Genet.* **2005**, *77*, 112 – 119

14) Out, A. A. *et al.* Leiden Open Variation Database of the MUTYH gene. *Hum. Mutat.* **2010**, *31*, 1205 – 1215.

15) Maio, N.; Rouault, T. A. Iron-sulfur cluster biogenesis in mammalian cells: New insights into the molecular mechanisms of cluster delivery. *Biochim. Biophys. Acta*, **2014**, *1853*, 1493 – 1512.

16) Alseth, I.; Eide, L.; Pirovano, M.; Rognes, T.; Seeberg, E.; Bjørås, M. The Saccharomyces cervisiae homologues of endonuclease III from Escherichia coli, Ntg1 and Ntg2, are both required for efficient repair of spontaneous and induced oxidative damage in yeast. *Mol. Cell. Biol.* **1999**, *19*, 3779 – 3787.

17) Trasvina-Areñas, C. H.; Lopez-Castillo, L. M.; Sanchez-Sandoval, E.; Brieba, L. G. Dispensability of the [4Fe-4S] cluster in novel homologues of adenine glycosylase MutY. *FEBS J.* **2016**, *283*, 521 – 540.

18) Cunningham, R. P. *et al.* Endonuclease III is an iron-sulfur protein. *Biochemistry* **1989**, *28*, 4450 – 4455.

19) Porello, S. L.; Cannon, M. J.; David, S. S. A substrate recognition role for the $[4Fe\text{-}4S]^{2+}$ cluster of the DNA repair glycosylase MutY. *Biochemistry* **1998**, *37*, 6465 – 6475.

20) Boal, A. K. *et al.* DNA-bound redox activity of DNA repair glycosylases containing [4Fe-4S] clusters. *Biochemistry* **2005**, *44*, 8397 – 8407.

21) Gorodetsky, A. A.; Boal, A. K.; Barton, J. K. Direct electrochemistry of endonuclease III in the presence and absence of DNA. *J. Am. Chem. Soc.* **2006**, *128*, 12082 – 12083.

22) Arnold, A. R.; Grodick, M. A.; Barton, J. K. DNA Charge Transport: from Chemical Principles to the Cell. *Cell Chem. Biol.* **2016**, *23*, 183 – 197.

23) O'Brien, E.; Silva, R. M.; Barton, J. K. Redox Signaling through DNA. *Isr. J. Chem.* **2016**, *56*, 705 – 723.

24) 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*, 3546 – 3551.

25) Boal, A. K. *et al.* Redox signaling between DNA repair proteins for efficient lesion detection. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15237 – 15242.

26) Miyoshi, Y. *et al.* Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. *Hum. Mol. Genet.* **1992**, *1*, 229 – 233.

27) Donnelly, M. I. *et al.* An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. *Protein Expr. Purif.* **2006**, *47*, 446 – 454.

28) Kapust, R. B.; Waugh, D. S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **1999**, *8*, 1668 – 1674.

29) D'Agostino, V. G. *et al.* Functional analysis of MUTYH mutated proteins associated with familial adenomatous polyposis. *DNA Repair (Amst)* **2010**, *9*, 700 – 707.

30) Whicher, J. R. *et al.* Cyanobacterial polyketide synthase docking domains: a tool for engineering natural product biosynthesis. *Chemistry & Biology* **2013**, *20*, 1340 – 1351.

31) Kundu, S.; Brinkmeyer, M. K.; Livingston, A. L.; David, S. S. 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*, 1400 – 1410.

32) Porello, S. L.; Leyes, A. E.; David, S. S. Single-turnover and pre-steady-state kinetics of the reaction of the adenine glycosylase MutY with mismatch-containing DNA substrates. *Biochemistry* **1998**, *37*, 14756 – 14764.

33) Rich, R. L.; Myszka, D. G. Higher-throughput, label-free, real-time molecular interaction analysis. *Anal. Biochem.* **2007**, *361*, 1 – 6.

34) Profrock, D.; Prange, A. Inductively coupled plasma-mass spectrometry (ICP-MS) for quantitative analysis in environmental and life sciences: a review of challenges, solutions, and trends. *Appl. Spectrosc.* **2012**, *66*, 843 – 868.

35) Pheeney, C. G.; Arnold, A. R.; Grodick, M. A.; Barton, J. K. Multiplexed electrochemistry of DNA-bound metalloproteins. *J. Am. Chem. Soc.* **2013**, *135*, 11869 – 11878.

36) Boon, E. M.; Salas, J. E.; Barton, J. K. An electrical probe of protein-DNA interactions on DNA-modified surfaces. *Nat. Biotechnol.* **2002**, *20*, 282 – 286.

37) Kelley, S. O.; Barton, J. K.; Jackson, N. M.; Hill, M. G. Electrochemistry of methylene blue bound to a DNA-modified electrode. *Bioconjug. Chem.* **1997**, *8*, 31 – 37.

38) Greeley, R. S.; Smith, W. T.; Stoughton, R. W.; Lietzke, M. H. Electromotive Force Studies in Aqueous Solutions at Elevated Temperatures .1. The Standard Potential of the Silver-Silver Chloride Electrode. *J. Phys. Chem-US* **1960**, *64*, 652 – 657.

39) Pope, M. A.; David, S. S. DNA damage recognition and repair by the murine MutY homologue. *DNA Repair (Amst)* **2005**, *4*, 91 – 102.

40) Johnson, M. K.; Duderstadt, R. E.; Duin, E. C. Biological and synthetic [Fe3S4] clusters. *Adv. Inorg. Chem.* **1999**, *47*, 1 – 82.

41) Sweeney, W. V.; Rabinowitz, J. C. Proteins containing 4Fe-4S clusters: an overview. *Annu Rev. Biochem.* **1980**, *49*, 139 – 161.

42) Duff, J. L. C.; Breton, J. L. J.; Butt, J. N.; Armstrong, F. A.; Thomson, A. J. Novel redox chemistry of [3Fe-4S] clusters: Electrochemical characterization of the all-Fe(II) form of the [3Fe-4S] cluster generated reversibly in various proteins and its spectroscopic investigation in Sulfolobus acidocaldarius ferredoxin. *J. Am. Chem. Soc.* **1996**, *118*, 8593 – 8603.

43) Netz, D. J. A. *et al.* Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat. Chem. Biol.* **2012**, *8*, 125 – 132.

44) Good, N. E. *et al.* Hydrogen Ion Buffers for Biological Research. *Biochemistry* **1966**, *5*, 467 – 477.

45) Ugwu, S. O.; Apte, S. P. The Effect of Buffers on Protein Conformational Stability. *Pharmaceutical Technology* **2004**, *28*, 86 – 108.

46) O'Handley, S.; Scholes, C. P.; Cunningham, R. P. Endonuclease III interactions with DNA substrates. 1. Binding and footprinting studies with oligonucleotides containing a reduced apyrimidinic site. *Biochemistry* **1995**, *34*, 2528 – 2536.

47) O'Handley, S.; Scholes, C. P.; Cunningham, R. P. Endonuclease III interactions with DNA substrates. 1. Binding and footprinting studies with oligonucleotides containing a reduced apyrimidinic site. *Biochemistry* **1995**, *34*, 2528 – 2536.

48) Dorn, J.; Ferrrari, E.; Imhof, R.; Ziegler, N.; Hübscher, U. Regulation of human MutYH DNA glycosylase by the E3 ubiquitin ligase. *J. Biol. Chem.*, **2014**, *289*, 7049 – 7058.

49) Asso, M.; Guigliarelli, B.; Yagi, T.; Bertrand, P. EPR and redox properties of Desulfovibrio vulgaris Miyazaki hydrogenase: comparison with the Ni-Fe enzyme from Desulfovibrio gigas. *Biochim. Biophys. Acta* **1992**, *1122*, 50 – 56.

50) Teixeira, M. *et al.* Redox intermediates of Desulfovibrio gigas [NiFe] hydrogenase generated under hydrogen. Mossbauer and EPR characterization of the metal centers. *J. Biol. Chem.* **1989**, *264*, 16435 – 16450.

51) Kowal, A. T. *et al.* Effect of cysteine to serine mutations on the properties of the [4Fe-4S] center in Escherichia coli fumarate reductase. *Biochemistry* **1995**, *34*, 12284 – 12293.

52) Golinelli, M. P.; Chmiel, N. H.; David, S. S. Site-directed mutagenesis of the cysteine ligands to the [4Fe-4S] cluster of Escherichia coli MutY. *Biochemistry* **1999**, *38*, 6997 – 7007.

53) Bai, H. *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 Res.* **2005**, *33*, 597 – 604.

54) Ali, M. *et al.* Characterization of mutant MUTYH proteins associated with familial colorectal cancer. *Gastroenterology* **2008**, *135*, 499 – 507.

55) Fleischmann, C. *et al.* Comprehensive analysis of the contribution of germline MYH variation to early-onset colorectal cancer. *Int. J. Cancer* **2004**, *109*, 554 – 558.

56) Bai, H. *et al.* Functional characterization of human MutY homolog (hMYH) missense mutation (R231L) that is linked with hMYH-associated polyposis. *Cancer Lett.* **2007**, *250*, 74 – 81.

57) Goto, M. *et al.* Adenine DNA glycosylase activity of 14 human MutY homolog (MUTYH) variant proteins found in patients with colorectal polyposis and cancer. *Hum. Mutat.* **2010**, *31*, E1861 – 1874.

Chapter 5

A Redox Role for the [4Fe4S] Cluster of Yeast DNA Polymerase δ

Adapted from: Bartels, P.L.; Stodola, J.L.; Burgers, P.M.J.; Barton, J.K. A Redox Role for the [4Fe4S] Cluster of Yeast DNA Polymerase δ. *J. Am. Chem. Soc.*, **2017**, *139*, 18339 – 18348.

P. Bartels synthesized DNA, carried out all electrochemical and spectroscopic characterization, and ran and analyzed activity assays. J. Stodola and P. Burgers provided purified proteins for experiments.

Introduction

During genomic replication, eukaryotic cells divide the task of DNA synthesis between three B-family DNA polymerases (Pols): Pol α, Pol δ, and Pol ε (*1*). In the most widely accepted model, the DNA primase-Pol α complex initiates 5′-3′ DNA synthesis by forming an RNA-DNA hybrid primer that is then extended by Pol ε on the continuously generated leading strand and by Pol δ on the discontinuously formed lagging strand (*2*). Additional roles for Pol δ during leading strand replication have been suggested (*3, 4*), and Pol δ is also involved in various DNA recombinatorial and repair processes (*5*). Structurally, each of the B-family polymerases forms a multi-subunit complex composed of a catalytic subunit and a B-subunit, with additional accessory subunits present in Pol δ and Pol ε (*6*). Recent work has shown that a $[4Fe4S]^{2+}$ cluster in the C-terminal domain (CTD) of the catalytic subunit is essential for the formation of multi-subunit complexes at least in the case of Pol δ (*6*).

While the C-terminal [4Fe4S] cluster is certainly important for complex formation (*6*), several lines of evidence suggest a more direct functional role. First, a 2.5 Å X-ray crystal structure of the yeast Pol α CTD in complex with its B-subunit contained zinc in place of a cluster, demonstrating that complex formation can be supported by simpler metals (*8*). Given the metabolic expense of [4Fe4S] cluster biosynthesis and loading into target proteins, the strict conservation of this cofactor in the B-family polymerases suggests that it serves an important function (*9*). Indeed, the importance of [4Fe4S] clusters in these enzymes is emphasized by the presence of an additional cluster in the unique Pol ε N-terminal domain (*10*).

The [4Fe4S] clusters perform a wide range of roles in biology including enzymatic catalysis and electron transfer (*11*). In the DNA polymerases, the cluster is not required for catalysis (*3, 12*). Many DNA-processing enzymes have now been shown to contain [4Fe4S]

clusters, and, in many cases, a DNA-bound redox activity of the cluster has been demonstrated (*13-16*). These diverse proteins include base excision repair glycosylases, repair helicases, and DNA primase. As in the Pol δ holoenzyme, the clusters are largely redox-inert in the absence of DNA (*17-20*). However, when bound to DNA, these protein cofactors undergo a significant negative shift in redox potential, activating the clusters toward oxidation (*21-23*). Electrochemical experiments with DNA-bound proteins show a reversible redox signal with potentials ranging from 65-95 mV versus the normal hydrogen electrode (NHE) (*13-16*). EPR studies support the assignment of the reversible signal to the $[4Fe4S]^{3+/2+}$ couple favored by high-potential iron proteins (HiPIP) that are electron carriers (*13, 24-25*). In addition to modulating redox potential, the π-stacked base pairs of DNA can act as a medium for long-range charge transport between redox-active proteins (*26*). DNA-mediated charge transport (DNA CT) is characterized by a shallow distance dependence and high sensitivity to base pair stacking, making it an excellent reporter of DNA integrity (*26*). Importantly, although DNA CT can be attenuated by proteins that bend the duplex or flip out bases, DNA CT can proceed unimpeded through nucleosome-wrapped DNA (*26*).

The redox activity of the [4Fe4S] cluster appears to be utilized in many of these proteins as a switch to regulate DNA binding and therefore activity. For the DNA repair enzyme, Endonuclease III (EndoIII), the negative shift in redox potential associated with DNA binding has been shown to lead to a 500-fold increase in DNA affinity for the oxidized $[4Fe4S]^{3+}$ cluster versus the reduced 2+ form (*27*). In the case of human DNA primase, the oxidation state of the [4Fe4S] cluster also controls template binding, and redox switching through electron transfer between clusters in primase and Pol α has been proposed to regulate RNA primer handoff (*16*).

Here we focus on Pol δ, a central B-family polymerase. We utilize a combination of electrochemical, spectroscopic, and biochemical techniques to investigate redox activity in this enzyme and to understand the consequences of redox switching for polymerase activity. These studies provide a new perspective on polymerase regulation under oxidative stress.

Materials and Methods

Protein Expression and Purification

Yeast Pol δ (WT and exo$^-$ D520V), RFC, RPA, PCNA, and *E. coli* EndoIII were expressed according to previously published protocols (*7*, *28*).

DNA Preparation

The DNA substrate for electrochemistry consisted of a 49:58-mer primer-template composed of three oligomers: a 20-mer with a 3′ thiol modification, a 38-mer, and a 49-mer complement; sequences are as follows (see also Figure 5.1c):

20-mer Thiol: 5′ - GCT GTC GTA CAG CTC AAT GC - 3′ - $(CH_2)_2O(CH_2)_3SH$

38-mer: 5′ - TAA CAG GTT GAT GCA TCG CGC TTC GGT GCT GCG TGT CT - 3′

49-mer: 5′ - GCA TT**G** AGC TGT ACG ACA GCA GAC ACG CAG CAC CGA AGC GCG ATG CAT C - 3′

The bold G of the 49-mer was changed to an A or an abasic (AP) site for CA mismatch and abasic site discrimination experiments.

Thiol-modified DNA sequences for electrochemistry were prepared by standard phosphoramidite chemistry on a DNA synthesizer (Applied Biosystems) using A, G, C, T phosphoramidites and the 3′-Thiol-Modifier 6-S-S CPG as purchased from Glen research. DNA substrates were cleaved and deprotected by 8-hour incubation in NH_4OH (Sigma-Aldrich) at 65 °C. Deprotected DNA was separated from truncation products by reverse-phase HPLC (Agilent PLRPS column, gradient of 5 – 75% ACN/95 - 25% 50 mM NH_4Ac over 30 minutes at a 2 mL/min flow rate). Thiol-modified DNA was reduced by dissolving in 50 μL Tris, pH 8.0 (Qiagen elution buffer), adding excess DTT (Sigma-Aldrich), and shaking for 45 minutes. DTT was removed by filtration through a NAP-5 column (GE Healthcare) prior to a final round of

HPLC purification (gradient of 5 – 15% ACN/95 – 85% 50 mM NH_4Ac over 35 minutes at 2 mL/min). Lastly, single-stranded DNA was desalted by standard ethanol precipitation (100 µl water, 1 mL 100% EtOH, 130 mM NaCl) and the identity of the substrate was confirmed by MALDI-TOF. Unmodified oligomers were ordered from IDT and purified by the DMT-free HPLC method. Desalted DNA was dissolved in a phosphate storage buffer (5 mM sodium phosphate, 50 mM NaCl, pH 7.0) and concentrations were determined by UV-visible spectroscopy using ε_{260} values estimated by Integrated DNA Technologies (IDT). Equimolar concentrations of single stranded DNA were then degassed and annealed (rapid heating to 95^{o} C, 5-minute incubation, and 1.5 hour cooling to 20^{o} C).

DNA replication assays used single-stranded M13mp18 plasmid purchased from New England Biolabs (NEB). Primers were purchased from IDT and purified by HPLC as described above. Primed DNA was formed by heating a 1:1 plasmid/primer mix in activity buffer (50 mM Tris-HCl, pH 7.8, 50 mM NaCl) to 90° C for 5′ and cooling to RT over several hours. The M13mp18 DNA primer had the following sequence (complementary to positions 6265-6235):

5′ - GAC TCT AGA GGA TCC CCG GGT ACC GAG CTC G - 3′

Primers were radiolabeled by incubating 10 pmol of 31-mer M13mp18 primer with T4 polynucleotide kinase (PNK) and [γ-^{32}P] ATP (Perkin Elmer) in T4 buffer (NEB) for 15 minutes at 37° C. Reactions were stopped by addition of EDTA to 10 mM and heating at 75^{o} C for 10 minutes. 2 log DNA ladder (NEB) was dephosphorylated by calf intestinal alkaline phosphatase (CIAP; 60 minutes, 37° C) prior to labeling in the same manner. As an additional size standard, duplexed M13mp18 DNA was linearized by digestion with HincII (60 minutes, 37° C) and dephosphorylated by CIAP prior to radiolabeling. Proteins and unincorporated ATP were removed using spin columns (BioRad Microspin6) equilibrated in Pol δ activity buffer (50 mM

Tris-HCl, pH 7.8, 50 mM NaCl). T4 PNK, CIAP, HincII, and dsM13mp18 DNA were purchased from NEB.

Preparation of DNA-Modified Gold Electrodes

Multiplexed chips containing 16 Au electrodes (0.015 cm^2 area) were prepared as described previously (*29*). Self-assembled monolayers (SAMs) were formed by incubating 25 μL of 25 μM duplexed DNA on the electrode overnight, after which electrodes were rinsed 3-5 times in phosphate buffer (5 mM sodium phosphate, pH 7.5, 50 mM NaCl) and backfilled for 45 minutes with 1 mM 6-mercapto-1-hexanol (Sigma-Aldrich) in the same buffer containing 5% (v/v) glycerol. Electrodes were then extensively rinsed in phosphate buffer followed by protein storage buffer (30 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.1 mM EDTA, 10% glycerol, and 0.01% decaethylene glycol monododecyl ether). Lastly, the absence of electroactive impurities was confirmed by scanning the surface with cyclic voltammetry (CV). Bulk electrolysis experiments were undertaken by droplet electrochemistry (30-40 μL solution) on Au rod electrodes of 0.0314 cm^2 electrode area (Pine Research Instrumentation). Electrodes were cleaned as previously described (*30*), and monolayers formed using the same procedure as the multiplexed chip.

Pol δ Electrochemistry

All electrochemical experiments were performed using a potentiostat equipped with a multiplexer, both from CH Instruments. Experiments used a standard 3-electrode cell composed of an Au working electrode, a Ag/AgCl reference electrode in 3 M NaCl (BASInc), and a 1 mm diameter Pt wire counter electrode (Lesker). Potentials were converted from Ag/AgCl to NHE by adding 212 mV to the potential as measured by Ag/AgCl; this correction accounted for both ambient temperature and the use of 3 M NaCl for reference storage (*31*). Experiments with Pol δ

were all run in polymerase storage buffer (30 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.1 mM EDTA, 10% glycerol v/v, and 0.01% w/v decaethylene glycol monododecyl ether).

Because PCNA can slide directly onto DNA with open ends, the clamp loader complex RFC was excluded from these experiments (*32*). WT Pol δ 3′-5′ exonuclease activity was prevented by excluding Mg^{2+} from the buffer. In initial experiments, 3-5 μM WT Pol δ or exonuclease-deficient Pol δ D520V (DV) (*33*) was incubated on the electrode for several hours in the presence of 5-10 μM PCNA. To spare enzyme, later experiments used Pol δ DV at 500 nM in combination with 5.0 μM PCNA, 80 μM dATP and 8 mM $MgAc_2$. Cyclic voltammetry (CV; 100 mV/s scan rate) and square wave voltammetry (SQWV; 15 Hz frequency, 25 mV amplitude) scans were taken once per hour for several hours. Between scans, electrodes were covered in Parafilm and stored in a humid environment to minimize evaporation. CV scan rate dependence was assessed after 3 hours using rates of 20, 50, 80, 100, 200, 500, 750, and 1000 mV/s. In experiments with abasic and CA mismatch DNA, signal attenuation was calculated as follows: [1 – ((peak area on abasic or CA mismatch DNA)/(peak area on unmodified DNA))]*100%

Pol δ concentrations are reported as the concentration of [4Fe4S] cluster, determined by UV-visible absorbance at 400 nm (ε_{400} = 13000 $M^{-1}cm^{-1}$) (*7*). The [4Fe4S] cluster loading was in the range of 70-85%, determined by dividing [4Fe4S] concentration by total protein concentration as measured by Bradford assay, Pierce BCA assay, and UV-visible absorbance at 280 nm (ε_{280} = 194100 $M^{-1}cm^{-1}$; estimated using the EXPasy ProtParam tool). Bradford and BCA assay standard curves were generated using a BSA standard, and both kits were purchased from Thermo Scientific. UV-visible spectra were taken on a Cary Varian instrument using 100 μL quartz cuvettes purchased from STARNA Cells.

When possible, diffusion coefficients were obtained from the scan rate dependence of the CV current using the Randles-Sevcik equation (*34*):

$$I_p = [0.4463(F^3/RT)^{1/2}](n^{3/2})(A)(D^{1/2})(C^\circ)\nu^{1/2} \qquad (2)$$

I_p is the peak current in amperes, F is Faraday's constant (96485 $C \cdot mol^{-1}$), R is the universal gas constant (8.314 $J \cdot (mol \cdot K)^{-1}$), T is temperature in K, n is the number of electrons transferred per CV peak, A is electrode area in cm^2, D is the diffusion coefficient in $cm^2 \cdot s^{-1}$, C° is bulk protein concentration in $mol \cdot cm^{-3}$, and ν is the scan rate in $V \cdot s^{-1}$. Experimental values of D were compared to those estimated by the Stokes-Einstein equation,

$$D = k_B T/6\pi\eta R \qquad (3)$$

where k_B is Boltzmann's constant (1.38 x 10^{-23} $J \cdot K^{-1}$), T is the incubation temperature (293 K), η is the solution viscosity (estimated to be 1.38 x 10^{-3} Pa·s for an aqueous solution with 10% glycerol), and R is the hydrodynamic radius. R was estimated to be ~26 Å from dimensions obtained from X-ray crystal structures of DNA-bound Pol3, Pol1-Pol12, and PCNA (PDB ID 3IAY, 3FLO, and 4YHR, respectively).

Electrochemical Oxidation and Spectroscopic Analysis of Pol δ

To prevent cluster degradation in the presence of O_2, bulk electrolysis was performed in an anaerobic glove bag (COY) under a 95% N_2, 5% H_2 atmosphere with an O_2-scavenging catalyst present. Buffers were degassed by bubbling in argon for several hours and stored open in the glove bag overnight prior to experiments. For spectroscopic characterization, a 150 μL sample of 1-2 μM Pol δ was added to two identical DNA-modified electrodes. On one electrode, a potential of 0.412 V vs NHE was applied for ~15 minutes, while no potential was applied to the other. Oxidation yields were estimated by taking the difference between the total charge obtained in the presence of Pol δ and that generated by electrolysis with buffer alone. Following

electrolysis, UV-visible and electron paramagnetic resonance (EPR) spectroscopy were used to confirm the integrity of the cluster after electrolysis. Samples were sealed in cuvettes for UV-visible spectroscopy, and subsequently returned to the glove bag and added to EPR tubes. Tubes were sealed by Parafilm and frozen in liquid nitrogen outside the bag. Continuous wave X-band EPR was performed at 10 K, and each experiment consisted of 9 sweeps taken at 12.88 mW microwave power, 2 G modulation amplitude, and 5.02 $x10^3$ receiver gain.

Pol δ Activity Assays

Immediately prior to assays, Pol δ DV was oxidized on Au rod electrodes exactly as described for spectroscopic characterization, but the sample was diluted to 190 nM in degassed storage buffer in a total volume of 30-40 μL. Reduction of oxidized sample was carried out at a potential of -0.188 V vs NHE for a similar length of time. In parallel with electrolysis, 140 μL reaction mixes (0.1mg/mL BSA, 80 μM each dNTP, 500 μM ATP, 2.0 nM M13mp18 with a ^{32}P-labeled primer, 8.0 mM $MgAc_2$, 500 nM RPA, 5.0 nM RFC and 5.0 nM PCNA, 50 mM NaCl, 50 mM Tris-HCl, pH 7.8) were prepared inside the glove bag. The PCNA sliding clamp was loaded onto the primer end by incubating the reaction mix with the RFC clamp loader and ATP for 1 minute at 30 °C. After clamp loading, reactions were initiated by the addition of 2 nM (final concentration) oxidized, untreated, or re-reduced Pol δ DV. Reactions were run at 30 °C, and 20 μL aliquots were removed and quenched at specific time points by adding 10 μL stop mix (10 mM EDTA and 0.1% v/v SDS final concentration). The polymerase was heat-inactivated for 10-20 minutes (55 °C), and samples were counted on a liquid scintillation counter to determine exposure time (1 hour per 300,000 counts). Samples were dried on a speed vacuum and dissolved in alkaline gel buffer (500 mM NaOH, 10 mM EDTA) with 1x alkaline loading dye (6x stock: 300 mM NaOH, 6.0 mM EDTA, 18% Ficoll w/v, 0.25% xylene cyanol w/v, and

0.15% bromocresol green w/v), and equivalent amounts of radioactivity were then loaded onto a 1% alkaline agarose gel and run at 30 V for 14-15 hours. Gels were neutralized in 7% TCA (w/v) in water for 30 minutes at RT and dried under mild pressure for several hours, exposed on a phosphor screen and visualized on a Typhoon phosphorimager (GE Healthcare). Products were analyzed using ImageQuant software (GE Healthcare). The relative amounts of DNA synthesis were determined by dividing the volume of the largest band in an oxidized sample by the equivalent band in the appropriate untreated sample.

To limit DNA synthesis to that of a single processive cycle by the PCNA-Pol δ complex, 0.01 mg/mL heparin was included in reactions (*35*) that were then analyzed on 5% polyacrylamide gels. In these instances, Pol δ was added after clamp loading and reactions were started by adding a mix of dNTPs and heparin. Quenched reactions were then counted, dried, and redissolved in 2.0 μL formamide loading dye. Immediately prior to gel loading, samples were heated at 90 ℃ for 10 minutes, and gels were run at ~50 W for 5 hours in 1x TBE buffer. Polyacrylamide gels were then exposed and imaged by phosphorimager analysis.

Tris-HCl, NaCl, $MgAc_2$, BSA, and heparin were purchased from Sigma-Aldrich, while dNTPs and ATP were from NEB. dNTPs, ATP, and $MgAc_2$ were thoroughly degassed prior to reaction, and the protein stocks were kept open during a series of vacuum/nitrogen/gas mix purges to minimize residual oxygen.

Chemical Oxidation of Pol δ

For photooxidation, the 31-mer M13mp18 primer was covalently modified with a 5′ anthraquinone (AQ). AQ was prepared as a phosphoramidite and added to the unmodified DNA on a DNA synthesizer according to previously reported procedures (*36*). The presence of AQ was verified by MALDI-TOF, and the modified primer was annealed to M13mp18 DNA in Pol δ

activity buffer (50 mM Tris-HCl, pH 7.8, 50 mM NaCl). Because the 5′ AQ modification prevented ^{32}P end-labeling, DNA was labeled by adding 2 μCi [α-^{32}P] dATP (Perkin Elmer) to the reaction, and incorporation of [α-^{32}P] dATP was facilitated by lowering the concentration of cold dATP from 80 μM to 10 μM.

Anaerobic reaction mixes lacking dNTPs were prepared in glass vials and incubated under a solar simulator equipped with a UVB/C long pass filter or in the dark for 30 minutes. To ensure complete oxidation, 2-fold molar excess of both PCNA and Pol δ DV were included. As controls, reactions were also run using unmodified DNA (no AQ) and AQ reactions were repeated with 140 nM Klenow fragment exo$^-$ (NEB). After treatment, samples were returned to the glove bag and transferred into Eppendorf tubes containing dNTPs to start the reaction. Free dNTPs were removed using BioRad Microspin 6 columns in SCC buffer (GE Healthcare) and sample radioactivity was quantified on a liquid scintillation counter. Samples were then run out on a 1% alkaline agarose gel and visualized by phosphorimaging analysis. Overall [α-^{32}P] dATP incorporation was used to compare overall DNA synthesis levels by dividing the total radioactivity counts in oxidized samples by those of dark controls.

Results

Electrochemical Characterization of Pol δ

To determine whether Pol δ holoenzyme was redox active in the presence of DNA, we carried out electrochemistry on DNA-modified gold electrodes. In initial experiments, 3 μM WT Pol δ in storage buffer was combined with 10 μM PCNA and incubated on the electrode for several hours. CV scans taken hourly reveal a reversible signal with a midpoint potential of 116 ± 3 mV vs NHE (Figure 5.1). This signal grew in over time to reach a maximum size of 41 ± 4 nC and -51 ± 2 nC for the reductive and oxidative peaks at a 100 mV/s scan rate after two hours of incubation (Figure 5.1). The CV current varied linearly with the square root of the scan rate (Figure 5.2), yielding a diffusion coefficient D on the order of 7 x 10^{-6} $cm^2 \cdot s^{-1}$ upon application of the Randles-Sevcik equation. A linear dependence on the root of the scan rate is characteristic of diffusive rather than adsorbed species, which instead show a linear relationship between scan rate and current (*34*). Diffusive behavior has been observed previously for DNA repair proteins under similar conditions, and would indeed be expected for a polymerase-sliding clamp complex on a free DNA end (*14*). No differences were observed between aerobic and anaerobic electrochemistry carried out in a glove bag, indicating that the cluster is relatively stable in air and consistent with the general long-term stability of B-family DNA polymerases (*37*). The redox couple observed was attributed to the $[4Fe4S]^{3+/2+}$ based on the fact that Pol δ is HiPIP-like, being EPR-silent unless oxidized (*7*). In addition, the electrochemical signal is similar to the DNA glycosylases EndoIII and MutY, in which the identity of the couple has been established by EPR (*13, 24*).

We could obtain quantifiable signals at lower concentrations by adding dNTPs and Mg^{2+} to enhance protein association with the DNA. To prevent degradation of the DNA substrate by

the 3′-5′ exonuclease activity of WT Pol δ, we turned to the exonuclease-deficient mutant Pol δ DV (D520V) for these experiments (*33*). At 113 ± 5 mV vs NHE, the midpoint potential of Pol δ DV is indistinguishable from WT (Figure 5.1). By adding 80 μM dATP (the incoming nucleotide), 8.0 mM $MgAc_2$, and 5.0 μM PCNA, we were able to see signals with Pol δ concentrations as low as 500 nM (Figure 5.1). Under these conditions, the maximum CV peak areas were 6.9 ± 1 nC and -7.5 ± 1 nC for the reductive and oxidative peaks at a scan rate of 100 mV/s.

The Pol δ midpoint potential as measured resides within the HiPIP regime, but it is slightly higher than the 65-95 mV vs NHE reported for DNA-bound repair proteins (*13, 28*). To determine if this was an intrinsic difference between these proteins or the result of different buffer conditions, we performed electrochemistry on the well-studied *E. coli* BER glycosylase EndoIII following exchange into Pol δ storage buffer (*28*). In standard phosphate storage buffer (20 mM sodium phosphate, pH 7.5, 100 mM NaCl, 1 mM EDTA, 10% glycerol v/v), the EndoIII midpoint potential is ~80 mV versus NHE (*28*). When we exchanged EndoIII into Pol δ storage buffer, CV and SQWV scans of 140 μM EndoIII resulted in a midpoint potential of 113 ± 3 mV versus NHE, which is indistinguishable from that of Pol δ (Figure 5.3). Notably, this result held even when EndoIII was diluted to 1.5 μM to more closely approximate the conditions used for polymerase experiments (Figure 5.3). UV-visible spectra confirmed the stability of EndoIII in Pol δ buffer (Figure 5.3a). That the EndoIII midpoint potential matches that of Pol δ confirms that there is no significant difference between the [4Fe4S] cluster between these two proteins. The potential increase in Pol δ storage buffer is thus most likely the result of higher ionic strength (350 mM NaAc in Pol δ buffer vs 150 mM NaCl in EndoIII buffer) (*38, 39*). In any

case, these results support the assertion that the [4Fe4S] cluster resides in the same narrow potential regime in a wide range of DNA-binding proteins.

Given that Pol δ ordinarily functions in complex with PCNA, we next asked what effect PCNA might have on the electrochemical properties of Pol δ. In the absence of PCNA, the midpoint potential is unaltered at 115 ± 8 mV versus NHE, but the signal was markedly smaller, reaching a maximum CV peak area of 0.4 ± 0.1 nC for the reductive peak and -0.7 ± 0.1 nC for the oxidative peak (Figure 5.4). The signal also decayed more rapidly with PCNA absent, suggesting lower polymerase stability. To more quantitatively compare the signals obtained in the presence and absence of PCNA, we measured diffusion coefficients (D) under both conditions, as well as with and without dATP, by applying the Randles-Sevcik equation to the scan rate dependence of the current (*34*). To confirm that our values were within the expected range, we also calculated D from the Stokes-Einstein equation, using a hydrodynamic radius obtained from several crystal structures. At maximum signal size, D was found to be 6.7 ± 3 x 10^{-6} $cm^2 \cdot s^{-1}$ with both PCNA and dATP present, which is within one order of magnitude of the Stokes-Einstein estimate (6.0 x 10^{-7} $cm^2 \cdot s^{-1}$). The difference between these values most likely reflects the use of simplified geometries calculated from partial crystal structures. In the absence of PCNA, D decreased by more than an order of magnitude to 1.2 ± 0.3 x 10^{-7} $cm^2 \cdot s^{-1}$, lower than either our experimental values with PCNA present or our Stokes-Einstein estimates, indicating a change in complex shape. As Pol δ has multiple contacts with PCNA, a change in shape is expected, and our result is fully consistent with earlier studies that indicates an elongate form for Pol δ alone in solution and a more compact form when bound to PCNA (*7, 40*). To see if dNTPs can also contribute to the shape of PCNA-bound Pol δ, we prepared a surface with Pol δ and PCNA but lacking dATP and Mg^{2+}. Under these conditions, the signal is comparable to that in

the absence of PCNA, giving a D value of $2.2 \pm 0.7 \times 10^{-7}$ $cm^2 \cdot s^{-1}$ and suggesting that, even with PCNA present, Pol δ is less compact without dNTPs. Taken together, these results indicate that PCNA does not affect the potential of Pol δ but is critical for effective DNA binding. The differences in signal size and stability suggest that, at least at nanomolar concentrations, the Pol δ [4Fe4S] cluster is most strongly coupled to the electrode when both PCNA and dNTPs are present. The importance of PCNA and dNTPs to DNA binding further suggest that the signals observed were from DNA-bound protein.

Given the experimental support for the observed electrochemical signals being DNA-bound, we next aimed to determine if Pol δ can signal in a DNA-mediated fashion. To this end, we prepared multiplexed chips containing DNA modified with either an abasic site or a CA mismatch 6 nucleotides from the thiolated end (Figure 5.1). Because DNA-bound proteins are capable of charge transport both through the DNA bases and directly through the monolayer surface, we also varied monolayer morphology in these experiments (Figure 5.5) (*28*). Thiolated DNA monolayers can be self-assembled in two basic ways: overnight incubation with 100 mM $MgCl_2$ forms closely packed islands of DNA (30-50 $pmol \cdot cm^{-2}$), while standard loosely packed monolayers (15-20 $pmol \cdot cm^{-2}$) form in the absence of Mg^{2+} (Figure 5.) (*41, 42*). Loosely packed monolayers had been used exclusively to this point to avoid steric crowding with large protein complexes, but the more densely covered closely packed monolayers have been shown to be more amenable to DNA-mediated signaling in experiments with EndoIII (*28*). To test each of these conditions, chips were prepared with one of the two monolayer forms, with the device divided evenly between well-matched and either abasic or CA mismatch DNA (Figure 5). All experiments were performed with 500 nM Pol δ DV in the presence of PCNA, dATP, and $MgAc_2$.

Consistent with previous studies on EndoIII (*28*), Pol δ redox potentials were identical on both film morphologies, supporting the assertion that all observed signals are from DNA-bound proteins (Figure 5). On closely packed DNA films, signal size was highly variable, but 46 ± 33% signal attenuation as determined by SQWV occurred on abasic DNA (Figure 5.5). No mismatch discrimination was observed, and even the abasic site discrimination decreased over time as more protein diffused to the surface. In general, the signals on closely packed monolayers are consistent with significant steric hindrance causing protein-DNA complexes to lie flat on the surface (Figure 5.5b). In contrast, signals on loosely packed monolayers were of very consistent size and showed significant charge attenuation with both abasic and CA-mismatch DNA, reaching maxima of 44 ± 16% on abasic DNA and 46 ± 29% on CA-mismatch DNA after 2 hours of incubation as measured by SQWV (Figure 5.5). Abasic site and mismatch discrimination remained stable over several hours, suggesting less steric hindrance not on loosely packed DNA. Together, these results confirm that Pol δ is capable of DNA-mediated signaling and emphasize the importance of substrate accessibility when assessing CT by large protein complexes.

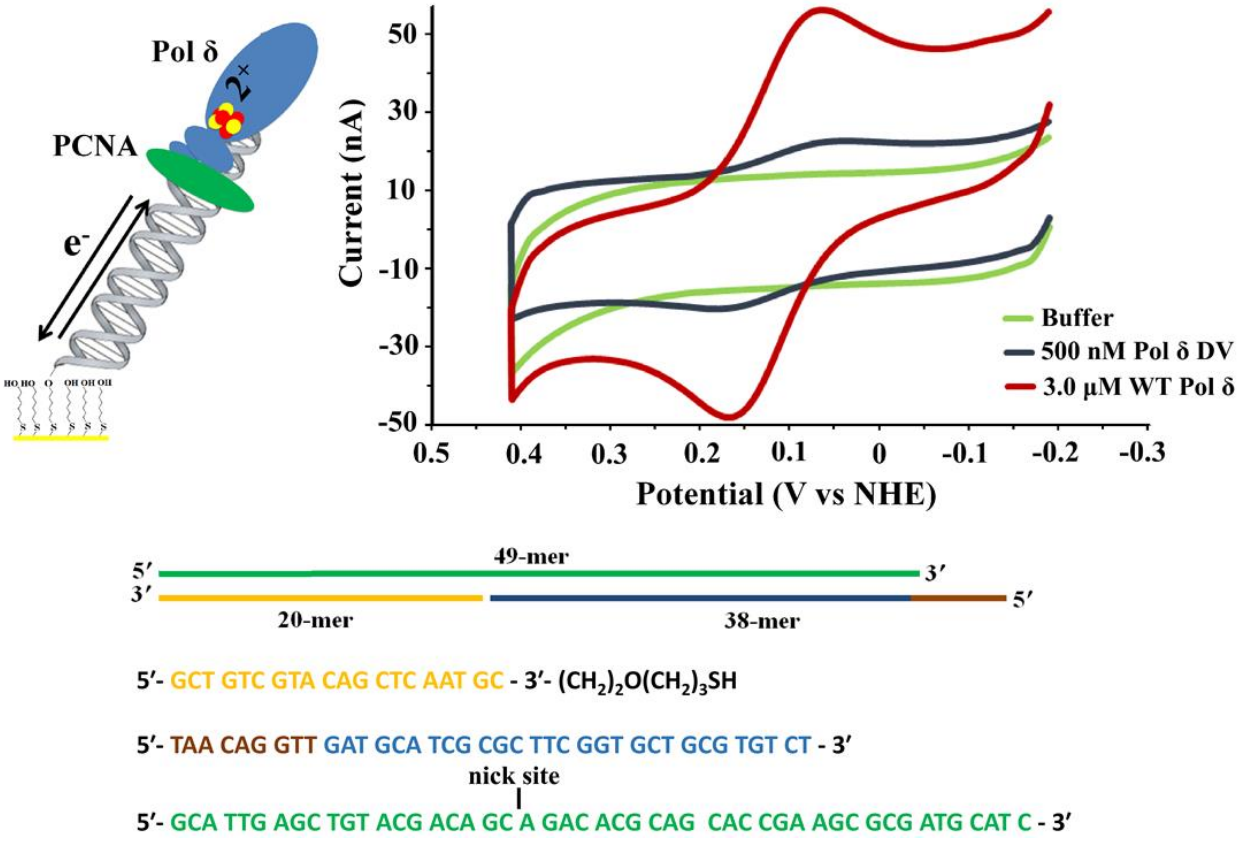


Figure 5.1 Cyclic voltammetry with WT and exonuclease-deficient Pol δ. (**Top**) The addition of 3.0 μM WT Pol δ and 10 μM PCNA to a DNA-modified gold electrode resulted in a reversible CV signal with a midpoint potential centered at 113 ± 4.7 mV versus NHE (red trace). By using the exonuclease deficient mutant Pol δ DV, which cannot degrade DNA in the presence of catalytic magnesium, signals could be obtained at concentrations as low as 500 nM when $MgAc_2$, 5 μM PCNA and excess dATP were included (blue trace). As expected, WT Pol δ and Pol δ DV have indistinguishable midpoint potentials at 116 ± 3 and 113 ± 3 mV versus NHE, respectively. (**Bottom**) DNA substrate design for electrochemical experiments. The substrate consists of 3 parts, a 20-mer thiol, a 38-mer and a 49-mer complement. Notably, the nick in the phosphate backbone is not expected to interfere with CT. Electrochemistry was performed in storage buffer (30 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.1 mM EDTA, 10% glycerol v/v, 0.01% decaethylene glycol monododecyl ether v/v), and the abasic site/mismatch discrimination experiments also included 5.0 μM PCNA, 80 μM dATP, and 8.0 mM $MgAc_2$. All CV scans were taken at a scan rate of 100 mV/s.

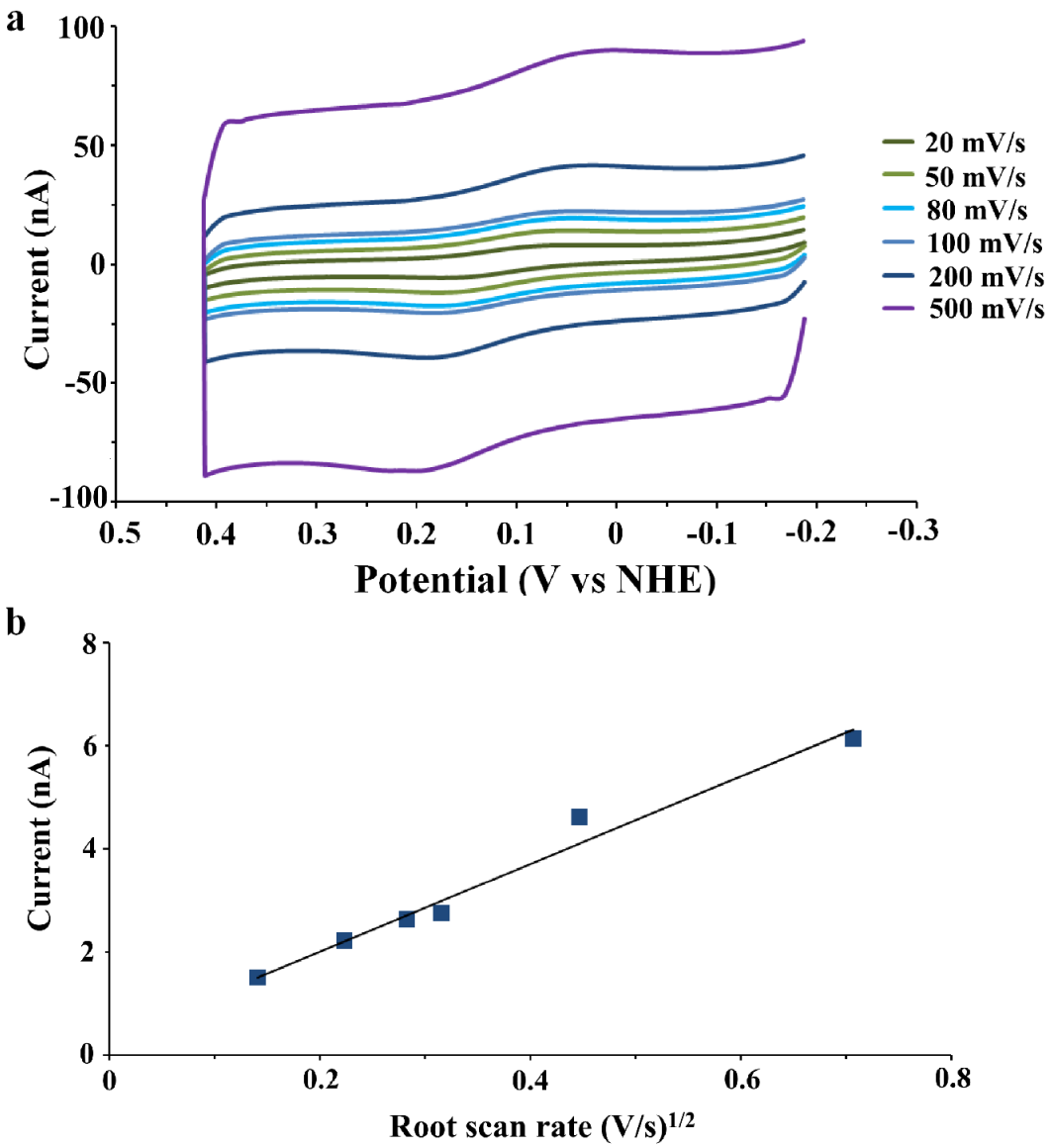


Figure 5.2 Scan rate dependence of the CV current in 500 nM Pol δ DV incubated with 5.0 μM PCNA, 80 μM dATP, and 8.0 mM $MgAc_2$. (**a**) The maximum peak current increases with increasing scan rate, coupled with an increase in peak splitting. (**b**) The current exhibits a linear dependence on the square root of the scan rate, characteristic of a diffusive rather than adsorbed species. The scan rates included are 20, 50, 80, 100, 200, and 500 mV/s. The line was fit to data averaged from 8 separate experiments, and the fit is $I = 7.7559\nu^{1/2} + 0.5725$ with an R^2 value of 0.9828.

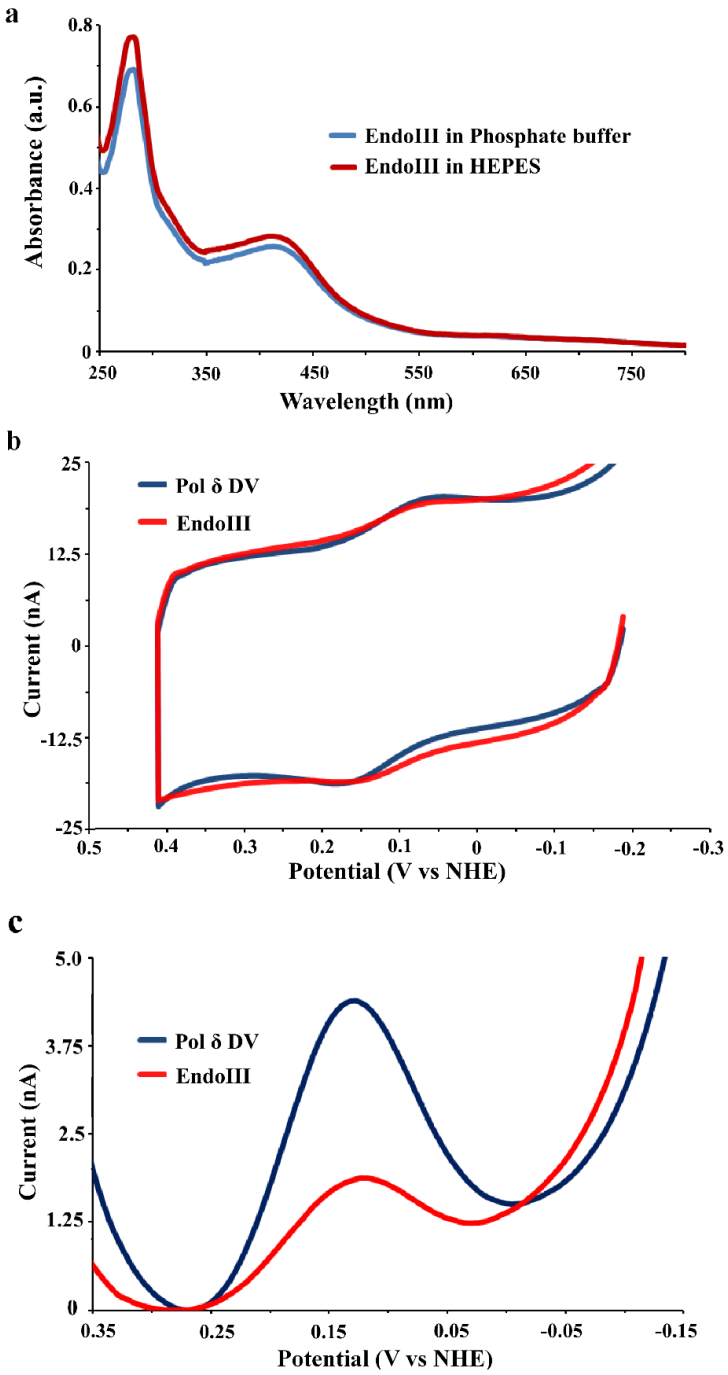


Figure 5.3 Pol δ and EndoIII electrochemistry compared. 1.5 μM EndoIII (stored in 20 mM sodium phosphate, pH 7.5, 150 mM NaCl, 1 mM EDTA) was exchanged into Pol δ storage buffer (30 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.1 mM EDTA, 10% v/v glycerol, 0.01% decaethylene glycol monododecyl ether w/v) and added to a multiplexed chip containing unmodified Pol δ DNA (49:58-mer substrate). (**a**)UV-visible spectra taken before and after buffer exchange confirm the stability of EndoIII in a HEPES-based buffer. (**b**) The midpoint potential as measured by CV is 113 ± 3 mV, virtually indistinguishable from Pol δ DV at 113 ± 5 mV versus NHE.

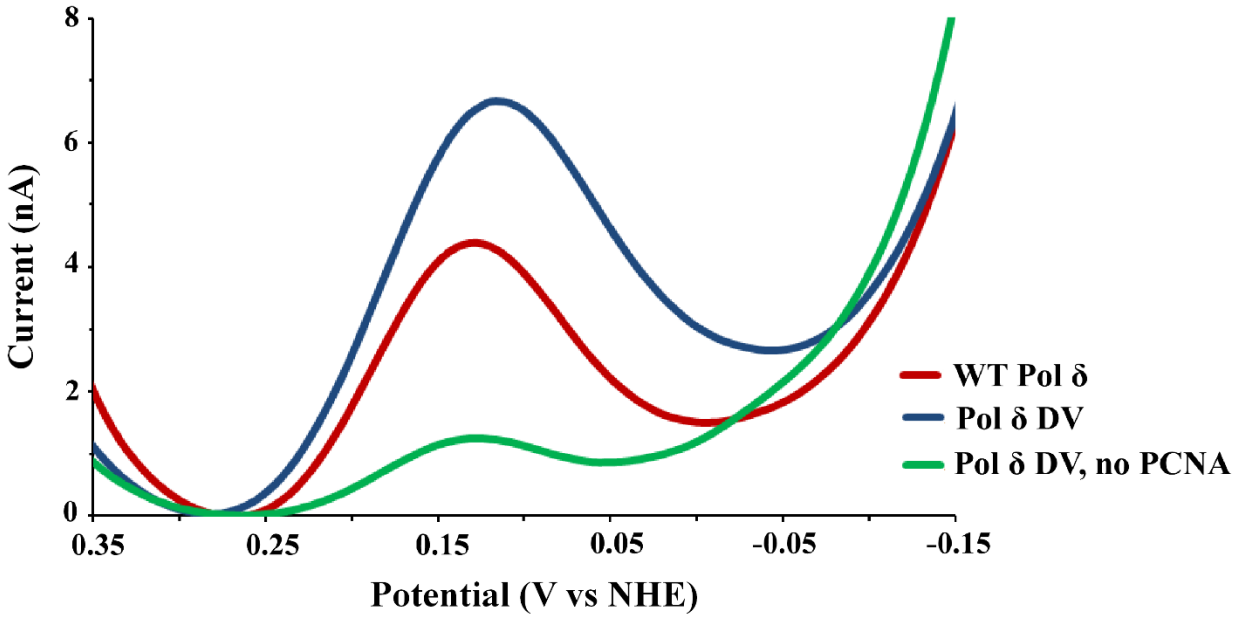


Figure 5.4 SQWV of 500 nM WT Pol δ and exonuclease-deficient Pol δ DV with and without 5.0 μM PCNA. WT and exo$^-$ Pol δ DV share the same potential, and both generate a substantial signal on a DNA-modified gold electrode; the smaller size of the WT signal may be due in part to DNA degradation by exonuclease activity. PCNA itself does not affect the potential, but its absence results in significantly decreased signal size and lower stability over time. SQWVs were taken at 15 Hz frequency and 25 mV amplitude, and electrochemistry was carried out in storage buffer (20 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.1 mM EDTA, 10% glycerol v/v, 0.01% decaethylene glycol monododecyl ether v/v) with 8.0 mM $MgAc_2$ and 80 μM dATP.

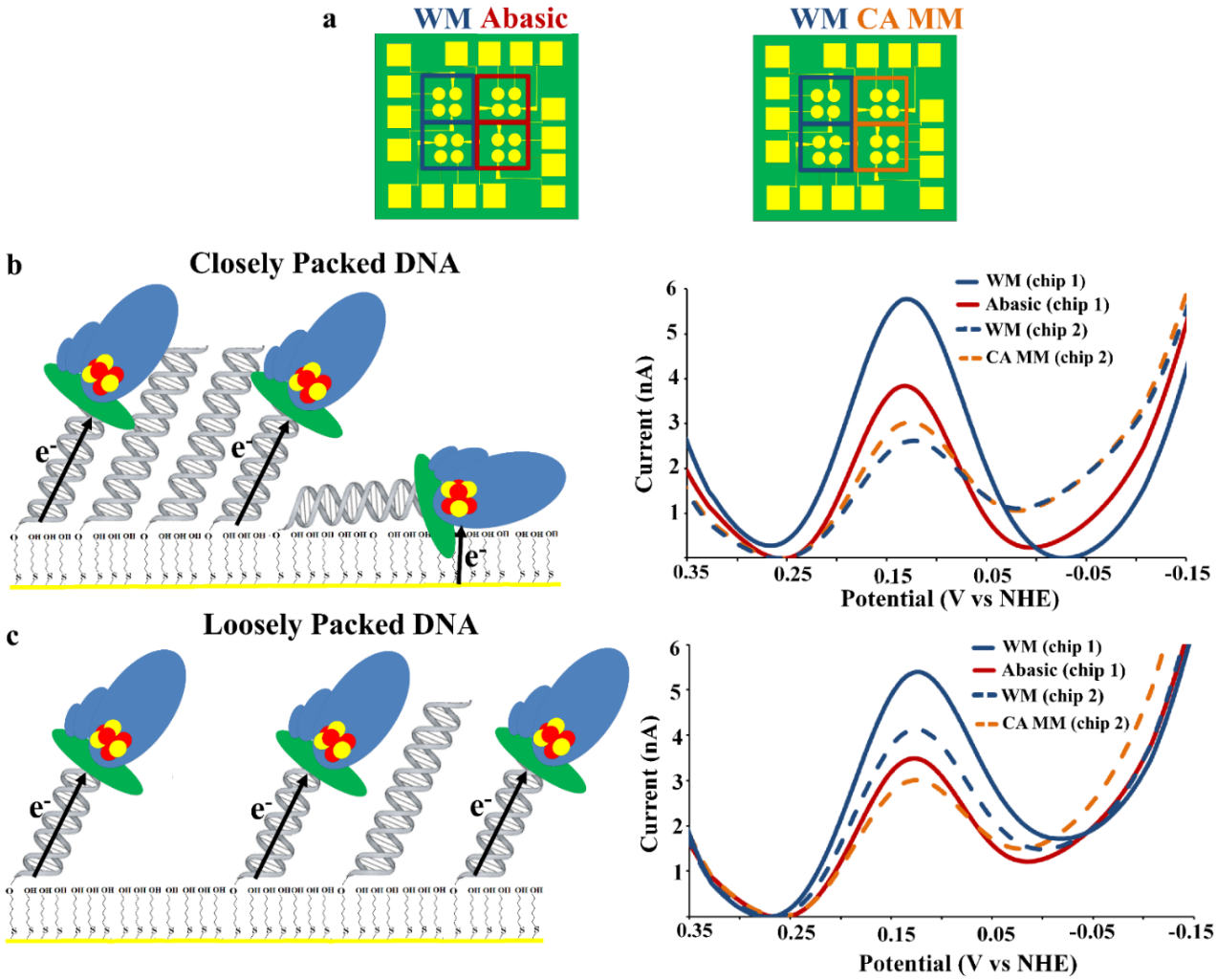


Figure 5.5 Pol δ electrochemistry on different DNA monolayer morphologies. (**a**) To find an optimal DNA monolayer morphology for Pol δ signaling, we prepared multiplexed chips containing either closely packed (assembled with 100 mM $MgCl_2$) or loosely packed (no $MgCl_2$) DNA films. Two chips were prepared for both morphologies, with one half of each chip consisting of well-matched (WM) DNA (dark blue) and the other containing DNA with either an abasic site (red) or a CA mismatch (orange) 6 nucleotides from the monolayer surface. (**b**) On closely packed films, Pol δ SQWV signals were highly variable and showed 46 ± 33% attenuation on abasic DNA (solid SQWV traces) but no significant mismatch discrimination (dashed SQWV traces). (**c**) In contrast, SQWV signals on loosely packed films were much more consistent between electrodes, with a 44 ± 16% signal loss on abasic DNA (solid traces) and 46 ± 29% signal loss with CA mismatch DNA (dashed traces). To minimize the effects of variability between devices, all direct comparisons were made on a single chip; scans that were directly compared are denoted by either solid or dashed lines in the SQWV signals shown. The SQWV traces shown are an average of 6 individual electrodes on a single device, with scans taken at 15 Hz frequency and 25 mV amplitude.

Activity Assays with Oxidized and Reduced Pol δ

Having seen that DNA binding can activate Pol δ for redox activity, we next asked if the cluster oxidation state might affect polymerase activity. As purified, Pol δ exists largely in the $[4Fe4S]^{2+}$ state (*7*), so any assessment of activity differences would require extensive oxidation to generate sufficient amounts of the $[4Fe4S]^{3+}$ cluster for comparison. For this task, we used bulk electrolysis on DNA-modified electrodes, applying an oxidizing potential of 0.412 V versus NHE for 15-20 minutes under an inert atmosphere (95% N_2/5% H_2) to prevent cluster degradation (Figure 5.6a).

High bulk electrolysis yields generally require an electrode with substantial surface area, so we switched from multiplex chips to single gold rod electrodes for these experiments. Multiplexed chips have many advantages for electrochemical characterization, but only a single electrode can be addressed at a time and each sample in a quadrant is distributed across 4 separate electrodes. To confirm the effectiveness of this system, a sample of concentrated (150 μL of 2.74 μM) Pol δ was oxidized on a DNA-modified electrode in a custom-made glass cell for several hours (Figure 5.6b). UV-visible spectra were taken before and after electrolysis, after which the sample was then frozen for EPR in parallel with untreated protein (Figure 5.6c, d). The [4Fe4S] cluster oxidation generally results in a broad increase in UV-visible absorbance from 300-450 nm with a less distinct peak at 410 nm in both the $[4Fe4S]^{3+}$ and $[3Fe4S]^{+}$ species (*43*–*45*). After bulk electrolysis, increased absorbance from 300-400 nm, consistent with cluster oxidation, was indeed observed. No significant increase in absorbance at 800 nm occurred after oxidation, and the 280 nm peak associated with aromatic and thiolated amino acid residues remained distinct. From our own observations, protein aggregation tends to generate a U-shaped curve with high absorbance at 800 nm and a shallow, poorly defined peak at 280 nm, and the

lack of these features in our spectra indicate that oxidized Pol δ did not aggregate (Figure 5.6c). EPR signals are small as a result of the low sample concentration, but clear signals at g = 2.08 and g = 2.02 are present in the oxidized sample (Figure 5.6d). These signals are consistent with a combination of $[4Fe4S]^{3+}$ and $[3Fe4S]^{+}$ cluster oxidation products (*24*, *44-45*). A smaller signal at g = 2.02 was also present in the native sample, consistent with earlier reports of residual $[3Fe4S]^{+}$ cluster in untreated Pol δ (*7*). That some $[3Fe4S]^{+}$ cluster would occur upon oxidation is not surprising, and similar results have been obtained for EndoIII and MutY (*13*, *24*). In earlier studies, loss of iron was attributed in part to damage incurred upon freezing, which may have also happened here. Furthermore, oxidized Pol δ was stored away from protective DNA long enough to take a UV-visible spectrum prior to freezing. In any case, some $[4Fe4S]^{3+}$ cluster was still observed, and the $[3Fe4S]^{+}$ cluster that did occur would have formed as a degradation product of the $[4Fe4S]^{3+}$ cluster (*13*).

Because activity assays require only low nanomolar polymerase concentrations, bulk electrolysis for these experiments was carried out with 190 nm Pol δ DV to minimize sample waste. Oxidation yields under these conditions were higher, typically around 75 – 90% as determined from the total charge passed. After electrolysis, untreated or oxidized Pol δ DV was added directly to pre-made reaction mixes to a final concentration of 2 nM (Figure 5.7a). When run out on an alkaline agarose gel, it is apparent that at early time points less DNA synthesis was carried out by oxidized Pol δ (Figure 5.7, 8). DNA synthesis can be more quantitatively compared for the oxidized versus untreated sample by dividing the amount of frontier products (highest molecular weight major products) in the oxidized sample by the amount present in untreated samples. Using this analysis, oxidized Pol δ at 60-80% yield forms only 30-50% as many large (~ 5 kb) DNA products as untreated Pol δ after 30 seconds (Figure 5.7c).

Significantly, higher oxidation yields lead to lower activity levels. In any case, this difference gradually decreases over the course of 10 minutes. Regardless of oxidation state, no DNA synthesis occurs in samples lacking PCNA, confirming that all observed DNA synthesis is processive (Figure 5.8c).

Reduction of the oxidized Pol δ stock by electrolysis at -0.188 V versus NHE effectively restores DNA synthesis, reaching 90% of untreated levels at early time points (Figure 5.7, 8). Critically, this result both confirms the reversibility of oxidative slowing and provides support for the $[4Fe4S]^{3+}$ cluster as the major oxidation product. As mentioned earlier, the reversible electrochemical signals are consistent with $[4Fe4S]^{3+/2+}$ cycling, but EPR spectroscopy with oxidized Pol δ showed evidence of both $[4Fe4S]^{3+}$ and $[3Fe4S]^{+}$ products in the sample, leading to some ambiguity. However, the nearly complete restoration of native activity levels upon re-reduction would not be expected if most of the cluster had degraded to the $[3Fe4S]^{+}$ state, supporting the $[4Fe4S]^{3+}$ cluster as the major oxidation product. These combined results thus indicate that the $[3Fe4S]^{+}$ cluster seen by EPR likely forms after the $[4Fe4S]^{3+}$ major product degrades over time in the absence of DNA; we have observed this previously with sample freezing for *E. coli* EndoIII and MutY following chemical oxidation (*13, 24*).

To gain further insight into the effect of oxidation, reaction rates were estimated by comparing frontier velocities. Velocities were calculated by dividing the amount of the largest quantifiable band of DNA by time and the number of polymerase molecules present (*46*). This method yields maximum rates of 118 ± 63 (SD, n = 7) nt/s per enzyme for untreated Pol δ and 21 ± 27 (SD, n = 5) nt/s for oxidized Pol δ at 2 minutes. The rates obtained for untreated Pol δ are consistent with previously published *in vitro* results (*35*), while the oxidized form is significantly slower. This calculation represents just an upper estimate, as typical bulk electrolysis fails to

oxidize around 20-30% of the enzyme; thus, some of the DNA synthesis observed in oxidized samples can be attributed to the non-oxidized population. Indeed, the comparable amounts of DNA synthesis observed after 5-10 minutes could have resulted from either slow oxidized polymerase catching up or redistribution by of residual native Pol δ in the oxidized sample. Overall, it is clear from these experiments that oxidation leads to a decrease in replication rate, but resolution on alkaline agarose gels is insufficient to distinguish between complete stalling or dramatic slowing of DNA synthesis.

To distinguish between stalling and slowing of DNA synthesis by oxidized Pol δ, reactions were analyzed on 5% polyacrylamide gels to obtain increased resolution in the 30-1000 nucleotides range. In addition, DNA synthesis was limited to that of a single processive cycle by the PCNA-Pol δ complex by adding heparin, which traps dissociated Pol δ (*35*). Without the heparin trap, products up to 7 kb were observed, due to multiple processive cycles of synthesis (Figure 5.7, 8). In order to visualize products at all sizes, reactions containing heparin were divided in two, with half loaded onto a polyacrylamide gel and half onto an alkaline agarose gel. With heparin present, alkaline agarose gels demonstrate a severe limitation to DNA synthesis, with no products larger than ~1 kb observed on alkaline agarose gels (data not shown). When these products are resolved on polyacrylamide gels, a greater proportion of both very small (~primer length) and very large (~1 kb) products are formed by untreated Pol δ, while the oxidized form generates more intermediate products between 30 and 1000 nucleotides (Figure 5.9).

These results illustrate several important points about the effects of oxidation on Pol δ DNA synthesis. First, they verify that the oxidized form remains active and does not completely stall. Second, the relatively greater amounts of very small products and unextended primers in

reactions with native sample indicate greater susceptibility to dissociation from DNA and trapping by heparin, while the native form that does associate produces longer products. In contrast, oxidized Pol δ leaves fewer primers unextended or fully extended, instead making more intermediate products between 150 bp and 1 kb. The greater proportion of extended primers is consistent with tighter DNA binding after cluster oxidation, as has been observed with both primase and DNA repair proteins (*16, 27*). However, the slower procession indicates that tighter binding impedes rapid procession, acting as a brake on PCNA-mediated DNA synthesis. These experiments suggest that the similar activity levels observed on alkaline agarose gels at time points beyond 5 minutes could be explained by either the oxidized form gradually catching up or by redistribution of the residual native enzyme in the sample. Regardless of the precise details, the overall impact of polymerase stalling, with a 6-fold decrease in rate, would be significant on the timescale of S-phase; unperturbed yeast S-phase lasts ~30 minutes, while using oxidized Pol δ moving at 20 nt/s to replicate the lagging strand of the yeast genome would itself require 27 minutes (*47-49*).

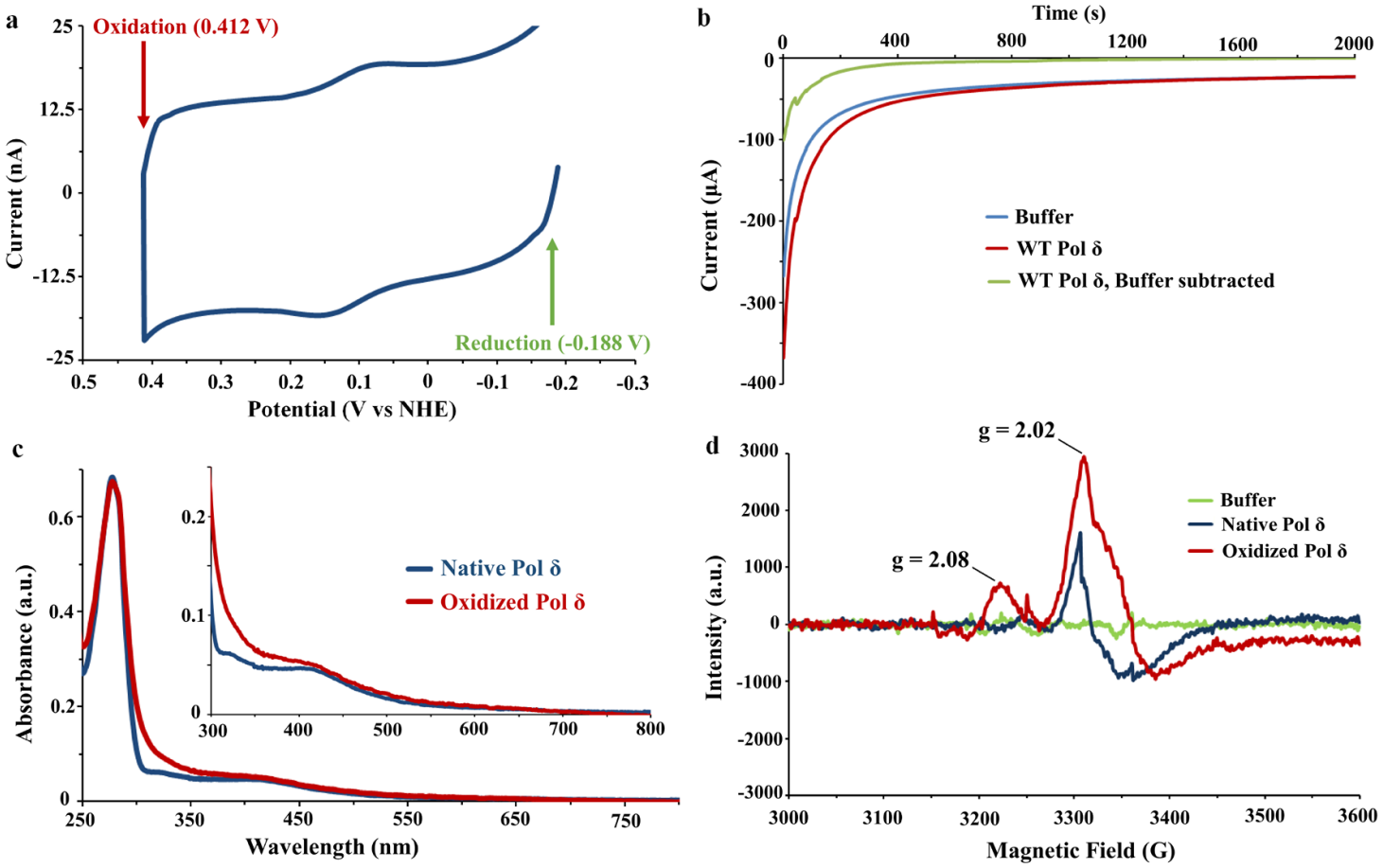


Figure 5.6. Characterization of electrochemically oxidized Pol δ. (**a**) Bulk electrolysis potentials were ~200 mV beyond the major oxidative and reductive peaks at 0.412 V (oxidation) and -0.188 V (reduction) versus NHE. (**b**) Yields were calculated by subtracting a background electrolysis (blue) from one containing protein (red) and taking the area under the resultant curve (green). Electrolysis of 150 µL of 2.74 µM Pol δ at 0.412 V gave ~35% oxidation yield. (**c**) UV-visible spectra reveal an increased absorbance from 300-400 nm consistent with cluster oxidation with no evidence of protein aggregation. (**d**) CW X-band EPR spectra at 10 K reveal the presence of both $[4Fe4S]^{3+}$ (g = 2.08) and $[3Fe4S]^{+}$ (g = 2.02) species in the oxidized sample, with a residual amount of $[3Fe4S]^{+}$ cluster present in the native sample. These results are consistent with the formation of $[4Fe4S]^{3+}$ cluster after anaerobic bulk electrolysis, with some degrading to form $[3Fe4S]^{+}$ cluster in the absence of DNA. As slight sample loss did occur following oxidation, the UV-visible spectrum of oxidized Pol δ has been normalized to native absorbance at 280 nm to afford a more direct comparison. EPR spectra were taken at 12.85 mW microwave power, 2 G modulation amplitude, and a receiver gain of 5.02 $x10^3$.

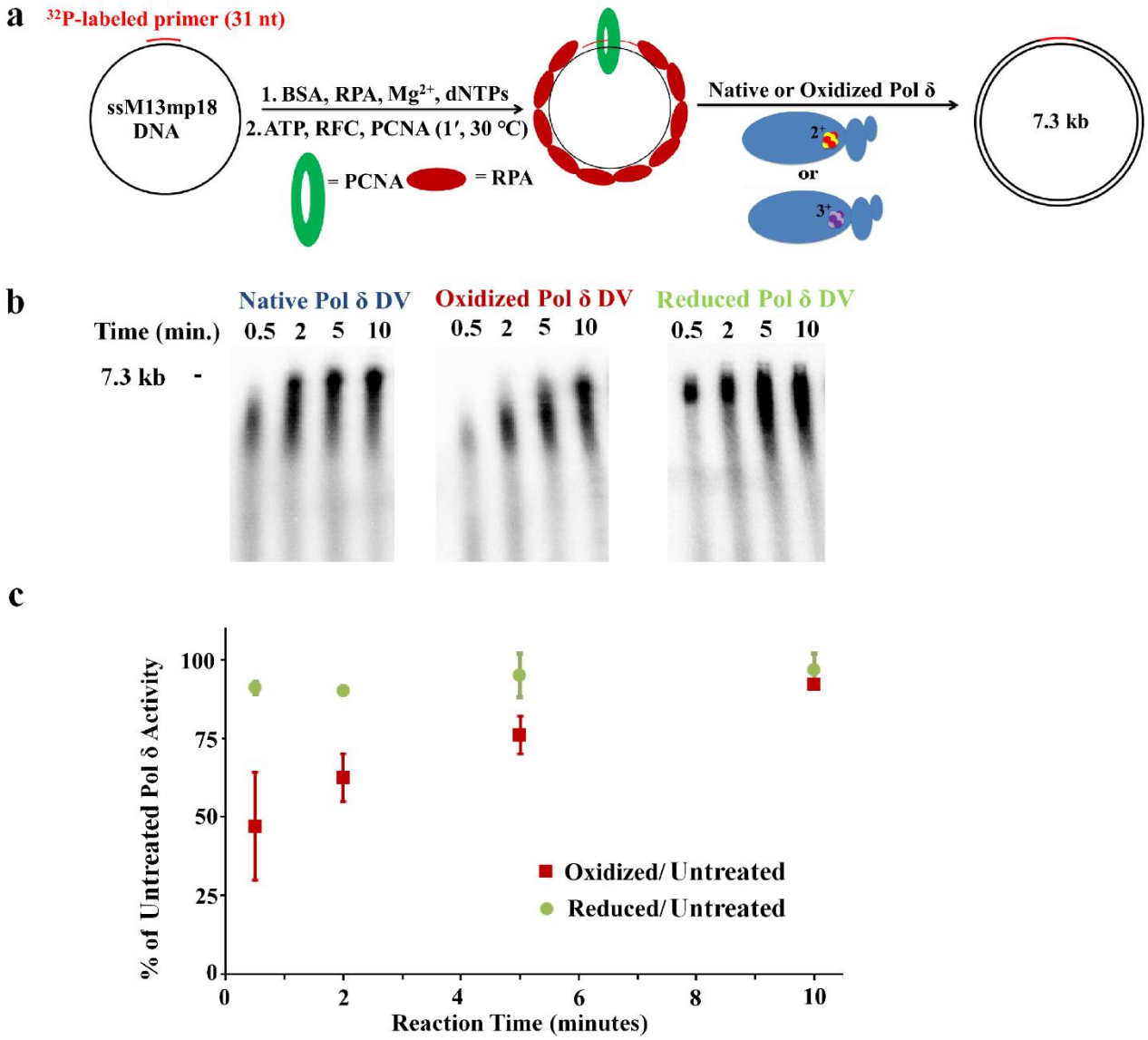


Figure 5.7 Activity assays with native and electrochemically oxidized Pol δ DV. (**a**) 190 nM Pol δ DV was oxidized or reduced by bulk electrolysis at potentials of 0.412 V and -0.188 V and subsequently diluted to 2 nM final concentration into reaction mixes containing radiolabeled M13mp18 DNA. (**b, c**) As seen on representative 1% alkaline agarose gels, oxidation lowers activity levels at early time points, while reduction restores activity to native levels. The degree of this effect can be quantified by dividing the amount of DNA synthesis in reactions with oxidized or reduced Pol δ by that from reactions with untreated enzyme. The oxidation yield for the gel shown in **b** is ~80%. Error bars are standard deviation of the mean ($n \geq 3$).

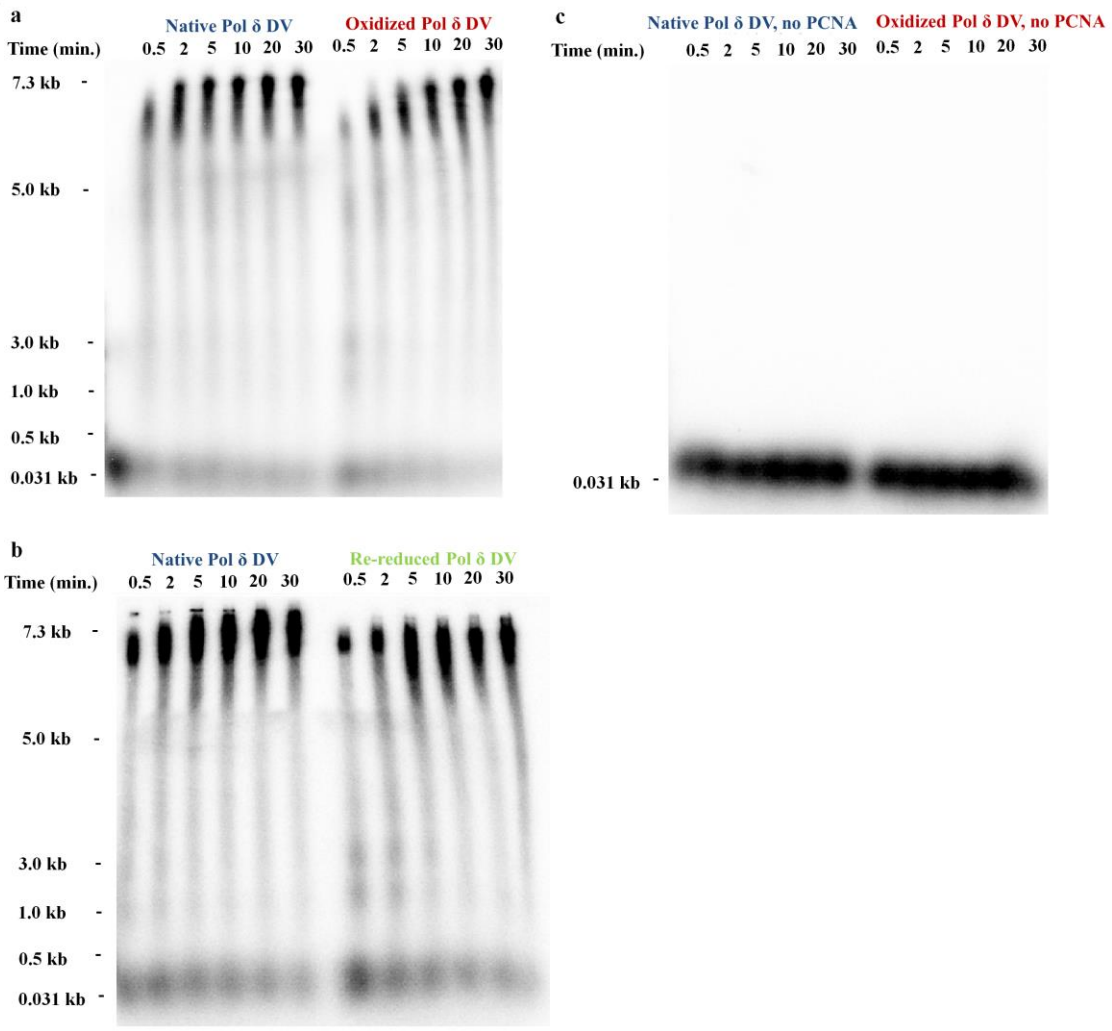


Figure 5.8 Complete alkaline agarose gels from Figure 5.7 and control lacking PCNA. The gels include untreated and oxidized Pol δ DV with 5.0 nM PCNA (**a**), untreated and re-reduced Pol δ DV (**b**), and untreated and oxidized Pol δ DV in the absence of PCNA (**c**). No DNA synthesis occurs in the absence of PCNA, confirming that the observed activity in native and oxidized samples is processive.

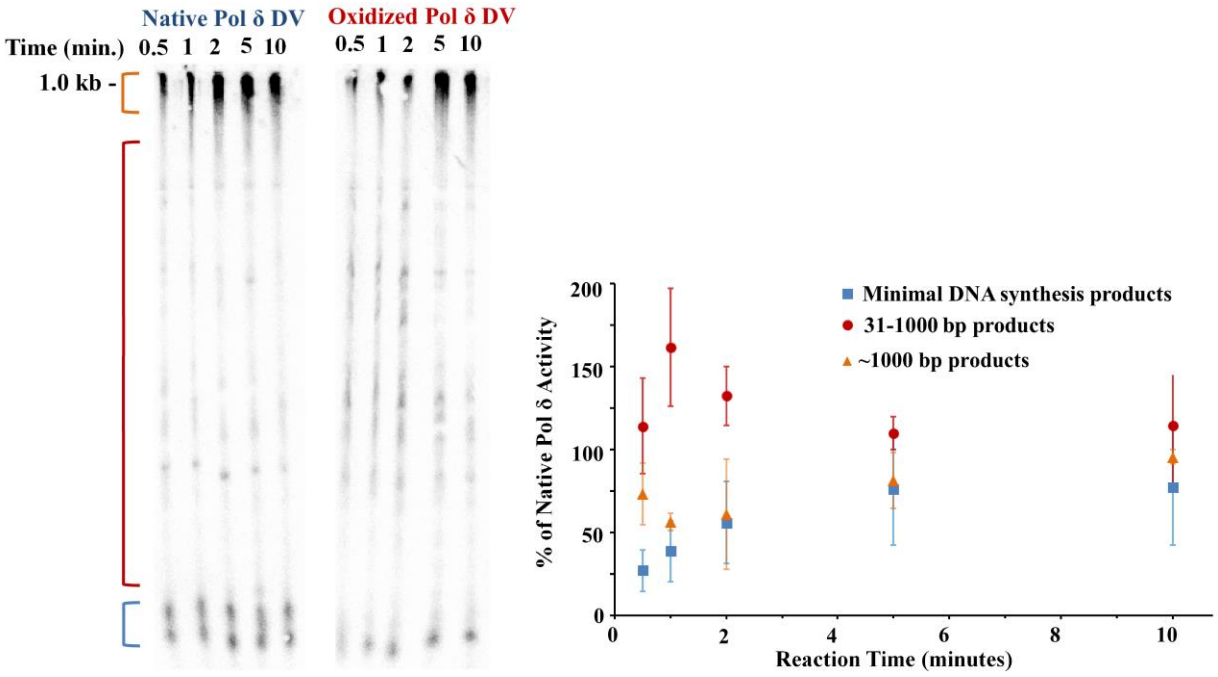


Figure 5.9 Establishment of activity by oxidized Pol δ. To see if oxidized Pol δ remained active or stalled completely, 0.01% heparin was included in reactions to challenge synthesis and products were analyzed on a 5% denaturing polyacrylamide gel to resolve DNA between 30 and 1000 bp (left). Pol δ remains active after oxidation, primarily forming intermediate-sized products (red range on gel). Native Pol δ is more sensitive to heparin, with more DNA close to primer length (blue range), but when it does associate with DNA, most products are around the maximum size (orange range). These results are consistent with tighter binding and slower processive DNA synthesis by the oxidized form. Gels were quantified using ImageQuant software; as synthesis appears as smears at this resolution, the total amount of background-subtracted radioactivity in each major range shown was compared between untreated and oxidized Pol δ. Error bars are standard deviation of the mean (n = 3).

Chemical Oxidation of Pol δ with Anthraquinone

Electrochemical oxidation provides clear advantages in estimating yields and re-reducing the oxidized sample, but the use of chemical oxidants is much more common. Thus, we were interested to see if chemical oxidation could yield an equivalent result. To this end, we used an anthraquinone (AQ)-derived photooxidant covalently tethered to the 5′ end of the DNA primer (*36, 50*). AQ has the advantage of oxidizing samples in a DNA-mediated fashion instead of the less effective, direct oxidation of the protein by oxidants in solution (*7, 50*). Irradiation at 350 nm of AQ generates an excited triplet state capable of oxidizing DNA bases, and AQ has been studied extensively in the context of DNA CT (*50-52*). The DNA base of lowest potential, guanine, has a redox potential of 1.29 V, which is considerably higher than that of the Pol δ [4Fe4S] cluster at 113 mV (*53*). The presence of AQ on the primer, however, prevents 5′ ^{32}P end-labeling, so [α-^{32}P] dATP added to reactions was used as an alternative label (Figure 5.10a). Since [α-^{32}P] dATP is not necessarily incorporated in a 1:1 ratio with DNA-primed ends, activity levels were compared by total scintillation counts rather than by directly quantifying gel bands.

Following irradiation, samples with AQ showed lower overall DNA synthesis relative to identical samples kept in the dark (Figure 5.10b, c). As with electrochemistry, the maximum differentials occur at earlier time points. After 30 and 60 seconds, irradiated Pol δ showed 40-60% of dark control DNA synthesis. In contrast, reaction mixes irradiated in the presence of unmodified DNA are not significantly different from dark controls at early time points, and remain equal or greater throughout the time course (Figure 5.10c, 11). Similarly, no significant differences are observed between reactions with irradiated or untreated *E. coli* Klenow fragment exo$^-$ on AQ-modified DNA (Figure 5.11), indicating that the attenuation observed in Pol δ DNA synthesis can be attributed to [4Fe4S] cluster oxidation. Assuming the activity differential at

early time points approximated the percentage of oxidized sample, as was the case electrochemically, photooxidation yields ranged from 40-50%. Overall, the pattern of attenuated activity after photooxidation is consistent with the results from electrochemical experiments (Figure 5.10c), independently confirming the slowing of Pol δ upon cluster oxidation.

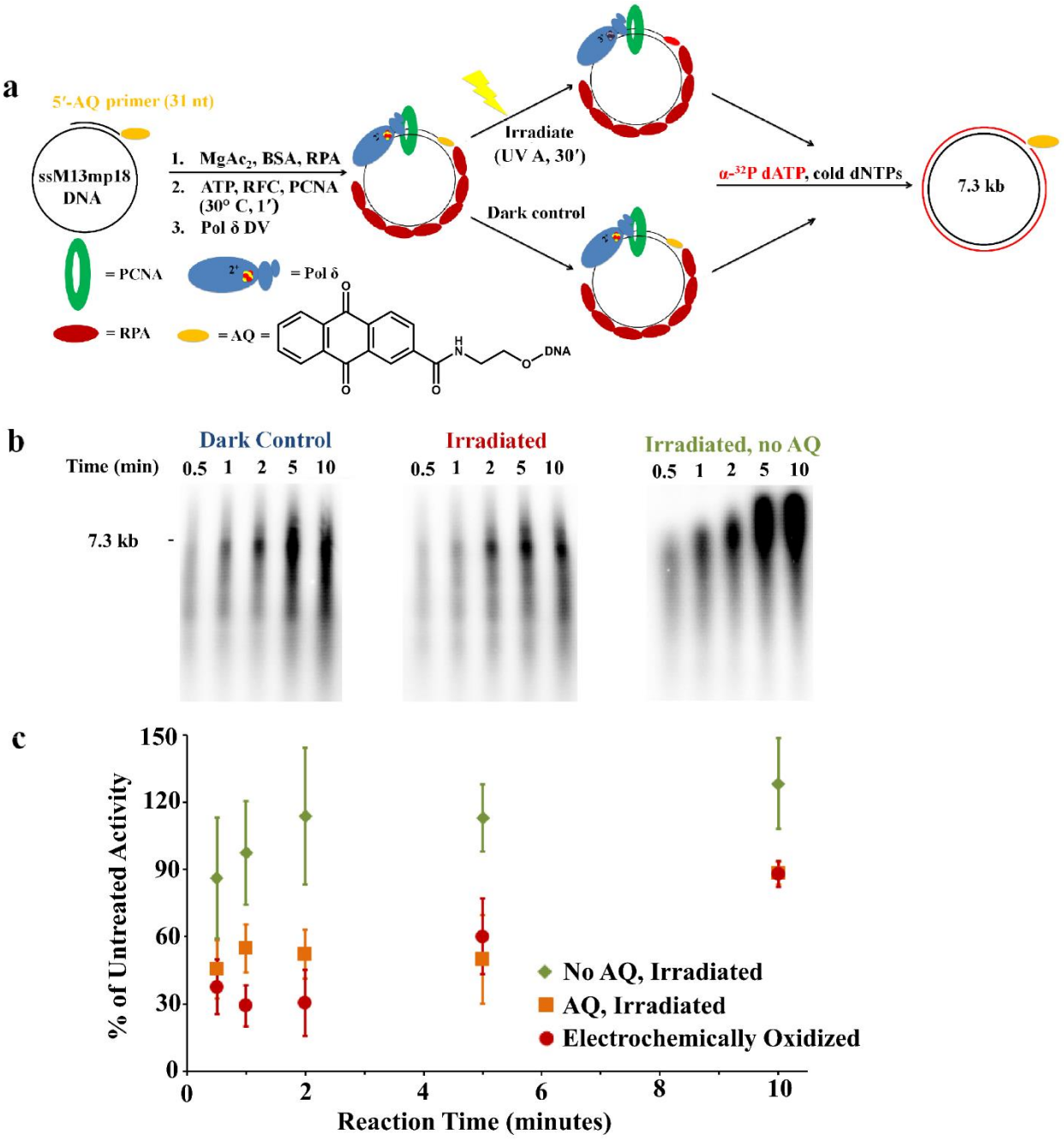


Figure 5.10 Pol δ activity assays with an AQ photooxidant. (**a**) Pol δ DV was added to a reaction mix lacking dNTPs and either irradiated under a solar simulator (UVA) or left in the dark, after which dNTPs (including [α-^{32}P] dATP) were added to start the reaction. (**b, c**) 1% alkaline agarose gels show less DNA synthesis by irradiated Pol δ at early time points, matching the pattern of electrochemical oxidation. Irradiation in the absence of AQ resulted in no significant effects, indicating that AQ was oxidizing the cluster. This pattern is apparent in relative radioactivity counts (or gel quantification for electrochemical oxidation), which further emphasize the similarity between oxidation methods. Error bars are standard deviation of the mean ($n \geq 3$).

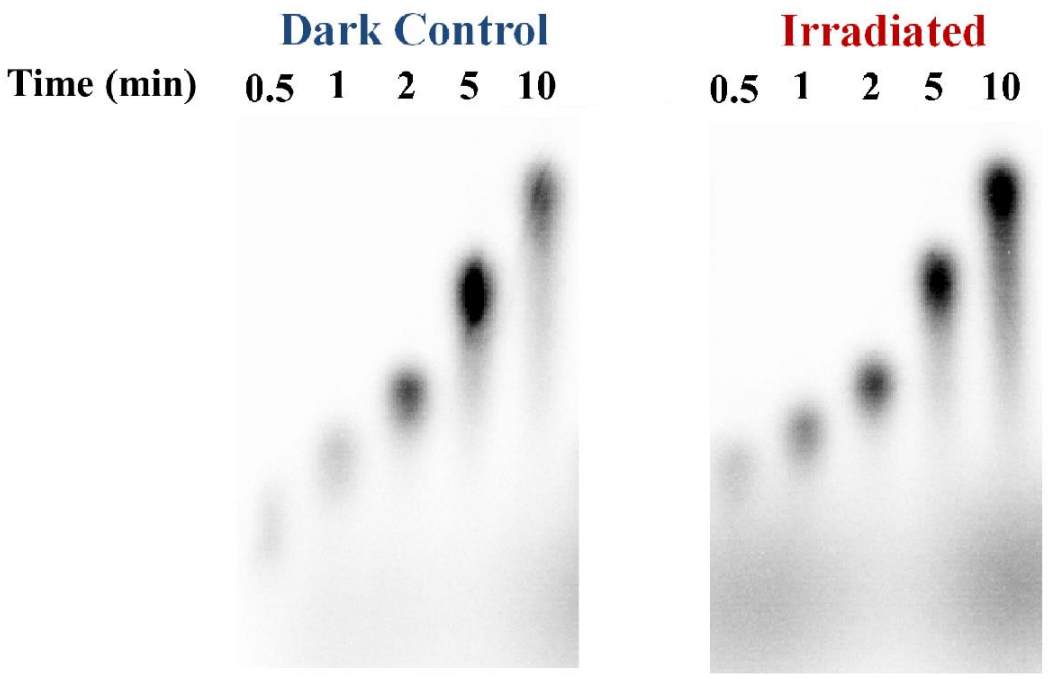


Figure 5.11 AQ assay controls with 140 nM *E. coli* Klenow fragment exo$^-$. UVA irradiation in the presence of AQ-primed DNA had no significant effect on DNA synthesis by Klenow fragment. The lack of difference confirms that irradiation in the presence of AQ does not adversely affect polymerase enzymes, and further supports the assignment of attenuated activity in Pol δ under the same conditions to [4Fe4S] cluster oxidation.

Discussion

In this work, we have demonstrated that the [4Fe4S] cluster of DNA polymerase δ can serve an important functional role as a redox-active cofactor that regulates enzymatic activity. On DNA-modified gold electrodes, Pol δ shows a reversible signal with a midpoint potential of 113 ± 5 mV versus NHE. Notably, this potential is comparable to that of the previously characterized [4Fe4S] protein EndoIII under the same buffer conditions. Charge attenuation in the presence of either an abasic site or a CA mismatch confirms that the redox signal is DNA-mediated. Activity assays carried out with electrochemically oxidized Pol δ demonstrate that oxidation results in significant slowing of processive DNA synthesis; the same result occurs following irradiation in the presence of an anthraquinone photooxidant. An assessment of both large and small DNA products indicates that the oxidized form remains active, but it is less processive. Given the retention of activity with decreased processivity, these results are consistent with an increase in DNA binding affinity upon oxidation, which would impede rapid sliding of PCNA-bound Pol δ. A significant increase in DNA binding is evident also with EndoIII upon oxidation (*27*). Critically, reduction by bulk electrolysis largely restores activity to native levels, confirming that cluster oxidation acts as a reversible switch. The reversibility of oxidation also lends further support to our electrochemical and spectroscopic evidence for the $[4Fe4S]^{3+}$ cluster as the biologically relevant oxidation product rather than degraded $[3Fe4S]^{+}$ cluster. Taken as a whole, these results suggest that reversible oxidation of the [4Fe4S] cluster in Pol δ could provide a rapid and reversible way to respond to replication stress.

Replication stress is a general term for fork slowing or stalling due to factors such as dNTP depletion, UV irradiation, and oxidative stress (*54*). Both replicative helicases and polymerases are known to be stabilized at stalled forks independently of checkpoint kinase

activity, although the mechanism of stabilization and its relationship to global checkpoint regulation remains incompletely understood (*55*). Pol δ slowing through cluster oxidation would be a straightforward way to stall replication on the lagging strand, and this form of slowing could also complement more standard regulatory mechanisms. In general, Pol δ slowing would lead to accumulation of single-stranded DNA and RPA on the lagging strand, activating checkpoint kinases that could then stall the helicase by phosphorylation of key subunits (*56, 57*). Among other types of stress, oxidative stress and ionizing radiation carry a heightened risk of double strand break formation (*58*). Furthermore, reactive oxygen species generate lesions such as 8-oxoguanine (OxoG), which is highly mutagenic due to the propensity of replicative polymerases to generate OxoG:A mispairs (*59*). Oxidation of the [4Fe4S] cluster from the 2+ to 3+ form could either occur directly, by reactive oxygen species, or, more likely, by charge transport (CT) communication through double stranded DNA with other oxidized species. Previously, we have seen that guanine radicals, the precursors to OxoG, can carry out DNA CT to oxidize the [4Fe4S] cluster of DNA-bound EndoIII (*24*). An additional benefit of polymerase stalling by DNA-mediated oxidation would be the prevention of excessive DNA synthesis under high risk circumstances. In addition to preventing damage, redox signaling to and from Pol δ could play a role in fork reversal and recombination events associated with replication stress (*54, 60*).

A general model for redox regulation of Pol δ is shown in Figure 5.12. In this model, processive lagging strand replication proceeds until replications stress occurs and the Pol δ $[4Fe4S]^{2+}$ cluster is oxidized either directly or by an electron acceptor, possibly an oxidized [4Fe4S] protein or a guanine radical formed during oxidative stress. We have found, in the case of EndoIII, that cluster oxidation promotes a substantial increase in binding affinity (*27*). Here, given the already tight binding to DNA of Pol δ with PCNA, a still tighter binding causes Pol δ

to slow its progression. The slowing of lagging strand synthesis would lead to RPA accumulation and an activation of checkpoint signaling, ultimately resulting in repair or replication fork collapse. Once the conditions of stress resolve, lagging strand synthesis can be restored by reduction of the Pol δ $[4Fe4S]^{3+}$ cluster, likely by another [4Fe4S] protein involved in DNA processing. Importantly, only Pol δ bound to the DNA in a complex with PCNA is readily oxidized, leaving the bulk of unbound Pol δ in the reduced 2+ form. Furthermore, this signaling can occur rapidly and at distance through DNA-mediated CT.

It is interesting to consider the many possible partners for DNA CT with Pol δ in the context of this model. First, Pol δ could be directly oxidized by species formed during oxidative stress, such as the guanine radical cation (*24*). After removal of reactive oxygen species, a partner [4Fe4S] protein could then re-reduce Pol δ and restore lagging strand replication. Alternatively, the entire redox cycle could be carried out between Pol δ and other [4Fe4S] proteins associated with the replication fork. The nuclease-helicase Dna2 is a prime example of such a partner, as it associates with the replication fork and is involved in both fork reversal and double strand break repair (*61, 62*). In human cells, the [4Fe4S] DNA repair glycosylase MUTYH associates with PCNA during S-phase; whether this association also occurs in the yeast homologue, Ntg2, is not known (*63*). If the glycosylases do generally associate with the fork, redox signaling could enable rapid communication between replication and repair pathways in eukaryotes. Finally, redox signaling between Pol δ and the B-family translesion DNA synthesis polymerase Pol ζ could help these proteins to hand off DNA containing bulky lesions. Indeed, a role for the cluster in the Pol δ – Pol ζ switch has already been suggested (*64*). In this view, Pol δ and Pol ζ switch by exchanging their shared B-subunits, although this would leave the cluster vulnerable to degradation (*65*). However, a redox handoff similar to that suggested between

primase and Pol α could allow such a transfer without requiring direct subunit exchange (*16*). While redox control of Pol δ would provide clear opportunities and is intriguing to consider, the *in vivo* mechanism and possible partners still require further investigation.

In summary, we have shown that Pol δ can use its [4Fe4S] cluster for reversible electron transfer along DNA and that oxidation of the cluster leads to reversible stalling. Taken together, our data suggest a model in which Pol δ uses redox signaling through DNA to sense oxidative stress, stalling replication under the mutagenic conditions, and potentially coordinating activities with repair and other replication proteins. Overall, the redox sensitivity of Pol δ reveals the oxidation state of the [4Fe4S] cluster to be a critical redox switch, and could provide a means to rapidly and reversibly to respond to replication stress.

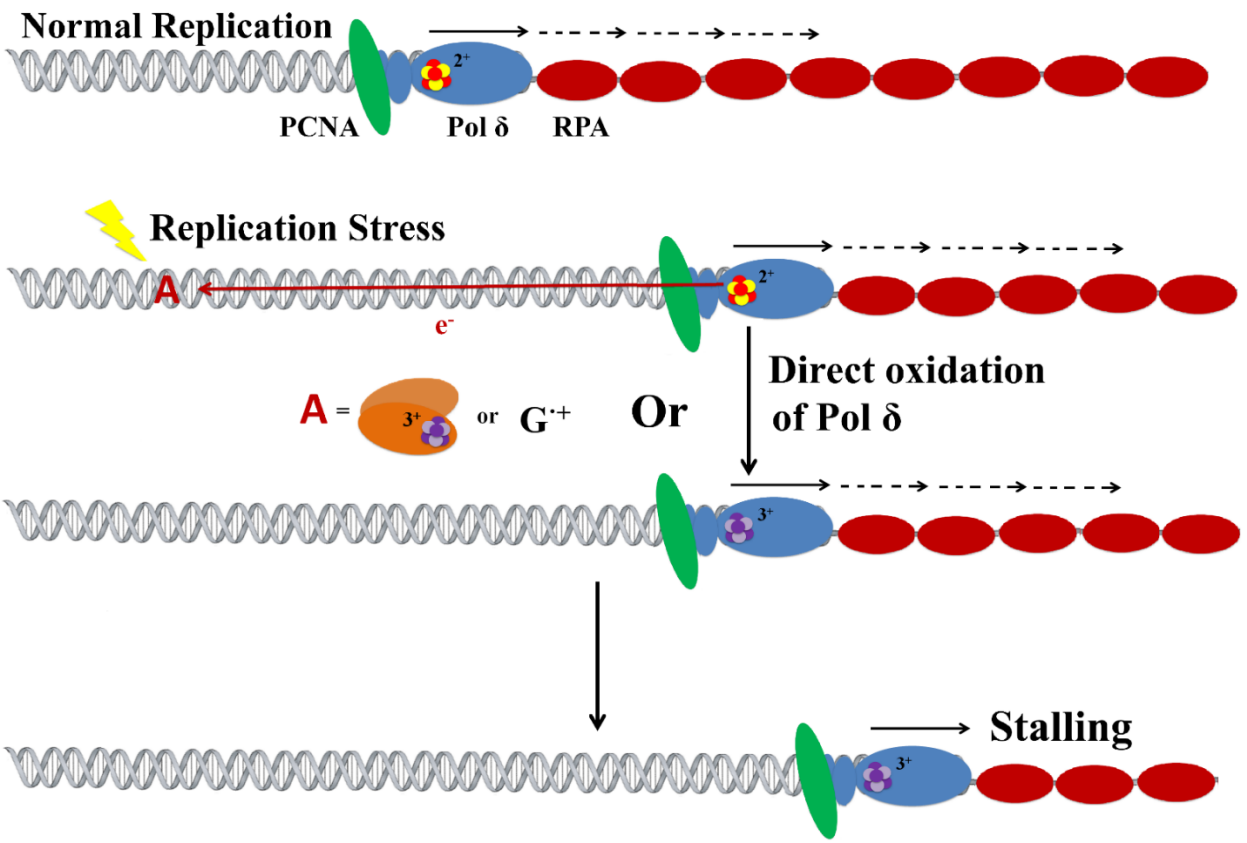


Figure 5.12 A model for redox-mediated regulation of Pol δ activity. Under ordinary circumstances, Pol δ forms a complex with PCNA and processively extends lagging strand DNA (top). When replication stress occurs, lagging strand synthesis could be stalled either by transfer of an electron from the Pol δ $[4Fe4S]^{2+}$ cluster to an acceptor (**A**) or by direct oxidation of the DNA-bound Pol δ-PCNA complex. In the case of DNA-mediated oxidation, the electron acceptor could be an oxidized [4Fe4S] protein or a guanine radical cation, $\mathbf{G}^{\cdot+}$, formed during oxidative stress. Once the cluster is oxidized to the $[4Fe4S]^{3+}$ form, Pol δ binds more tightly to DNA and synthesis slows. After damage resolution, lagging strand replication could be restored by reduction of the $[4Fe4S]^{3+}$ cluster by another [4Fe4S] proteins.

References

1) Lujan, S. A.; Williams, J. S.; Kunkel, T. A. DNA Polymerases Divide the Labor of Genome Replication. *Trends Cell Biol.* **2016**, *26*, 640 – 654.

2) Burgers, P.M.J.; Kunkel, T.A. Eukaryotic DNA Replication Fork. *Annu. Rev. Biochem.* **2017**, *86*, 417 – 438.

3) Johnson, R. E.; Klassen, R.; Prakash, L.; Prakash, S. A Major Role of DNA Polymerase δ in Replication of Both the Leading and Lagging DNA Strands. *Mol. Cell* **2015**, *59*, 163 – 175.

4) Yeeles, J. T. P.; Janska, A.; Early, A.; Diffley, J. F. X. How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. *Mol. Cell* **2017**, *65*, 105 – 116.

5) Prindle, M.J.; Loeb, L.A. DNA Polymerase Delta in DNA Replication and Genome Maintenance. *Environ. Mol. Mutagen.* **2012**, *53*, 666 – 682.

6) Johansson, E.; MacNeill, S. A. The eukaryotic replicative DNA polymerases take shape. *Trends Biochem. Sci.* **2010**, *35*, 339 – 347.

7) Netz, D. J. A.; Stith, C. M.; Stümpfig, M.; Köpf, G.; Vogel, D.; Genau, H. M.; Stodola, J. L.; Lill, R.; Burgers, P. M. J.; Pierik, A. J. Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat. Chem. Biol.* **2012**, *8*, 125 – 132.

8) Klinge, S.; Núñez-Ramírez, R.; Llorca, O.; Pellegrini, L. 3D architecture of DNA Pol α reveals the functional core of multi-subunit replicative polymerases. *EMBO J.* **2009**, *28*, 1978 – 1987.

9) Netz, D. J. A.; Mascarenhas, J.; Stehling, O.; Pierik, A. J.; Lill, R. Maturation of cytosolic and nuclear iron-sulfur proteins. *Trends Cell Biol.* **2014**, *24*, 303 – 312.

10) Jain, R.; Vanamee, E. S.; Dzikovski, B. G.; Buku, A.; Johnson, R. E.; Prakash, L.; Prakash, S.; Aggarwal, A. K. An Iron-Sulfur Cluster in the Polymerase Domain of Yeast DNA Polymerase ε. *J. Mol. Biol.* **2014**, *426*, 301 – 308.

11) Johnson, D.C.; Dean, D.R.; Smith, A.D.; Johnson, M.K. Structure, Function, and Formation of Biological Iron-Sulfur Clusters. *Annu. Rev. Biochem*. **2005**, *74*, 247 – 281.

12) Swan, M. K.; Johnson, R. E.; Prakash, L.; Prakash, S.; Aggarwal, A. K. Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. *Nat. Struct. Mol. Biol.* **2009**, *16*, 979 – 986.

13) Boal, A. K.; Yavin, E.; Lukianova, O. A.; O'Shea, V. L.; David, S. S.; Barton, J. K. DNA-Bound Redox Activity of DNA Repair Glycosylases Containing [4Fe4S] Clusters. *Biochemistry* **2005**, *44*, 8397 – 8407.

14) Mui, T. P.; Fuss, J. O.; Ishida, J. P.; Tainer, J. A.; Barton, J. K. ATP-Stimulated, DNA-Mediated Redox Signaling by XPD, a DNA Repair and Transcription Helicase. *J. Am. Chem. Soc.* **2011**, *133*, 16378 – 16381.

15) Grodick, M. A.; Segal, H. M.; Zwang, T. J.; Barton, J. K. DNA-Mediated Signaling by Proteins with 4Fe-4S Clusters is Necessary for Genomic Integrity. *J. Am. Chem. Soc.* **2014**, *136*, 6470 – 6478.

16) O'Brien, E.; Holt, M. E.; Thompson, M. K.; Salay, L. E.; Ehlinger, A. C.; Chazin, W. J.; Barton, J. K. The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport. *Science* **2017**, *355*, 813.

17) Porello, S. L.; Cannon, M. J.; David, S. S. A Substrate Recognition Role for the [4Fe-4S]$^{2+}$ Cluster of the DNA Repair Glycosylase MutY. *Biochemistry* **1998**, *37*, 6465 – 6475.

18) Fan, L.; Fuss, J. O.; Cheng, Q. J.; Arvai, A. S.; Hammel, M.; Roberts, V. A.; Cooper, P. K.; Tainer, J. A. XPD Helicase Structures and Activities: Insights into the Cancer and Aging Phenotypes from XPD Mutations. *Cell* **2008**, *133*, 789 – 800.

19) Ren, B.; Duan, X.; Ding, H. Redox control of the DNA damage-inducible protein DinG helicase activity via its iron-sulfur cluster. *J. Biol. Chem.* **2009**, *284*, 4829 – 4835.

20) Weiner, B. E.; Huang, H.; Dattilo, B. M.; Nilges, M. J.; Fanning, E.; Chazin, W. J. An Iron-Sulfur Cluster in the C-terminal Domain of the p58 Subunit of Human DNA Primase. *J. Biol. Chem.* **2007**, *282*, 33444 – 33451.

21) Gorodetsky, A.A.; Boal, A.K.; Barton, J.K. Direct Electrochemistry of Endonuclease III in the Presence and Absence of DNA. *J. Am. Chem. Soc.* **2006**, *128*, 12082 – 12083.

22) Bartels, P.L.; Zhou, A.; Arnold, A.R.; Nuñez, N.N.; Crespilho, F.N.; David, S.S.; Barton, J.K. Electrochemistry of the [4Fe4S] Cluster in Base Excision Repair Proteins: Tuning the Redox Potential with DNA. *Langmuir* **2017**, *33*, 2523 – 2530.

23) Ha, Y.; Arnold, A.R.; Nuñez, N.N.; Bartels, P.L.; Zhou, A.; David, S.S.; Barton, J.K.; Hedman, B.; Hodgson, K.O.; Solomon, E.I. Sulfur K-Edge XAS Studies of the Effect of DNA Binding on the [Fe_4S_4] Site in EndoIII and MutY. *J. Am. Chem. Soc.* **2017**, *139*, 11434 – 11442.

24) Yavin, E.; Boal, A. K.; Stemp, E. D. A.; Boon, E. M.; Livingston, A. L.; O'Shea, V. L.; David, S. S.; Barton, J. K. Protein-DNA Charge Transport: Redox Activation of a DNA Repair Protein by Guanine Radical. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 3546 – 3551.

25) Dey, A.; Jenney, F. E.; Adams, M. W. W.; Babini, E.; Takahashi, Y.; Fukuyama, K.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. Solvent Tuning of Electrochemical Potentials in the Active Sites of HiPIP vs Ferredoxin. *Science* **2007**, *318*, 1464 – 1468.

26) Arnold, A. R.; Grodick, M. A.; Barton, J. K. DNA Charge Transport: from Chemical Principles to the Cell. *Cell Chem. Biol*. **2016**, *23*, 183 – 197.

27) Tse, E.C.M.; Zwang, T.J.; Barton, J.K. The Oxidation State of [4Fe4S] Clusters Modulates the DNA-Binding Affinity of DNA Repair Proteins. *J. Am. Chem. Soc.* **2017**, *139*, 12784 – 12792.

28) Pheeney, C. G.; Arnold, A. R.; Grodick, M. A.; Barton, J. K. Multiplexed Electrochemistry of DNA-Bound Metalloproteins. *J. Am. Chem. Soc.* **2013**, *135*, 11869 – 11878.

29) Slinker, J. D.; Muren, N. B.; Gorodetsky, A.A.; Barton, J. K. Multiplexed DNA-Modified Electrodes. *J. Amer. Chem. Soc.* **2010**, *132*, 2769 – 2774.

30) Barton, J.K.; Bartels, P.L.; Deng, Y.; O'Brien, E. Electrical Probes of DNA-Binding Proteins. *Methods Enzymol.* **2017**, *591*, 355 – 414.

31) Greeley, R. S.; Smith, W. T.; Stoughton, R. W.; Lietzke, M. H. Electromotive Forces in Aqueous Solutions at Elevated Temperatures. I. The Standard Potential of the Silver-Silver Chloride Electrode. *J. Phys. Chem.* **1960**, *64*, 652 – 657.

32) Burgers, P. M. J.; Yoder, B. L. ATP-Independent Loading of the Proliferating Cell Nuclear Antigen Requres DNA Ends. *J. Biol. Chem.* **1993**, *268*, 19923 – 19926.

33) Garg, P.; Stith, C. M.; Sabouri, N.; Johansson, E.; Burgers, P. M. Idling by DNA polymerase δ maintains a ligatable nick during lagging-strand DNA replication. *Genes Dev.* **2004**, *18*, 2764 – 2773.

34) Kissinger, P. T.; Heineman, W. R. Large-Amplitude Controlled-Potential Techniques in *Laboratory Techniques in Electroanalytical Chemistry*, *2nd ed.* (Kissinger, P.T. and Heineman, W.R., eds.), **1996**, pp. 51 – 139, Marcel Dekker, New York.

35) Stodola, J.L.; Burgers, P.M. Resolving individual steps of Okazaki-fragment maturation at a millisecond time scale. *Nat. Struct. Mol. Biol.* **2016**, *23*, 402 – 408.

36) Schaefer, K. N.; Barton, J. K. DNA-Mediated Oxidation of p53. *Biochemistry* **2014**, *53*, 3467 – 3475.

37) Johnson, R. E.; Prakash, L.; Prakash, S. Yeast and Human Translesion DNA Synthesis Polymerases: Expression, Purification, and Biochemical Characterization. *Methods Enzymol*. **2006**, *408*, 390 – 407.

38) Tomoaki, O.; Nakamura, H., Katsumi, N.; Cusanovich, M.A.; Akutsu, H. Ionic Strength-Dependent Physicochemical Factors in Cytochrome c_3 Regulating the Electron Transfer Rate. *Biophys. J.* **1998**, *75*, 1483 – 1490.

39) Lode, E.T.; Murray, C.L.; Rabinowitz, J.C. Apparent oxidation-reduction potential of *Clostridium acidiurici* ferredoxin. Effect of pH, ionic strength, and amino acid replacements. *J. Biol. Chem.* **1976**, *251*, 1683 – 1687.

40) Johansson, E.; Majka, J.; Burgers, P.M. Structure of DNA Polymerase δ from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **2001**, *276*, 43824 – 43828.

41) Boon, E. M.; Salas, J. E.; Barton, J. K. An Electrical Probe of Protein-DNA Interactions On DNA-Modified Surfaces. *Nat. Biotechnol.* **2002**, *20*, 282 – 286.

42) Kelley, S. O.; Barton, J. K.; Jackson, N. M.; Hill, M. G. Electrochemistry of Methylene Blue bound to a DNA-Modified Electrode. *Bioconjug. Chem.* **1997**, *8*, 31 – 37.

43) Sweeney, W. V; Rabinowitz, J. C. Proteins Containing 4Fe-4S Clusters: An Overview. *Annu. Rev. Biochem.* **1980**, *49*, 139 – 161.

44) Johnson, M. K.; Duderstadt, R. E.; Duin, E. C. Biological and Synthetic $[Fe_3S_4]$ Clusters. *Adv. Inorg. Chem.* **1999**, *47*, 1 – 82.

45) Duff, J. L. C.; Breton, J. L. J.; Butt, J. N.; Armstrong, F. A.; Thomson, A. J. Novel Redox Chemistry of [3Fe-4S] Clusters: Electrochemical Characterization of the All-Fe(II) Form of the [3Fe-4S] Cluster Generated Reversibly in Various Proteins and Its Spectroscopic Investigation in *Sulfolobus acidocaldarius* Ferredoxin. *J. Am. Chem. Soc.* **1996**, *118*, 8593 – 8603.

46) Hedglin, M.; Pandey, B.; Benkovic, S. J. Stability of the human polymerase δ holoenzyme and its implications in lagging strand DNA synthesis. *Proc. Natl. Acad. Sci.* **2016**, *13*, E1777 – 1786.

47) Kulak, N.A.; Pichler, G.; Paron, I.; Nagaraj, N.; Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in yeast. *Nat. Methods* **2014**, *11*, 319 – 324.

48) Goffeau, A.; Barrell, B.G.; Bussey, H.; Davis, R.W.; Dujon, B.; Feldmann, H.; Galibert, F.; Hoheisel, J.D.; Jacq, C.; Johnston, M.; Louis, E.J.; Mewes, H.W.; Murakami, Y.; Philippsen, P.; Tettelin, H.; Oliver, S.G. A Gene Map of the Human Genome. *Science* **1996**, *274*, 546 – 567.

49) Lengronne, A.; Pasero, P.; Bensimon, A.; Schwob, E. Monitoring S phase progression globally and locally using BrdU incorporation in TK^+ yeast strains. *Nucleic Acids Res.* **2001**, *29*, 1433 – 1422.

50) Henderson, P.T.; Jones, D.; Hampikian, G.; Kan, Y.; Schuster, G.B. Long-distance charge transport in duplex DNA: The phonon-assisted polaron-like hopping mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8353 – 8358.

51) Armitage, B.; Yu, C.; Devadoss, C.; Schuster, G.B. Cationic Anthraquinone Derivatives as Catalytic DNA Photonucleases: Mechanisms for DNA Damage and Quinone Recycling. *J. Am. Chem. Soc.* **1994**, *116*, 9847 – 9859.

52) Gasper, S.M.; Schuster, G.B. Intramolecular Photoinduced Electron Transfer to Anthraquinones Linked to Duplex DNA: The Effect of Gaps and Traps on Long-Range Radical Cation Migration. *J. Am. Chem. Soc.* **1997**, *119*, 12762 – 12771.

53) Nuñez, M.E.; Hall, D.B.; Barton, J.K. Oxidative Charge Transport through DNA in Nucleosome Particles. *Chem. Biol.* **1999**, *6*, 85 – 97.

54) Berti, M.; Vindigni, A. Replication stress: getting back on track. *Nat. Struct. Mol. Biol.* **2016**, *23*, 103 – 109.

55) De Piccoli, G.; Katou, Y.; Itoh, T.; Nakato, R.; Shirahige, K.; Labib, K. Replisome stability at defective DNA replication forks is independent of S-phase checkpoint kinases. *Mol. Cell.* **2012**, *45*, 696 – 704.

56) Zeman, M.K.; Cimprich, K.A. Causes and consequences of replications stress. *Nat. Cell Biol.* **2014**, *16*, 2 – 9.

57) Saldivar, J.C.; Cortez, D.; Cimprich, K.A. The essential kinase ATR: ensuring faithful duplication of a challenging genome. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 622 – 636.

58) Azzam, E.I.; Jay-Gerin, J.P.; Pain, D. Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. *Cancer Lett.* **2012**, *327*, 48 – 60.

59) McCulloch, S.D.; Kokoska, R.J.; Garg, P.; Burgers, P.M.; Kunkel, T.A. The efficiency and fidelity of 8-oxo-guanine bypass by DNA polymerases δ and η. *Nucleic Acids Res.* **2009**, *37*, 2830 – 2840.

60) Zellweger, R.; Dalcher, D.; Mutreja, K.; Berti, M.; Schmid, J. A.; Herrador, R.; Vindigni, A.; Lopes, M. Rad51-Mediated Fork Reversal is a Global Response to Genotoxic Treatments in Human Cells. *J. Cell Biol.* **2015**, *208*, 563 – 579.

61) Thangavel, S.; Berti, M.; Levikova, M.; Pinto, C.; Gomathinayagam, S.; Vujanovic, M.; Zellweger, R.; Moore, H.; Lee, E. H.; Hendrickson, E. A.; Cejka, P.; Stewart, S.; Lopes, M.; Vindigni, A. DNA2 Drives Processing and Restart of Reversed Replication Forks in Human Cells. *J. Cell Biol.* **2015**, *208*, 545 – 562.

62) Nimonkar, A. V.; Genschel, J.; Kinoshita, E.; Polaczek, P.; Campbell, J. L.; Wyman, C.; Mdrich, P.; Kowalczykowski, S. C. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-

MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev.* **2011**, *25*, 350 – 362.

63) Markkanen, E.; Dorn, J.; Hübscher, U. MUTYH DNA glycosylase: the rationale for removing undamaged bases from the DNA. *Front. Genet.* **2013**, *4*, 18.

64) Baranovskiy, A. G.; Lada, A. G.; Siebler, H. M.; Zhang, Y.; Pavlov, Y. I.; Tahirov, T. H. DNA polymerase δ and ζ switch by sharing accessory subunits of DNA polymerase δ. *J. Biol. Chem.* **2012**, *287*, 17281 – 17287.

65) Johnson, R. E.; Prakash, L.; Prakash, S. Pol31 and Pol32 subunits of yeast DNA polymerase δ are also essential subunits of DNA polymerase ζ. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 12455 – 12460.

Chapter 6

Evaluating the Charge Transport Pathway Between DNA and the [4Fe4S] Cluster of Yeast DNA Polymerase δ

This work was done as part of a collaboration with Joseph Stodola, Carrie Stith, and Prof. Peter Burgers at the Washington University School of Medicine in St. Louis, Missouri.

P. Bartels identified possible mutant residues and performed all electrochemistry-based experiments. J. Stodola and C. Stith working with P. Burgers prepared all proteins, carried out basic characterization, and performed all *in vivo* experiments.

Introduction

A [4Fe4S] cluster in the C-terminal domain of yeast DNA polymerase δ is essential for the formation of an active multi-subunit complex and has recently been shown to be redox active in the presence of DNA (*1, 2*). Notably, electron transfer in Pol δ is thought to be primarily DNA-mediated, with electrons moving toward or away from the cluster through the π-stacked bases of the DNA in a process known as DNA-mediated charge transport (DNA CT) (*2, 3*). However, studies performed on a wide variety of other DNA-binding [4Fe4S] proteins that participate in DNA CT suggest that it is unlikely that electron transfer takes place directly between the cluster and the DNA (*4 -10*). Unfortunately, no structural data is available for this region in Pol δ, but these other cases may provide some useful general insights. Among these proteins, complete or partial structures are known for bacterial endonuclease III (EndoIII), MutY, archaeal XPD, and human DNA primase, and DNA-bound structures are also available for EndoIII, MutY, and mouse Dna2 (*4-10*). In all cases, the [4Fe4S] cluster is buried and thus shielded from solvent, and, for those with DNA-bound structures, the cluster is typically positioned at a distance of 15 Å or more from the DNA duplex (*8-10*). Aromatic residues have been shown to be critical for CT in EndoIII, and detailed mutagenesis studies in DNA primase reveal a string of tyrosine residue bridging the cluster and DNA (*11, 12*). Aromatic residues, specifically tryptophan and tyrosine, are well-known to act as bridges for electron hopping within proteins, and their use for this purpose in DNA-binding [4Fe4S] proteins is thus unsurprising (*13*).

In redox-active DNA-binding proteins, indirect electron transfer mediated by aromatic amino acid residues serves two main purposes. First, keeping the cluster away from the DNA may minimize the risk of DNA damaging Fenton chemistry if the cluster is degraded to release

free iron (*3*). Second, in order to communicate with one another in reversible long-range CT, all of these proteins must have similar potentials (*3*). Solvent exposure and proximity to the polyanionic DNA backbone are thought to be the primary factors in tuning redox potential, and thus these factors must be roughly equivalent between these proteins (*14, 15*). Pol δ has a measured redox potential that is indistinguishable from the well-characterized EndoIII in the same buffer conditions and is similarly unresponsive to in-solution oxidants and reductants, making it likely that the cluster environment is like that of other [4Fe4S] proteins that bind DNA (*1, 2*).

Previous work has shown that the oxidation state of the Pol δ [4Fe4S] cluster is important in regulating activity, with oxidation to the 3+ state markedly slowing DNA synthesis in a reversible manner (*2*). Such a redox switch would be very useful in responding to replication stress *in vivo,* especially in cases of oxidative stress. While *in vitro* work has been very useful in analyzing the redox capabilities of the cluster and its effect on enzymatic activity, addressing the actual function requires CT-deficient mutants that can be used in *in vivo* experiments. In these studies, we aimed to identify, prepare, and characterize several such mutants in Pol δ using *in vitro* techniques before proceeding to assess the sensitivity of yeast cells harboring these mutations to various stressors.

Materials and Methods

Protein Expression and Purification

Pol δ DV (D520V) and the mutants W1053A and Y1078A were prepared using a previously published yeast overexpression system (*16*). PCNA was prepared by overexpression in yeast, while RFC and RPA were overexpressed in *E. coli* following previous protocols (*17-19*).

DNA Preparation

Thiol-modified DNA was prepared on an Applied Biosystems automated DNA synthesizer using a 3′ thiol modifier phosphoramidite and standard reagents from Glen Research. Unmodified DNA was purchased from IDT. All DNA sequences were purified by HPLC as described previously (*2*), and annealed in 1:1 molar ratios for electrochemistry. The sequences used were as follows:

20-mer 3′ thiol: 5′ - GCT GTC GTA CAG CTC AAT GC - 3′ - $(CH_2)_2O(CH_2)_3SH$

38-mer: 5′ - TAA CAG GTT GAT GCA TCG CGC TTC GGT GCT GCG TGT CT - 3′

49-mer: 5′ - GCA TT**G** AGC TGT ACG ACA GCA GAC ACG CAG CAC CGA AGC GCG ATG CAT C - 3′

DNA replication assays utilized the single-stranded plasmid M13mp18 DNA annealed to a 31-mer primer in a 1:1 ratio (5′ at 90 °C followed by slow cooling to RT). The primer was complementary to positions M13mp18 positions 6265-6234, and had the following sequence:

5′ - GAC TCT AGA GGA TCC CCG GGT ACC GAG CTC G - 3′ (1)

For activity assays, a 5′ ^{32}P label was appended to the primer before annealing using T4 polynucleotide kinase (PNK) in PNK buffer with [γ-^{32}P] ATP. Product sizes were determined using HincII-digested M13mp18 RFII DNA and 2 log DNA ladder, radiolabeled with T4 PNK following dephosphorylation by calf intestinal alkaline phosphatase (CIAP; 60 minutes, 37 °C).

DNA and enzymes were purchased from New England Biolabs (NEB), and radiolabeled ATP was from Perkin Elmer.

Electrochemical Characterization on DNA-Modified Gold Electrodes

Electrochemical characterization was performed on multiplexed chips with 0.015 cm^2 electrode area (*20*). Loosely-packed DNA self-assembled monolayers (SAMs) were formed by incubating 25 μL of 25 μM duplexed DNA on the electrode overnight, after which electrodes were rinsed 3-5 times in phosphate buffer (5 mM sodium phosphate, pH 7.5, 50 mM NaCl) and backfilled for 45 minutes with 1 mM 6-mercapto-1-hexanol in the same buffer containing 5% (v/v) glycerol. After backfilling, electrodes were extensively rinsed in phosphate buffer followed by protein storage buffer (30 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.5 mM EDTA, 10% glycerol, and 0.01% decaethylene glycol monododecyl ether), and the absence of electroactive impurities was confirmed by cyclic voltammetry (CV).

Electrochemical experiments used a standard 3-electrode cell consisting of a Au working electrode with a Ag/AgCl reference electrode in 3 M NaCl (BASInc), and a 1 mm diameter Pt wire as the counter electrode (Lesker). Potentials were converted from Ag/AgCl to NHE by adding 212 mV to the potential measured by Ag/AgCl (*2*). Because the [4Fe4S] absorbance was too low to determine concentration, Pol δ mutant concentrations in electrochemical experiments are reported as total protein concentration determined by Bradford assay. For unmodified Pol δ DV, [4Fe-4S] cluster concentration was determined by UV-visible spectroscopy using an extinction coefficient of 13000 $M^{-1}cm^{-1}$ at 400 nM, and a cluster loading of ~85% was determined by dividing cluster concentration by total protein concentration as determine by Bradford assay. For comparison with mutants, similar loading levels were assumed and equivalent amounts of loaded protein were added to electrodes. BSA and reagents were included

in kits purchased from Thermo Scientific, and UV-vis spectra were taken on a Cary Varian instrument using 100 μL quartz cuvettes (STARNA Cells).

Electrochemistry was performed using 500 nM Pol δ DV (unmodified or mutant) in combination with 5.0 μM PCNA, 80 μM dATP and 8 mM $MgAc_2$ in storage buffer (30 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.1 mM EDTA, 10% glycerol v/v, and 0.01% w/v decaethylene glycol monododecyl ether). Surfaces were scanned by CV (100 mV/s scan rate) and square wave voltammetry (SQWV; 15 Hz frequency and 25 mV amplitude) once per hour for several hours.

Activity Assays with Oxidized and Reduced Pol δ Mutants

Unless otherwise noted, bulk electrolysis was carried out in an anaerobic glove bag (95% N_2/5% H_2) using droplet electrochemistry (30-40 μL 190 nM sample diluted into degassed storage buffer) on Au rod electrodes of 0.0314 cm^2 electrode area (Pine Research Instrumentation). Oxidation was performed using a potential of 0.412 V vs NHE, while reduction was done at -0.188 V. Oxidation yields were estimated by taking the difference between the total charge obtained in the presence of Pol δ and that generated by electrolysis with buffer alone.

140 μL reaction mixes (0.1mg/mL BSA, 80 μM each dNTP, 500 μM ATP, 2.0 nM M13mp18 with a ^{32}P-labeled primer, 8.0 mM $MgAc_2$, 500 nM RPA, 5.0 nM RFC and 50 nM PCNA in 50 mM Tris-HCl with 50 mM NaCl, pH 7.8) were prepared inside the glove bag. All buffer components were purchased from Sigma-Aldrich.

PCNA was loaded onto primed ends by incubation with RFC (1′, 30 °C), and reactions were started by addition of Pol δ to 2 nM. Reactions were run at 30 °C, and 20 μL aliquots were removed at the specified time points and quenched with 10 μL stop mix (10 mM EDTA and 0.1% v/v SDS, final concentrations). Samples were counted on a liquid scintillation counter and

dried on a speed vacuum before being dissolved in alkaline gel buffer (500 mM NaOH, 10 mM EDTA) with 6x alkaline loading dye (300 mM NaOH, 6.0 mM EDTA, 18% Ficoll w/v, 0.25% xylene cyanol w/v, and 0.15% bromocresol green w/v) diluted to 1x in buffer. Equivalent amounts of radioactivity were loaded onto a 1% alkaline agarose gel and run at 30 V for 14-15 hours. Gels were neutralized in 7% TCA w/v in water (30 minutes, RT), dried, and imaged on a Typhoon phosphorimager. Gels were quantified using ImageQuant software (GE Healthcare). Because oxidation affected the mutants differently (described below), the total amount of DNA synthesis (identified as radioactivity above unextended primer) was compared.

***In Vivo* Assays**

Yeast strains containing either WT or mutant Pol δ were treated with exogenous agents to test the response to replication stress. Treatments included UV irradiation, camptothecin, and hydroxyurea, and treated cells were spot plated for an assessment of growth. To test sensitivity to oxidative stress, a mutant yeast strain defective in glutathione synthase (Δ*gsh*) was prepared, and cells were treated with either hydrogen peroxide (2 mM) or menadione (50 mM). Analysis was the same as for more generic treatments.

Results

Identification of Putative CT-Relevant Residues in Pol δ and Purification of Mutants

Sequence alignments of the catalytic subunits of Pol δ and Pol α (Pol3 and Pol1, respectively) were used as first step in identifying conserved aromatic residues that might be involved in the Pol δ CT pathway. Because no structure of Pol δ containing the C-terminal domain that harbors the cluster exists, we relied upon a structure of the Pol α CTD/B-subunit complex (PDB ID 3FLO) to estimate distances between the cluster and possible CT-relevant residues. To facilitate this analysis, we limited ourselves to residues conserved between both Pol δ and Pol α. It should be noted that Pol1 differs from Pol3 in the spacing of ligating cysteine residues, and the structure used contains zinc in place of a [4Fe4S] cluster (*21*), so some limitations to this strategy are present. If anything, these factors would lead to an overestimation of distances; nonetheless, given a lack of alternatives and the high degree of conservation in this region between the B-family polymerases, we considered the Pol1 structure sufficient for our purposes. Proceeding with this strategy, we identified two conserved aromatic residues within 15 Å of the Pol1 zinc (and thus likely closer in Pol3 with an intact [4Fe4S] cluster): W1053 and Y1078.

Notably, these residues are on opposite sides of the cluster, with W1053 facing the external environment and Y1078 more deeply buried (Figure 6.1). Thus, if these residues are involved in CT, they are likely to represent redundant pathways rather than a single, coherent path between DNA and the cluster; alternatively, one or both may be entirely uninvolved in CT. While the position of DNA relative to the cluster is unknown, the buried position of Y1078 makes this residue unlikely to be in direct contact with the DNA; however, it could still serve as the final residue in a string originating elsewhere. To maximize our chances of finding relevant

residues and to address this latter point, we also targeted Y1035, a residue in one of the alpha helices midway between the CysA zinc finger and the CysB [4Fe4S] cluster site (*1, 21*).

Once identified, constructs were designed with W or Y → A mutations in Pol δ DV, and all three mutants were overexpressed in yeast and purified according to standard protocols. Yields were lower than expected, although FPLC traces confirmed that intact 3-subunit complexes did form for both W1053A and Y1078A (Figure 6.2). Nonetheless, the proportion of degraded complexes was considerably higher than seen in unaltered Pol δ DV, and yields of intact complexes were too low to obtain useful quantities of Y1035A. The low yields and tendency of these mutants to degrade even when purified under anaerobic conditions suggested that the alanine substitution was likely perturbing structural integrity to some extent, although complexes could clearly still form. Supporting the fully intact nature of these collected fractions, the major peak also included absorbance at 410 nm, characteristic of [4Fe4S] clusters (Figure 6.2). Since we were still able to obtain intact complexes for W1053A and Y1078A at ~500 nM, we decided to proceed to assess their effects on activity and electrochemical properties in addition to their effect on cell viability under different stress conditions.

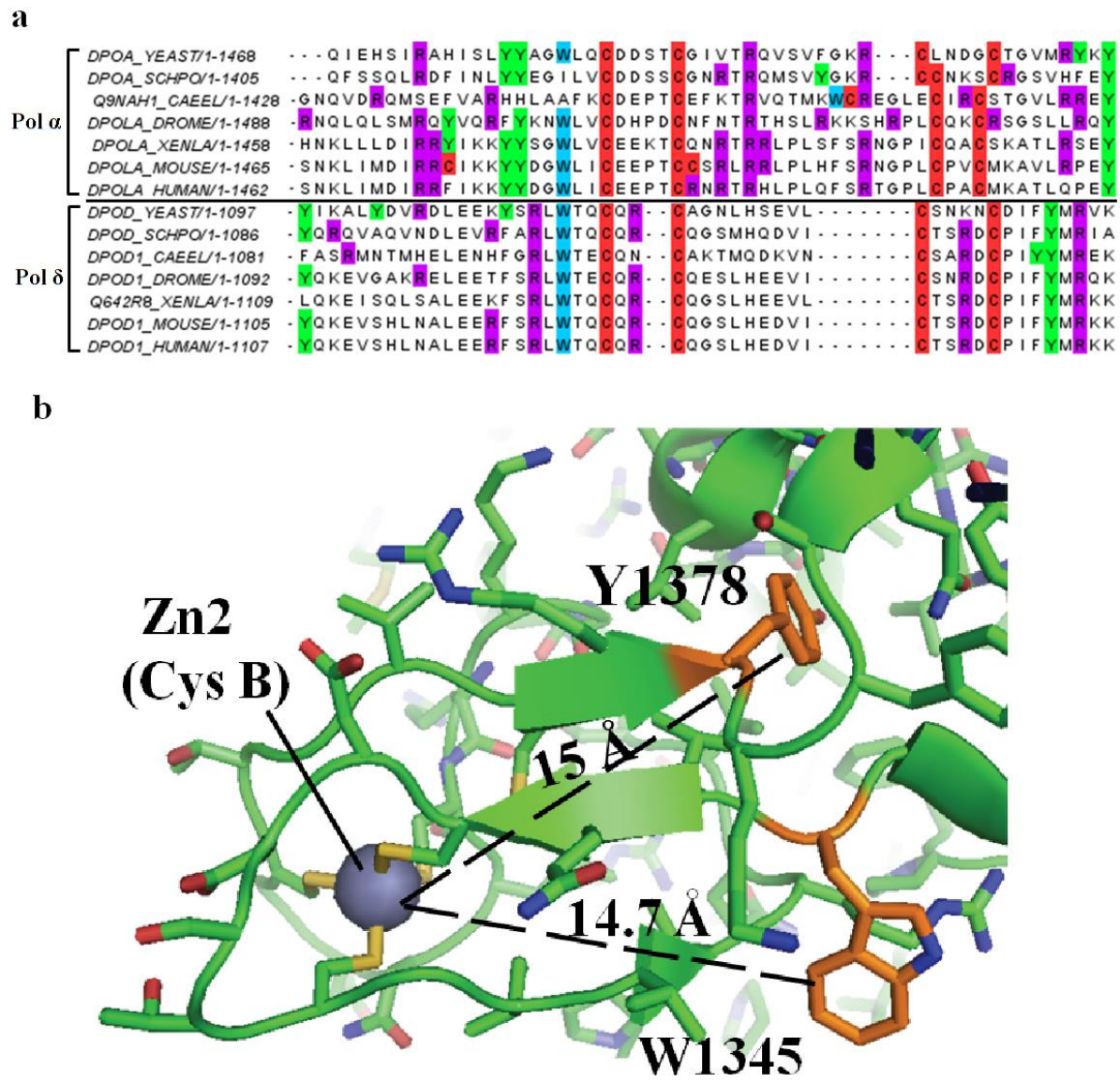


Figure 6.1 Identification of possible CT-relevant residues in Pol δ using sequence alignments and a Pol α C-terminal domain X-ray crystal structure (*21*). (a) Sequence alignments of Pol α and Pol δ from a variety of organisms. Cluster-ligating cysteines are highlighted in red, tyrosine in green, tryptophan in blue, and arginine in violet. Arginine is included to illustrate the proximity of a disease-relevant mutation in human Pol δ (ref. *27*). (b) Yeast Pol α C-terminal domain structure (PDB ID 3FLO) highlighting Y1378 and W1345, which correspond to Y1078 and W1053 in yeast Pol δ. Notably, both residues are within electron tunneling distance (~15 Å) from a zinc that occupies the CysB site in this structure, and these distances would be even shorter if the native [4Fe4S] cluster were present.

Characterization of Pol δ Mutants

Yeast strains containing Pol δ W1053A or Y1078A exhibited unperturbed growth, indicating that these mutants were able to form complexes effectively enough to support normal DNA replication. To see if replication stress inducers had a differential effect on these strains, cells were subjected to UV, camptothecin, and hydroxyurea treatment and spot plated (Figure 6.3; only hydroxyurea treatment is shown). Among these treatments, the Y1078A strain showed slight sensitivity to 75 mM hydroxyurea, although neither mutant was sensitive to lower concentrations or to any other stressor applied.

In an effort to accurately determine cluster loading levels, both mutants were concentrated to the extent possible (500 nM) for UV-visible spectroscopy (Figure 6.4). A clear absorption peak at 280 nm attributable to aromatic and thiolated amino acid residues was present, and the absence of elevated absorption at 800 nm indicated that aggregrates were not forming. At these concentrations, the low extinction coefficient of the cluster (13000 $M^{-1}cm^{-1}$ at 400 nm) makes the presence of a peak difficult to discern. To work around this problem, absorbance at 280 nm was normalized to 1 using an unmodified Pol δ DV spectrum; BSA, which should have no absorbance in the 400 nm region, was included as a control. When this analysis was carried out, both mutants show absorption elevated beyond BSA that is generally shaped like the peak in unmodified Pol δ DV, supporting similar cluster loading. However, with absolute absorbance so low, these results are not definitive. Although not quantitative, the best evidence for comparable cluster loading in the mutants remains the FPLC evidence for the formation of 3-subunit complexes with 410 nm absorption.

In vitro, both Pol δ DV mutants remained active in standard DNA replication assays on primed ssM13mp18 plasmid DNA (Figure 6.5). They did, however, exhibit lower processivity

than unaltered Pol δ DV, although native activity could be restored by adding a 10-fold molar excess of PCNA (Figure 6.5). Pol δ typically requires only stoichiometric amounts of PCNA to completely extend a primed ssM13mp18 plasmid, and these results indicate that the interaction of these mutants with PCNA is disrupted to some extent. PCNA binding involves the CysA zinc finger at the other side of the C-terminal domain from the [4Fe4S] cluster (*1*), so these results suggest that this entire region of the protein has undergone some reorganization upon these mutations. That excess PCNA can compensate demonstrates that PCNA binding remains possible, and suggests that, although structurally perturbed, the mutants remain in intact 3-subunit complexes.

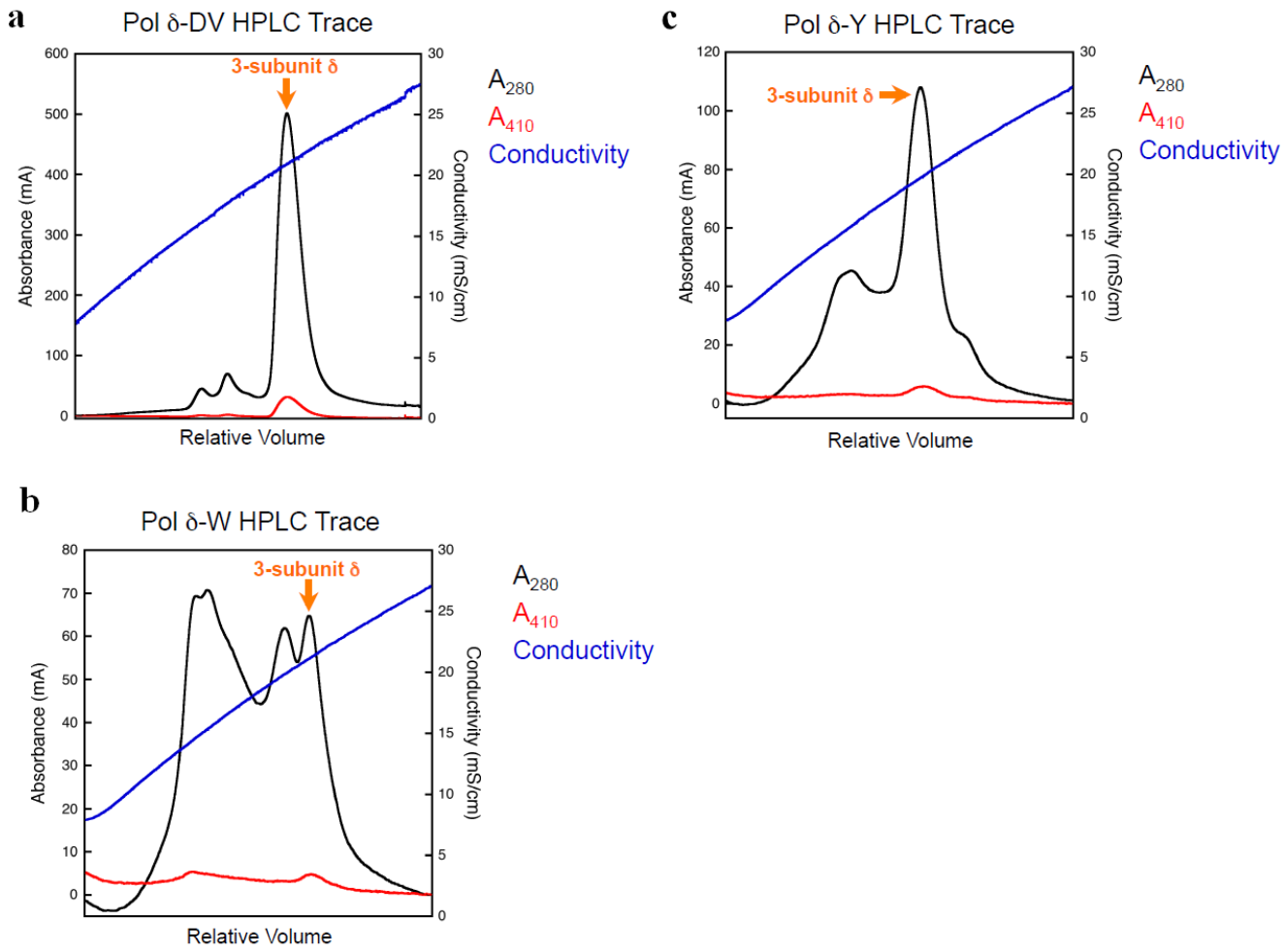


Figure 6.2 HPLC traces from Pol δ purification on a MonoS column in a 150 – 750 mM NaAc gradient. (**a**) Pol δ DV without any additional mutations elutes primarily as a three-subunit complex that contains substantial absorbance at 410 nm associated with the presence of a [4Fe4S] cluster. (**b**, **c**) Preps from Pol δ DV W1053A (**b**) and Y1078A (**c**), in contrast, are lower yielding (note the scale) and contain a much higher proportion of degraded complexes eluting at earlier time points, indicative of lower stability. Despite low stability, some absorbance at 410 nm remains visible, supporting the assumption that intact complexes bind a [4Fe4S] cluster.

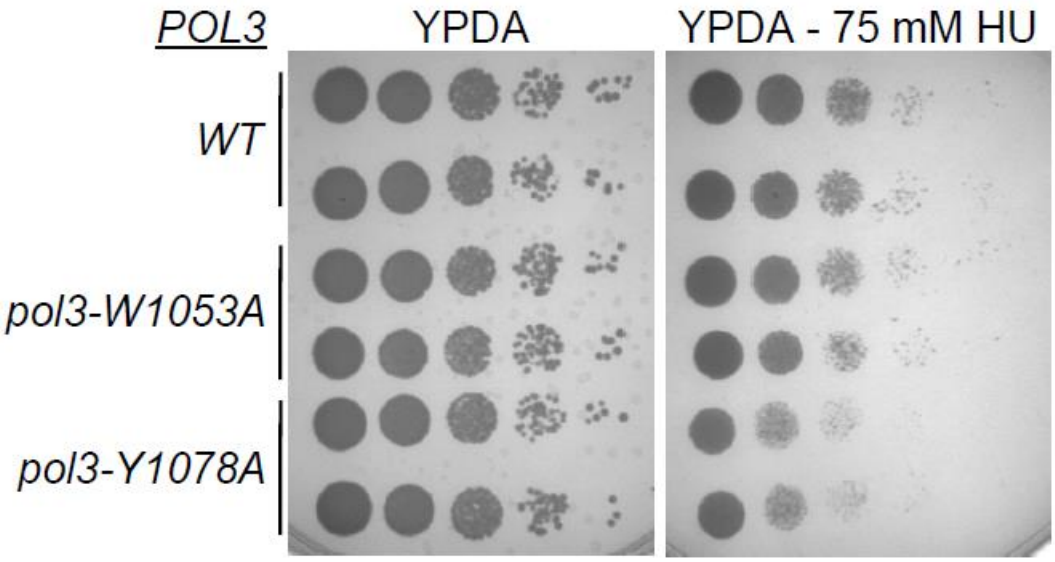


Figure 6.3 Growth of budding yeast containing WT, W1053A, or Y1078A Pol δ on YPDA medium with and without 75 mM hydroxyurea. Y1078A showed a slightly enhanced sensitivity to hydroxyurea, although neither mutant was sensitive to UV irradiation or camptothecin treatment (not shown).

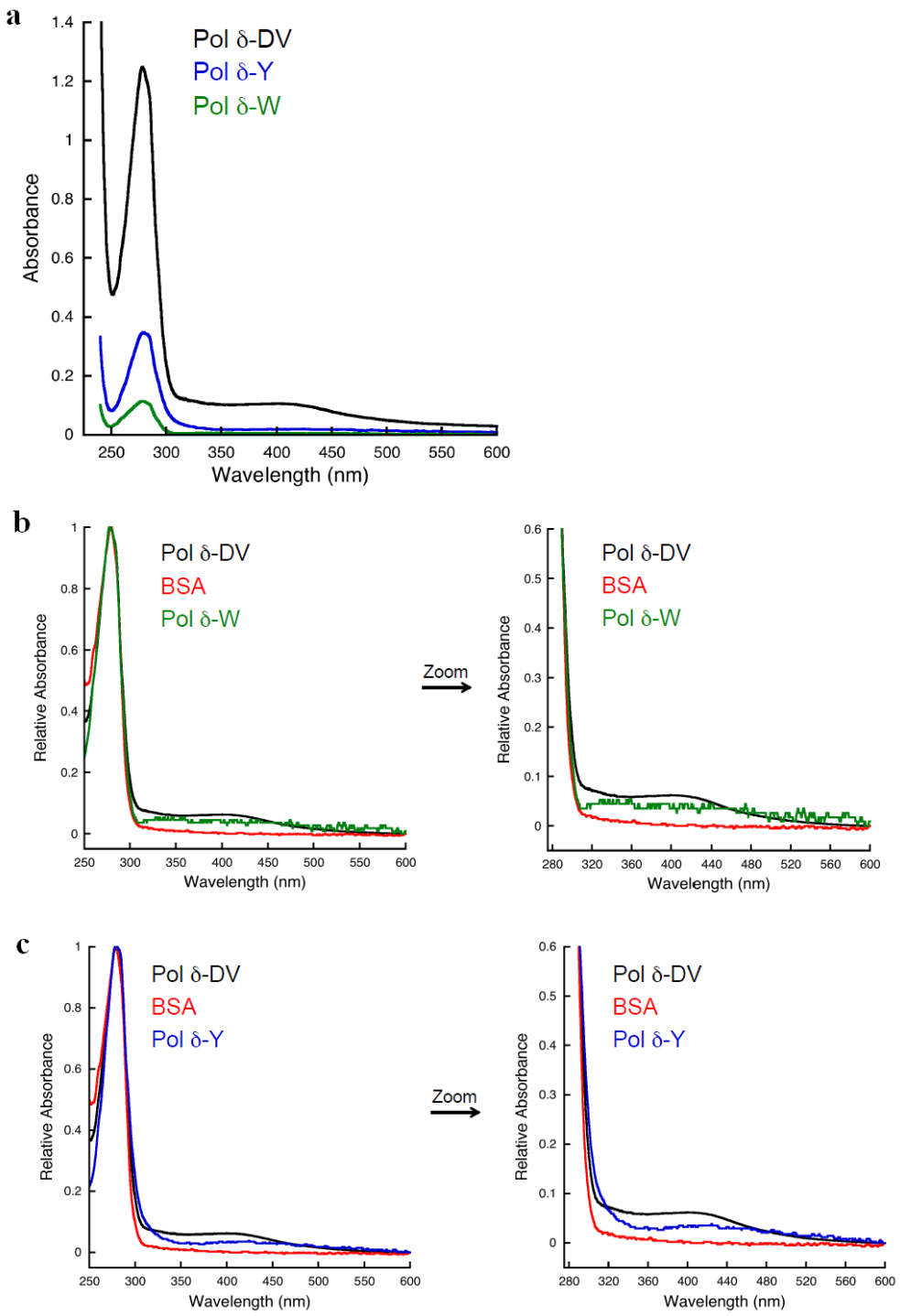


Figure 6.4 UV-visible spectra of Pol δ DV and the mutants W1053A and Y1078A. (**a**) Both mutants were much more dilute than Pol δ DV, making [4Fe4S] cluster concentration difficult to calculate. (**b**, **c**) To see if [4Fe4S] cluster absorbance at 410 nm was present, spectra were normalized to A280 and compared with BSA, which does not bind a cluster. Although not ideal, this analysis does suggest [4Fe4S] binding, in agreement with the fact that both form intact complexes and show 410 nm absorbance by HPLC during purification.

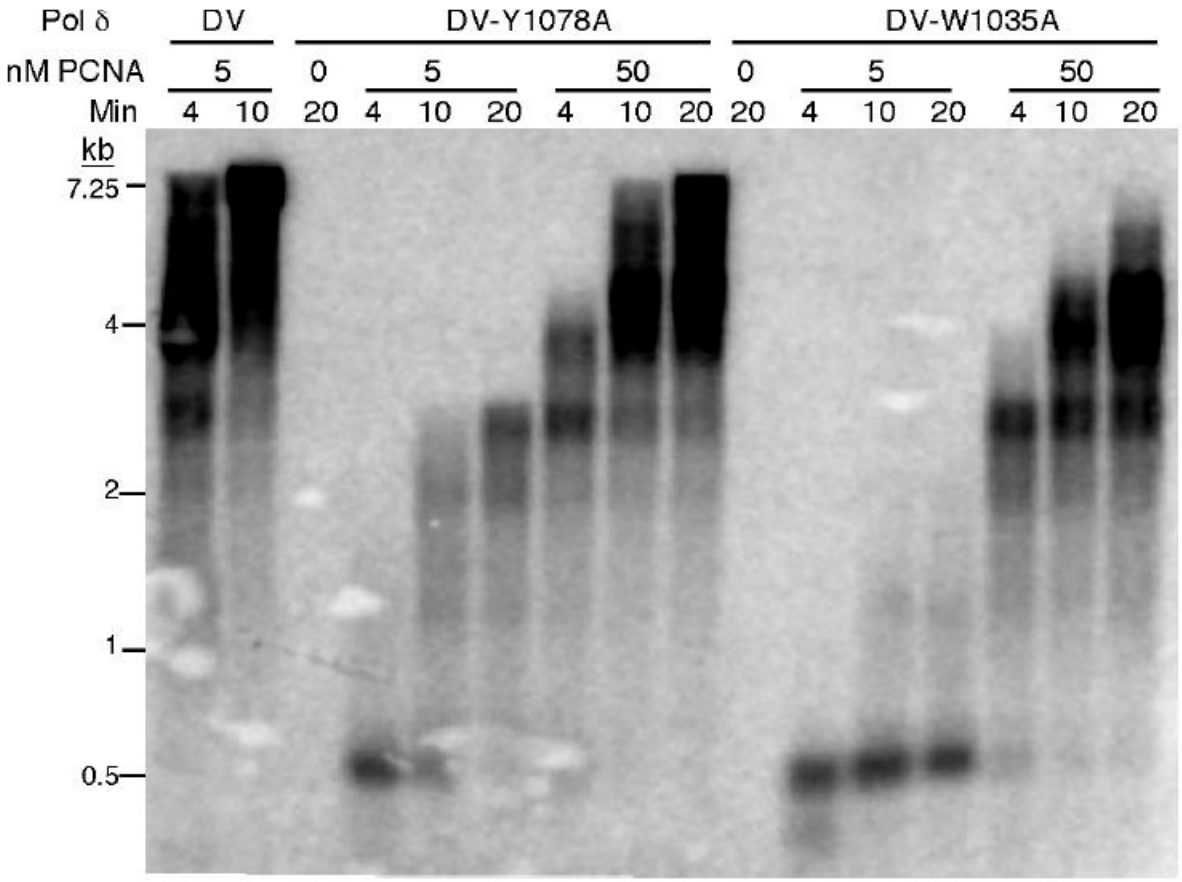


Figure 6.5 Primed ssM13mp18 plasmid DNA replication by Pol δ variants in a standard lagging strand replication assay. Pol δ DV (left) shows complete replication of the 7.3 kb substrate by 10 minutes with stoichiometric PCNA. Both Y1078A (center) and W1053A (right) are less processive, and require a 10-fold excess of PCNA for complete replication. Notably, the processivity defect is more pronounced in W1053A, which also appeared less stable during purification.

Electrochemical Characterization of Pol δ Mutants

With basic cell growth assays and biochemical characterization completed, we next proceeded to investigate the redox properties of these mutants using electrochemistry on DNA-modified gold electrodes. All electrochemical experiments were run on multiplexed chips in parallel with Pol δ DV (Figure 6.6). To maximize signal size at these low concentrations, we used 500 nM Pol δ pre-mixed with 5.0 μM PCNA, 80 μM dATP, and 8 mM $MgAc_2$. CT-proficient Pol δ DV gives substantial signals under these conditions (~5-10 nC maximum CV peak charge), and signals remain reliably detectable at least down to a tenth of this size (*2*). Thus, we reasoned that these conditions should be sufficient to observe the mutants even if they were nearly completely CT-deficient relative to WT. However, given that both mutants showed lower processivity in activity assays and were thus less likely to bind DNA as effectively as unmodified Pol δ DV, one half of each chip contained DNA with a 3′ ddC substrate trap; the remaining quadrants contained unmodified DNA.

Small but quantifiable electrochemical signals occur with both mutants; the midpoint potentials are 121 ± 4 mV and 120 ± 3 mV versus NHE for W1053A and Y1078A, respectively. As expected, the potentials are very similar to Pol δ DV (113 ± 5 mV). Because the relatively large CV capacitance can impede signal quantification and make comparison difficult for small signals, SQWV, which minimizes background current, was used for this purpose instead. By SQWV, the W1053A and Y1078A peak areas are 21 ± 7 % and 25 ± 4 % of unmodified Pol δ DV signals (Figure 6.6). Interestingly, no obvious differences occurred in the presence of the ddC substrate trap, although signal sizes were somewhat less variable on this DNA. This result can most likely be attributed to the saturation of available DNA with protein: loosely-packed monolayers contain 0.2 – 0.3 pmol DNA per electrode (*22*), while 10 pmol of Pol δ was present.

Like Pol δ DV, the Y1078A signal grows in over time, although it never reaches the same extent (Figure 6.7). In contrast, W1053A remains small and even decreases over time, possible indicative of sample degradation over extended time in aerobic conditions.

Interestingly, after overnight incubation, a secondary, irreversible reductive peak centered around -40 mV versus NHE develops in Y1078A; at the same time point, the W1053A signal is too small to see if the same additional peak develops. In unstable human MUTYH mutants, an irreversible peak at the same potential is attributable to oxidative degradation of the [4Fe4S] cluster to the $[3Fe4S]^{+}$ state (*23*). Such degradation in the less stable Pol δ mutants is a reasonable explanation for this peak, although the low concentrations of these mutants precluded UV-visible and EPR spectroscopic identification. In MUTYH mutants, cluster degradation is dependent on O_2, and the secondary peak can be significantly amplified by bulk electrolysis in an aerobic environment (*23*). Given the impossibility of spectroscopic analysis in the case of Pol δ DV Y1078A, we undertook aerobic bulk electrolysis as an alternative to see if the pattern matched that of MUTYH (Figure 6.8). As a control, we repeated this experiment with unmodified Pol δ DV (Figure 6.8). In these experiments, UV-visible spectra were taken before and after oxidation, and one half of the sample was left untreated and incubated on a multiplexed chip. CV and SQWV scans were taken of both samples on a multiplexed chip before and after electrolysis. After 1 hour, aerobic oxidation had reached its maximum extent, and the oxidized sample was transferred back to the multiplexed chip once a UV-visible spectrum had been taken. Immediately after electrolysis, the Y1078A signal was slightly smaller than the untreated control, although this may have been in part due to minor sample loss. Despite sample loss, the secondary peak was proportionally larger relative to the main peak at this point (Figure 6.8). All signals disappear entirely within 2 hours of incubation in the oxidized sample; signals in the untreated

sample gradually decrease, but not to the same extent (Figure 6.8). UV-visible spectra taken after electrolysis show an increase in absorbance at 800 nm, and the entire spectrum is U-shaped, consistent with aggregation (Figure 6.8). When repeated with unmodified Pol δ DV, the UV-visible spectrum post-electrolysis shows a general increase in cluster absorbance associated with oxidation with no evidence for aggregation (Figure 6.8). Furthermore, electrochemical signals remain indistinguishable from untreated sample even after extended incubation, and no peak indicative of $[3Fe4S]^{+}$ cluster is observed. Overall, these results can best be explained as a difference in stability between unmodified Pol δ DV and the Y1078A mutant. The cluster in unmodified Pol δ DV is stable in air even when oxidized to the $[4Fe4S]^{3+}$ state, likely as a result of tight binding with the B-subunit. In contrast, Y1078A, although capable of forming intact complexes, is structurally perturbed, making the cluster more susceptible to oxidative degradation in air. The loss of signal and evidence for aggregation suggest that the cluster in this mutant degrades upon oxidation, causing complexes to come apart and aggregate. Pol3 is known to be aggregation-prone in the absence of B-subunits, providing further support for our assessment of these results.

Notably, absent forced oxidation by bulk electrolysis, cluster degradation as observed by electrochemistry takes several hours to develop in the Pol δ mutants, with Y1078A appearing as the more stable of the two. This observation suggests that, while these mutants are less stable than WT Pol δ, at least Y1078A is stable on short time scales. Overall, CT deficiency is apparent in both mutants, although whether it is the result of involvement of the residue in the CT pathway or simply complex degradation is ambiguous. Since Y1078A appears somewhat more stable than W1053A, however, a case can be made that this residue is part of a CT pathway. Even with Y1078A, however, some signal loss due to complex degradation cannot be excluded.

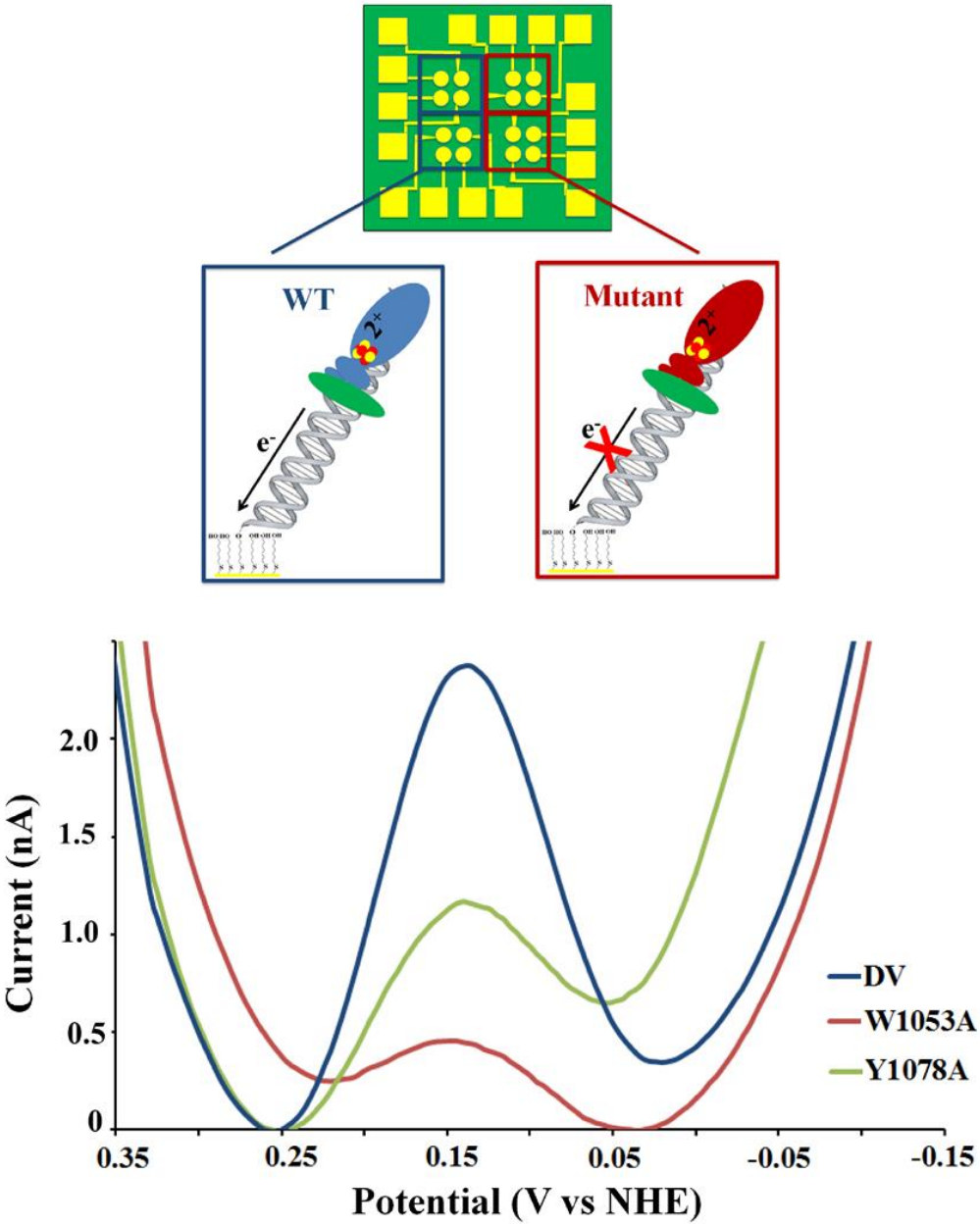


Figure 6.6 Square wave voltammetry (SQWV) of Pol δ DV and the W053A and Y1078A mutants. Polymerase variants were incubated in parallel on a multiplexed chip (**top**), and each quadrant contained 25 µL of 500 nM Pol δ DV, 5.0 µM PCNA, 80 µM dATP, and 8.0 mM $MgAc_2$ in storage buffer (20 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, 0.01% decaethylene glycol monododecyl ether v/v). Both mutants showed marked deficiency in redox activity relative to Pol δ DV, with signal area only 20-25% of the unmodified protein (**bottom**). SQWV scans run from left to right, and were taken at 15 Hz frequency and 25 mV amplitude.

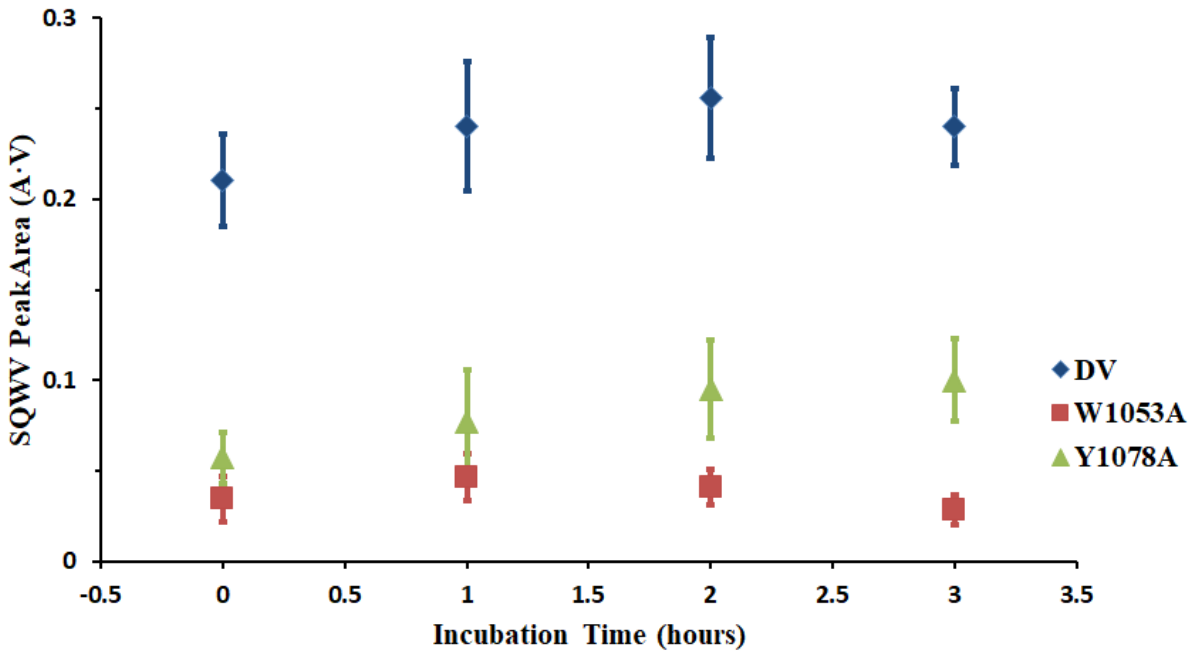


Figure 6.7 SQWV signal growth and stability over time for 500 nM Pol δ DV and the mutants W1053A and Y1078A. Pol δ DV equilibrates rapidly and the signal remains stable for several hours. Y1078A has a similar patter, although signals are much smaller and equilibration is more prolonged. W0153A plateaus early, after which the signal steadily decreases. The loss of signal in W1053A is most likely attributable to oxidative degradation in the aerobic environment. Experiments were performed in storage buffer (20 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, 0.01% decaethylene glycol monododecyl ether v/v) with 5.0 μM PCNA, 80 μM dATP, and 8.0 mM $MgAc_2$.

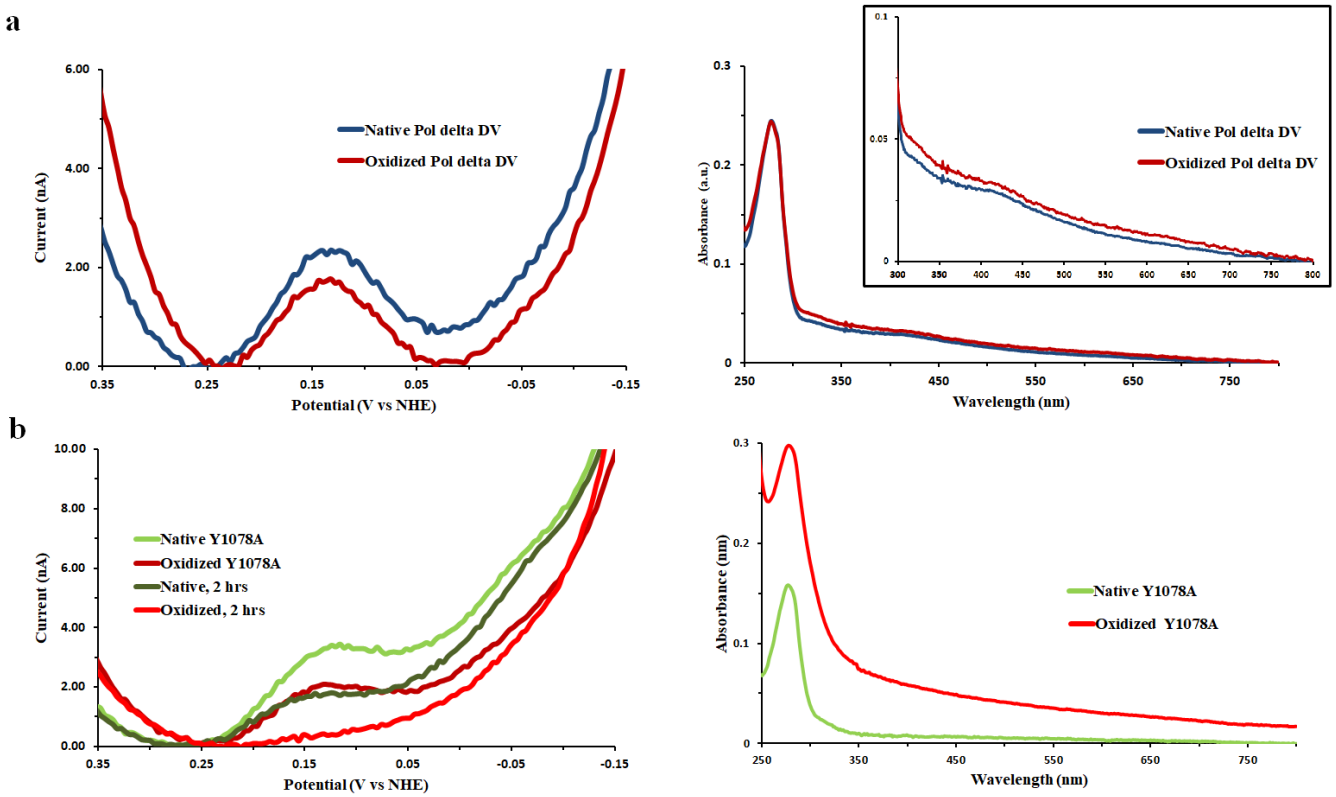


Figure 6.8 SQWV (left) and UV-visible spectroscopy (right) of untreated and aerobically oxidized Pol δ DV (**a**) and Y1078A (**b**). (**a**) Oxidation by bulk electrolysis (0.412 mV vs NHE) in air did not result in any appreciable changes in Pol δ DV electrochemistry (the smaller signal can be attributed to a minor loss of sample), and the UV-visible spectrum was unaltered except for a broad increase in [4Fe4S] cluster absorbance associated with oxidation (inset at right). (**b**) Y108A showed a broad, irreversible peak around -50 mV associated with $[3Fe4S]^{+}$ cluster formation; this peak underwent a proportional increase in size relative to the main peak after oxidation, and the signal decreased and fully disappeared after two hours of incubation. The UV-visible spectra of oxidized Y1078A showed evidence of aggregation (an elevated absorbance at 800 nm and a U-shaped spectrum), consistent with cluster degradation and aggregation of disassembled complexes. Electrolysis was performed with 500 nM Pol δ in storage buffer and 5.0 μM PCNA present. SQWV was performed at 15 Hz frequency and 25 mV amplitude.

Activity Assays with Electrochemically Oxidized Pol δ Mutants

To see if Y1078A and W1053A were capable of the reversible oxidative stalling seen in unmodified Pol δ DV, we carried out activity assays following electrolysis under anaerobic conditions. Anaerobic conditions were used to prevent the degradation seen in aerobically oxidized Y1078A, and all assays were run in previously degassed buffers. In contrast to aerobic oxidation experiments with Y1078A, bulk electrolysis at 0.412 mV vs NHE gives minimal yields for this mutant in an anaerobic environment (Figure 6.9). Consistent with minimal or no cluster oxidation, activity levels are indistinguishable from those of untreated anaerobic samples (Figure 6.10). Unsurprisingly, reduction of the same protein stock at -0.188 mV vs NHE had no effect, and DNA synthesis by all three samples overlay almost perfectly even across three replicates (Figure 6.10). Although the stability of this mutant is certainly lower than WT, the fact that activity levels do not change between a 20-minute oxidation and an additional 20-minute reduction indicates that complex formation was not affected under these conditions.

Unexpectedly, W1053A underwent extensive oxidation at 0.412 V vs NHE based on bulk electrolysis yields (Figure 6.9). As with unmodified Pol δ DV, oxidized W1053A shows a decrease in activity, although the difference is much greater (Figure 6.10). Notably, re-reduction does not restore activity in this mutant. Taken together, the high susceptibility to oxidation and its irreversible nature suggest complex degradation. This was somewhat unexpected under anaerobic conditions, but follows the general trend seen in aerobic electrochemistry of this mutant.

In summary, the inertness of Y1078A in the absence of oxygen is consistent with disruption of the CT pathway by this mutation. Considered alongside our aerobic electrolysis experiments, it appears that this mutant may be more sensitive to oxygen than WT Pol δ, and we

were interested in observing the response of Y1078A yeast cells to oxidative stress conditions. In contrast, all of our experiments point to CT deficiency in W0153A as being due to structural instability rather than a break in the CT pathway. Nonetheless, we were still interested in the ability of Pol δ W1053A to permit yeast survival under oxidative stress, and we thus proceeded to *in vivo* experiments with both mutants.

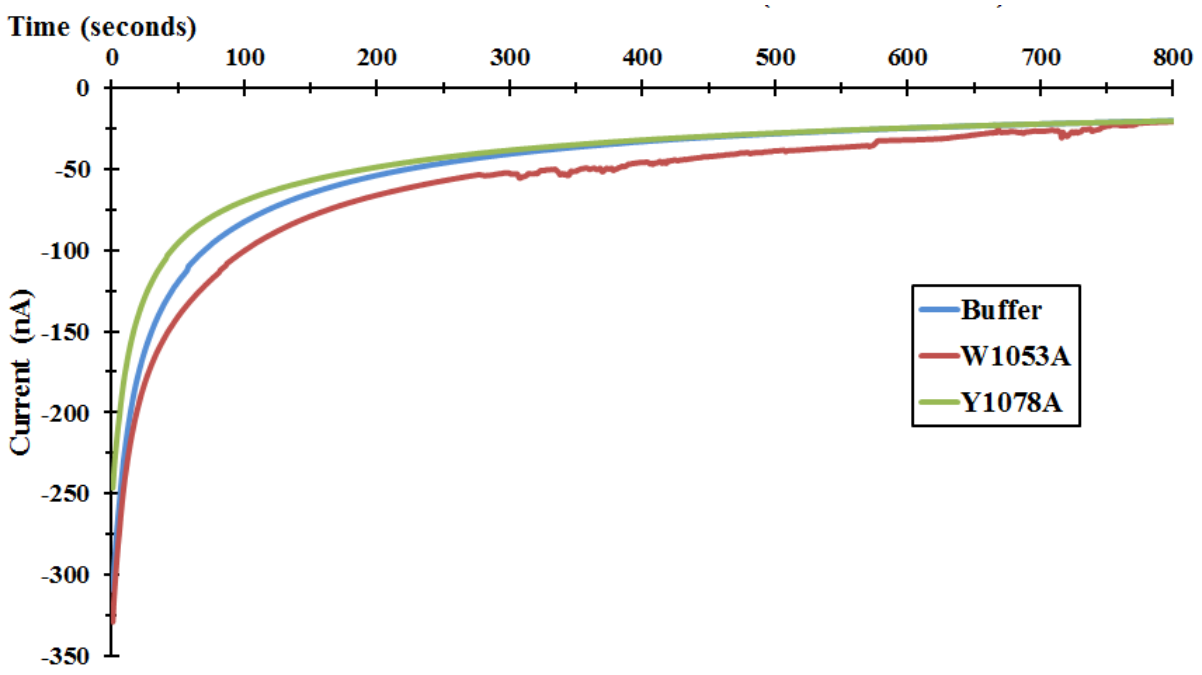


Figure 6.9 Anaerobic bulk electrolysis traces of 190 nM Pol δ DV W1053A (red) and Y1078A (green) compared to buffer trace (blue). Under anaerobic conditions, Y1078A was insensitive to oxidation, with charge yields below buffer background. W1053A unexpectedly underwent near-complete oxidation, as evidenced by the increased area under the curve relative to buffer alone. Electrolysis was carried out at 0.412 V vs NHE on a gold rod electrode, with polymerase diluted to 190 nM in degassed storage buffer (20 mM HEPES, pH 7.4, 350 mM NaAc, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, 0.01% decaethylene glycol monododecyl ether v/v).

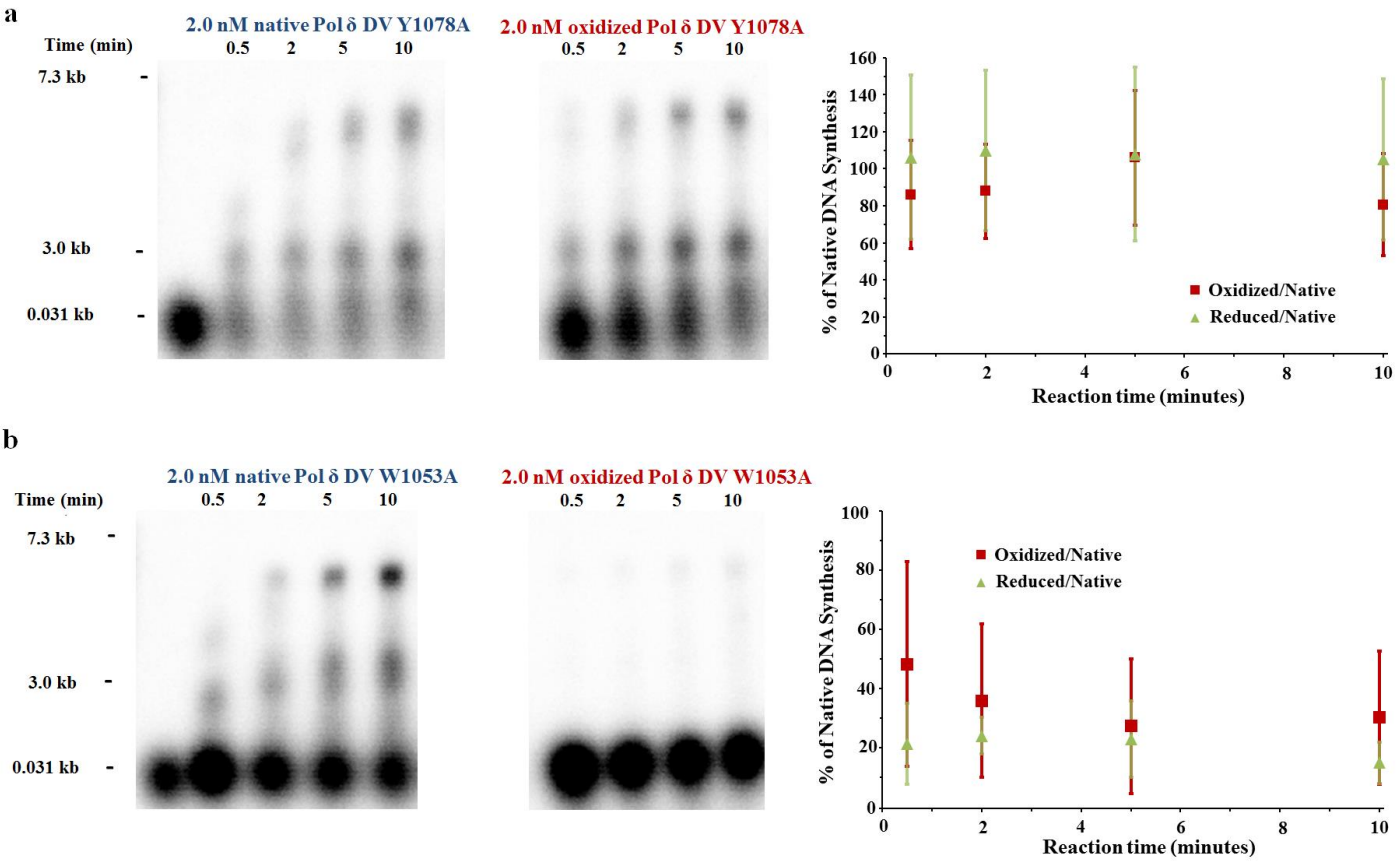


Figure 6.10 Alkaline agarose gels (left) and quantification of DNA synthesis (right) by oxidized or re-reduced Pol δ DV Y1078A (**a**) and W1053A (**b**). (**a**) Y1078A was insusceptible to oxidation (Figure 6.9), and activity of both oxidized and "re-reduced" samples were indistinguishable. (**b**) W1053A, which underwent complete oxidation (Figure 6.9), showed a pronounced loss of activity. Unlike unmodified Pol δ DV, re-reduction could not restore activity, indicating that this mutant is susceptible to cluster degradation and complex disassembly even under anaerobic conditions. Oxidation was performed by bulk electrolysis at 0.412 V vs NHE, while reduction was carried out on the same electrode at -0.188 V vs NHE. Quantification is the total DNA synthesis by oxidized or re-reduced sample divided by that of native, or untreated, sample. Error bars are the standard deviation of the mean of 3 separate experiments.

Evaluating the Sensitivity of Yeast Containing Pol δ Mutations to Oxidative Stress

In order to determine the effect of Y1078A and W1053A in cells under conditions of oxidative stress, cells were treated with exogenous H_2O_2 and spot plated. Unfortunately, no significant differences in cell survival were apparent even at low (2 mM) concentrations (data not shown), so we sought an alternative method to sensitize cells to oxidative stress. In an effort to enhance the differential in survival, the GSH2 gene was knocked out from yeast cells. GSH2 encodes glutathione synthetase; since glutathione plays a critical role in maintaining a reducing environment, this modification was expected to make cells much more susceptible to H_2O_2 treatment. Once prepared, *Δgsh* cells were treated with either 2 mM H_2O_2 or 50 mM menadione salt to induce oxidative stress (Figure 6.11). Menadione, which is a superoxide dismutase inhibitor, was included to provides an internally focused source of oxidative stress, which we thought might be more effective than H_2O_2. *Δgsh* cells proved to be very sensitive to both treatment, but cell death is so rapid in all cases that no consistent differences in survival can be observed even at H_2O_2 concentrations as low as 2 mM (Figure 6.11). In summary, these results are equivocal due to extremely rapid and non-selective cell death. Indeed, the cell-wide effects of the oxidants used is certain to enhance background effects, making the likelihood of observing differences in DNA replication responses much lower than ideal.

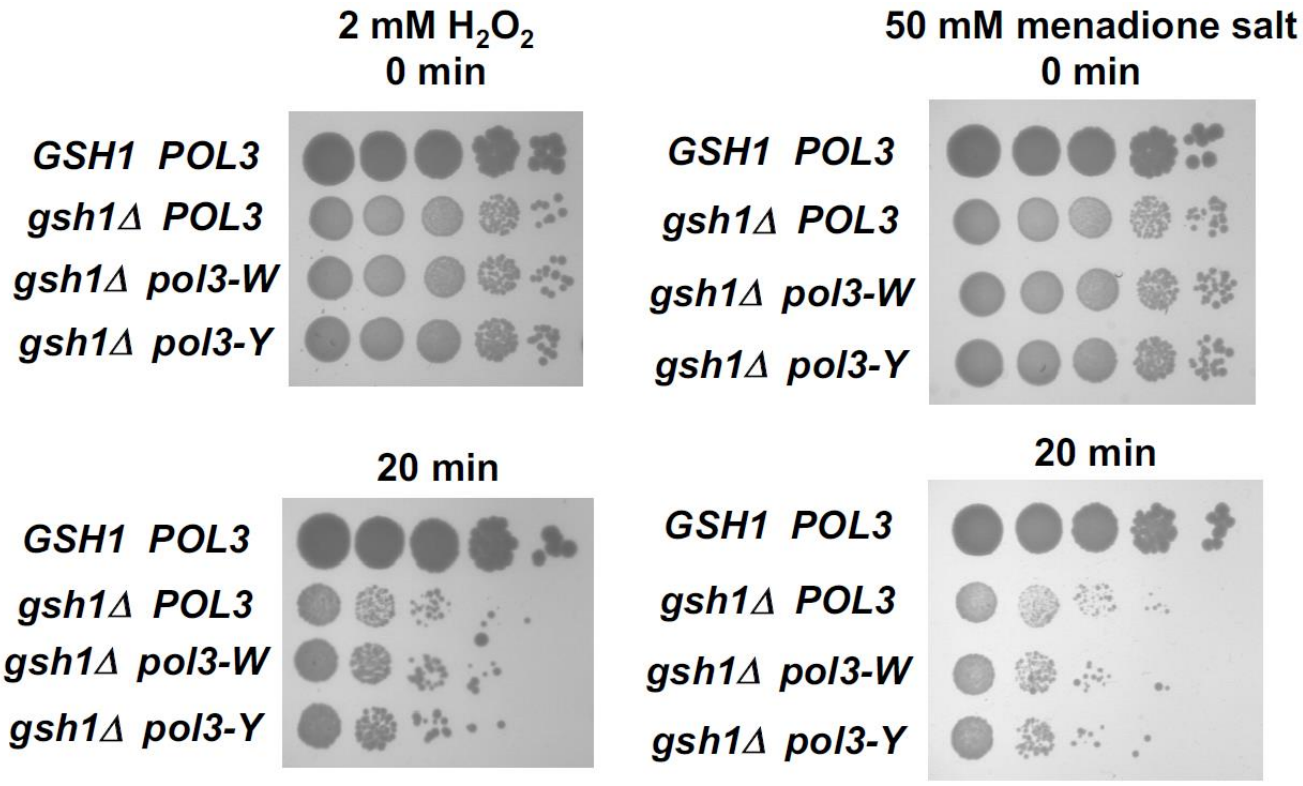


Figure 6.11 Growth of glutathione-deficient yeast strains containing WT or CT-defective Pol δ before (0 min) and after (20 min) treatment with either 2 mM H_2O_2 (left) or 50 mM menadione (right) to induce oxidative stress conditions. No consistent growth defect was observed for either Y1078A or W1053A cells, although both strains were more generally sensitive to oxidants than cells with no defects in glutathione production (top row). The lack of differences is most likely attributable to indiscriminate and delocalized cellular damage, with any actual sensitivity due to polymerase mutations masked by these non-specific effects.

Discussion

In these studies, we have attempted to identify and evaluate aromatic residues near the Pol δ [4Fe4S] cluster that might play a role in the electron transfer pathway between the cluster and DNA. From sequence alignments and comparison with a Pol α structure, we identified two residues within 15 Å of the cluster, W1053 and Y1078. The alanine mutations of both residues could be purified as intact 3-subunit complexes and retained processive activity in the presence of DNA, although multiple lines of evidence point to these mutants being less stable than WT Pol δ. Electrochemically, both were markedly CT-deficient relative to unmodified Pol δ DV, and both showed some oxygen sensitivity. Y1078A was largely redox-inert in the absence of oxygen, and showed no differences in activity after attempted anaerobic oxidation. In contrast, W1053A degraded upon bulk electrolysis even under anaerobic conditions, making the role of this residue in CT unclear. Preliminary *in vivo* studies designed to test the effect of these mutations on the oxidative stress response in yeast were similarly equivocal, with both exogenous and endogenously-induced oxidants indiscriminately killing cells too quickly for survival differences to be assessed.

Overall, these studies illustrate the challenges associated with making mutations in aromatic residues in the Pol3 C-terminal domain. While inconclusive, our work nonetheless represents a start to addressing interesting questions about Pol δ, including the role of the cluster in responding to oxidative stress and the composition of the electron transfer pathway between DNA and the cluster. It is likely that the structural instability that hindered these studies could be mitigated by using mutations to phenylalanine rather than alanine, as has been successfully done for human DNA primase (*12*). In particular, our results with Y1078A suggest that future studies would benefit from a more isomorphic mutation. With respect to the *in vivo* assays, the overall

strategy is worth pursuing with more stable mutants. However, our current results do not provide conclusive evidence for a role of the [4Fe4S] cluster in responding to oxidative stress. These failures can, however, provide a guide for more successful future experiments. The major problem with the studies as described is in the generic nature of the cellular damage. The nucleus is not a specific target for either H_2O_2 or menadione-generated superoxide, and both will indiscriminately damage biomolecules throughout the cell (*24*). One possible means of enhancing the readout of Pol δ-specific effects is S-phase synchronization and release, which would at least ensure that all affected cells were undergoing DNA replication (*25*). In addition, switching the oxidant from indiscriminate oxidizing species to photosensitizers, such as the singlet oxygen-generators Rose Bengal and methylene blue, may be of benefit in addressing the issue of the *in vivo* function of the cluster (*26*).

The importance of building on this work is emphasized by the occurrence of disease-relevant mutations in this region of Pol δ. Mutations in this critical region are rare, although the yeast Pol3-13 allele, which contains a serine in place of cysteine as one [4Fe4S] cluster ligand, is synthetic lethal with mutations in ISC machinery (*1*). Most relevant to the mutants investigated here, a case of pediatric lymphoid cancer with an R1053C mutation in Pol3 has recently been described (*27*). Arginine 1053 is two residues away from W1055, and thus has the potential to be involved in DNA binding or hydrogen bonding networks near the cluster. It will thus be of interest to repeat the experiments described herein with phenylalanine mutants, and to conduct similar experiments on Pol3-13 and R1053C.

References

1) Netz, D. J. A.; Stith, C. M.; Stümpfig, M.; Köpf, G.; Vogel, D.; Genau, H. M.; Stodola, J. L.; Lill, R.; Burgers, P. M. J.; Pierik, A. J. Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat. Chem. Biol.* **2012**, *8*, 125 – 132.

2) Bartels, P.L.; Stodola, J.L.; Burgers, P.M.J.; Barton, J.K. A Redox Role for the [4Fe4S] Cluster of Yeast DNA Polymerase δ. *J. Am. Chem. Soc.* **2017**, *139*, 18339 – 18348.

3) Arnold, A.R.; Grodick, M.A.; Barton, J.K. DNA Charge Transport: from Chemical Principles to the Cell. *Cell Chem. Biol.* **2016**, *23*, 183 – 197.

4) Thayer, M.M.; Ahern, H.; Xing, D.; Cunningham, R.P.; Tainer, J.A. Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *The EMBO J.* **1995**, *14*, 4108 – 4120.

5) Guan, Y.; Manuel, R.C.; Arvai, A.S.; Parikh, S.S.; Mol, C.D.; Miller, J.H.; Lloyd, R.S.; Tainer, J.A. MutY catalytic core, mutant, and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nat. Struct. Biol.* **1998**, *5*, 1058 – 1064.

6) Fan, L.; Fuss, J.O.; Cheng, Q.J.; Arvai, A.S.; Hammel, M.; Roberts, V.A.; Cooper, P.K.; Tainer, J.A. XPD Helicase Structure and Activities: Insights into the Cancer and Aging Phenotypes from XPD Mutations. *Cell* **2008**, *133*, 789 – 800.

7) Vaithiyalingam, S.; Warren, E.M.; Eichman, B.F.; Chazin, W.J. Insights into Eukaryotic DNA priming from the structure and functional interactions of the 4Fe-4S cluster domain of human DNA primase. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 13684 – 13689.

8) Fromme, J.C.; Verdine, G.L. Structure of a trapped endonuclease III-DNA intermediate. *The EMBO J.* **2003**, *22*, 3461 – 3471.

9) Fromme, J.C.; Bannerjee, A.; Huang, S.J.; Verdine, G.L. Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase. *Nature* **2004**, *427*, 652 – 656.

10) Zhou, C.; Pourmal, S.; Pavletich, N.P. Dna2 nuclease-helicase structure, mechanism, and regulation by RPA. *eLIFE* **2015**, 4:e09832.

11) Romano, C.A.; Sontz, P.A.; Barton, J.K. Mutants of the Base Excision Repair Glycosylase, Endonuclease III: DNA Charge Transport as a First Step in Lesion Detection. *Biochemistry* **2011**, *50*, 6133 – 6145.

12) O'Brien, E.; Holt, M.E.; Thompson, M.K.; Salay, L.E.; Ehlinger, A.C.; Chazin, W.J.; Barton, J.K. The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport. *Science* **2017**, *355*, 813.

13) Gray, H.B.; Winkler, J.R. Electron flow through metalloproteins. *Biochimica et Biophysica Acta* **2010**, *1797*, 1563 – 1572.

14) Dey, A.; Jenney, F.E., Jr.; Adams, M.W.W.; Babini, E.; Takahashi, Y.; Fukuyama, K.; Hodgson, K.O.; Hedman, B.; Solomon, E.I. Solvent Tuning of Electrochemical Potentials in the Active Sites of HiPIP Versus Ferredoxin. *Science* **2007**, *318*, 1464 – 1468.

15) Ha, Y.; Arnold, A.R.; Nuñez, N.N.; Bartels, P.L.; Zhou, A.; David, S.S.; Barton, J.K.; Hedman, B.; Hodgson, K.O.; Solomon, E.I. Sulfur K-Edge XAS Studies of the Effect of DNA Binding on the [Fe4S4] Site in EndoIII and MutY. *J. Am. Chem. Soc.* **2017**, *139*, 11434 – 11442.

16) Fortune, J.M.; Stith, C.M.; Kissling, G.E.; Burgers, P.M.; Kunkel, T.A. RPA and PCNA suppress formation of large deletion errors by yeast DNA polymerase δ. *Nucleic Acids Res.* **2006**, *34*, 4335 – 4341.

17) Ayyagari, R.; Impellizzeri, K.J.; Yoder, B.L.; Gary, S.L.; Burgers, P.M. A mutational analysis of the yeast proliferating cell nuclear antigen indicates distinct roles in DNA replication and repair. *Mol. Cell. Biol.* **1995**, *15*, 4420 – 4429.

18) Gomes, X.V.; Gary, S.L.; Burgers, P.M. Overproduction in *Escherichia coli* and characterization of yeast replication factor C lacking the ligase homology domain. *J. Biol. Chem.* **2000**, *275*, 14541 – 14549.

19) Henricksen, L.A.; Umbricht, C.B.; Wold, M.S. Recombinant replication protein A: expression, purification, and functional characterization. *J. Biol. Chem.* **1994**, *269*, 11121 – 11132.

20) Slinker, J.D.; Muren, N.B.; Gorodetsky, A.A.; Barton, J.K. Multiplexed DNA-Modified Electrodes. *J. Am. Chem. Soc.* **2010**, *32*, 2769 – 2774.

21) Klinge, S.; Núñez-Ramírez, R.; Llorca, O.; Pellegrini, L. 3D architecture of DNA Pol α reveals the functional core of multi-subunit replicative polymerases. *EMBO J.* **2009**, *28*, 1978 – 1987.

22) Boon, E.M.; Salas, J.E.; Barton, J.K. An Electrical Probe of Protein-DNA Interactions on DNA-Modified Surfaces. *Nat. Biotechnol.* **2002**, *20*, 282 – 286.

23) McDonnell, K.J.; Chemler, J.A.; Bartels, P.L.; O'Brien, E.; Marvin, M.L.; Ortega, J.; Stern, R.H.; Raskin, L.; Li, G.; Sherman, D.H.; Barton, J.K.; Gruber, S.B. A Human MUTYH Variant Linking Colonic Polyposis to Redox Degradation of the $[4Fe4S]^{2+}$ Cluster. **2018**, *Submitted*.

24) Imlay, J.A. Pathways of Oxidative Damage. *Annu. Rev. Microbiol.* **2003**, *57*, 395 – 418.

25) Breeden, L.L. Apha-Factor Synchronization of Budding Yeast. *Methods Enzymol.* **1997**, *283*, 332 – 341.

26) Kochevar, I.E.; Redmond, R.W. Photosensitized Production of Singlet Oxygen. *Methods Enzymol.* **2000**, *319*, 20 – 28.

27) Ma, X.; Edmonson, M.; Yergeau, D.; Muzny, D.M.; Hampton, O.A.; Rusch, M.; Song, G.; Easton, J.; Harvey, R.C.; Wheeler, D.A.; Ma, J.; Doddapaneni, H.; Vadodaria, B.; Wu, G.; Nagahawatte, P.; Carroll, W.L.; Chen, I.-M.; Gastier-Foster, J.M.; Relling, M.V.; Smith, M.A.; Devidas, M.; Guidry Auvil, J.M.; Downing, J.R.; Loh, M.L.; Willman, C.L.; Gerhard, D.S.; Mullighan, C.G.; Hunger, S.P.; Zhang, J. Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. *Nat. Commun.* **2016**, *6*, 6604.

Chapter 7

An Electrochemical Assay for the Detection of DNA Polymerase Activity

This work was done as a collaboration with Joseph Stodola and Prof. Peter Burgers at the Washington University School of Medicine in St. Louis.

P. Bartels fabricated devices and performed all experiments and associated analysis. Initial versions of the integrated device were designed and tested by C. Pheeney, who assisted in design for the device described herein. J. Stodola prepared DNA polymerase and PCNA proteins, and A. Arnold purified endonuclease III.

Introduction

Electrochemical detection of enzymatic activity is of great interest for diagnostic purposes, and a range of assays have been developed for these purposes (*1*). Given their importance in disease, DNA-processing enzymes make particularly intriguing targets, and assays that take advantage of DNA-mediated charge transport (DNA CT) have been successfully employed to detect protein binding and activity (*2-4*). DNA CT involves the long-range transfer of electrons through the π-stacked base pairs, and is sharply attenuated in the presence of lesions or mismatches that perturb base stacking (*5*). By appending covalent redox probes to one end of DNA tethered to a gold electrode, it is thus possible to assess the activity of enzymes that repair lesions or modify bases (i.e., methylation) using electron transfer to the probe as a readout. With careful choice of DNA substrate and reporter, the activity of a specific enzyme can be detected even at low concentrations and from heterogeneous samples (*4*). Although developed for diagnostic purposes, similar systems can also be used to provide insight into fundamental biochemical processes. This is well illustrated by the used of DNA-modified gold electrodes to monitor such activities as thymine dimer repair by photolyase and DNA unwinding activity by ATP-dependent [4Fe4S] helicases (*6-8*).

Although electrochemical activity assays have been less frequently performed for [4Fe4S] proteins, the redox behavior of a range of such enzymes in DNA repair and replication has been extensively studied on DNA-modified electrodes (*9*). Earlier studies with these proteins have generally used standard in-solution activity assays with a radioactivity-based readout to assess the role of the [4Fe4S] cluster (*10*), but electrochemical activity assays would provide some appealing advantages. Beyond bypassing the need for radioactive reagents, electrochemical

assays provide a conveniently rapid readout (minutes to hours versus days to run and expose gels) and require no additional sample preparation steps.

Among the [4Fe4S] proteins studied recently, DNA polymerase (Pol) δ has been shown to be redox active and to undergo reversible stalling upon oxidation of the cluster to the $[4Fe4S]^{3+}$ state (*11*). In these electrochemical experiments, Pol δ was incubated on a gold electrode containing substrate DNA with a 9-base 5′ single-stranded overhang. Given this substrate design, we surmised that it might be possible to use a covalent redox probe such as methylene blue (MB′) tethered to the end of the overhang to report on polymerase activity. In such a system, the readout would be a change from a rapid, surface-mediated methylene blue electron transfer process before reaction to a slower, DNA-mediated process as the probe intercalated into newly-synthesized duplexed DNA after reaction. In addition to activity detection, we also hoped to use this assay as an alternative to standard gel-based assays that compare the activity of oxidized and untreated Pol δ.

Biochemical assays are ideally done with minimal usage of precious protein sample. Up to this point, most electrochemical experiments were performed on multiplexed devices containing 16 independent electrodes separable into four quadrants (*12*). These devices provide the ability to work with up to four distinct monolayers in parallel in addition to providing built-in experimental replicates. While each quadrant requires only 20 μL solution, experiments generally involved the addition of 100-200 μL bulk buffer to avoid having to move an external reference electrode between quadrants. To simplify scanning while maintaining multiplexing ability and keeping reaction volumes low, we designed an 8-electrode multiplexed chip with internally integrated Pt auxiliary and pseudo-reference electrodes.

In this work, we aimed to develop a polymerase assay with a rapid readout that could serve as an alternative to radioactivity-based assays. Furthermore, we hoped to detect differences in activity between oxidized and reduced Pol δ on this platform. To aid in these goals, we designed and tested a new multiplexed device optimized for protein experiments.

Materials and Methods

Design and Fabrication of an Integrated Multiplexed Device

Masks for gold and Pt components of a multiplexed device with internal reference and auxiliary electrodes were designed using AutoCAD software. Multiplexed chips were patterned and cleaved from wafers containing 9 chips each using clean room facilities in the Kavli Nanoscience Institute at Caltech. Boron-doped silicon wafers were oxidized in large batches of 25 in a Tystar furnace for 3 hours to generate SiO surfaces. Prior to patterning, all wafers were dehydrated by baking on a hot plate at 150 C for 10 minutes, after which the cooled wafers were primed for photoresist annealing by a 5-minute incubation with HMDS. S1813 positive photoresist (Shipley) was then applied uniformly across the surface using a photoresist spinner, and wafers were immediately transferred to a hot plate and baked for 1 minute at 115 ℃ to set the photoresist. Patterning was accomplished by exposure to UV light using the g-line of a Karl-Suss mask aligner using a mask containing the pattern for the Pt or Au electrodes; exposed photoresist was then removed with a 30 second bath in basic MF19 developer followed by a 1-minute water bath, yielding the desired electrode pattern.

After the pattern was set in photoresist, the wafers were stored in the clean room until metal deposition. Metal deposition was done with batches of 8 wafers (cleaned prior to this step in a plasma chamber with 150 sccm O_2 at 1 mtorr) in a CHA metal evaporator; specifically, 30 Å Ti was deposited first at 0.5 Å/s, after which 1000 Å of Pt or Au was deposited at 1 Å/s. Following metal deposition, metal on photoresist outside of the pattern was removed by a 60 C metal lift off step carried out in PG remover. The metal lift-off step took about 1 hour for ready removal of unwanted Pt, while Au took from 1-1.5 hours for a complete liftoff. Following this step, wafers were rinsed with acetone to physically remove any remaining photoresist/metal, and

then placed in subsequent baths of acetone, water, and isopropanol. This process was performed for the Pt pattern first and was then repeated for the Au pattern.

After metal application, the wafers were coated with insulating SU-8, a negative photoresist. Wafers were dehydrated and plasma cleaned prior to photoresist application. SU-8 2002 was applied on a photoresist spinner, and the wafers were then baked for 2 minutes at 95 C. Exposure was undertaken as before, but using the i-line UV wavelength on the mask aligner along with a mask set to expose only the Au and Pt surfaces to UV. Importantly, a post-exposure bake (2 minutes, 95 ℃) was required to prevent the complete removal of SU-8 during development. Wafers were developed in SU-8 developer for 80 seconds, then placed in an isopropanol bath for 1 minute and dried under an inert gas (either Ar or N_2). The pattern was examined under a microscope to confirm complete development and cleanliness of the surfaces, and the SU-8 pattern was then set in a 15 minute hard bake at 150 C. Finally, individual chips were cleaved from the wafers using a Dynatex scriber-breaker. Chips were stored in sealed plastic containers in a dessicator under vacuum until use.

DNA Synthesis and Purification

Initial characterization of the integrated chip platform used the following DNA substrates:

17-mer thiol: Thiol – 5′– GAC TGA GTA CGG TCG CA – 3′

17-mer MB′: 5′ – T_{MB}GC GAC CGT ACT CAG TC – 3′

17-mer unmodified complement: 5′ – TGC GAC CGT ACT CAG CT – 3′

For polymerase assays, a DNA construct consisting of the following three components served as a primed DNA replication substrate:

20-mer 3′ thiol: 5′– GCT GTC GTA CAG CTC AAT GC – 3′- thiol linker

38-mer MB′: 5′– T_{MB}AA CAG GTT GAT GCA TCG CGC TTC GGT GCT GCG TGT CT – 3′

49-mer complement: 5′ – GCA TTG AGC TGT ACG ACA GCA GAC ACG CAG CAC CGA AGC GCG ATG CAT C– 3′

Comparisons of this substrate with fully duplexed DNA used the same components but included a 58-mer complement with the following sequence:

58-mer complement: 5′ – GCA TTG AGC TGT ACG ACA GCA GAC ACG CAG CAC CGA AGC GCG ATG CAT CAA CCT GTT A – 3′

Unmodified DNA was purchased from IDT, and phosphoramidites for modified DNA were purchased from Glen Research. The 3′ thiol linker used for these reactions was 3-oxahexyl-1-thiol, a higher-yielding modification of the 6 carbon linker typically used for 5′ thiol modifications. All DNA synthesized in lab was made using an ABI DNA synthesizer employing standard phosphoramidite chemistry.

DNA oligomers were removed from CPG beads by 8-hour incubation in NH_4OH at 60 C, then rinsed, filtered, and dried on a speed vacuum. DNA was then dissolved in 600 μL 5 mM sodium phosphate buffer with 50 mM NaCl (pH 7.0) and purified with HPLC using a PLRPS column with a gradient of ACN and 50 mM NH_4OAc (5-75% ACN/95-25% NH_4OAc over 30 minutes). Collected DNA was then dried on a lyophilizer, and the 5′ DMT group was removed by 25-minute incubation in 250 μL of an 80% aqueous acetic acid solution followed by EtOH precipitation. DNA was again dried on a speed vacuum, and purified by another round of HPLC

(5 – 15% ACN/95 – 85% 50 mM NH_4Ac over 35 minutes). DNA strands containing a disulfide modifier were reduced using 100 mM DTT (45 minutes, RT), skipping the acetic acid step. DTT was removed on a NAP5 column, and the oligomers were HPLC purified.

MB′-modified DNA was prepared by activation of the MB′-acid to an NHS ester followed by coupling to 5′-amino-modified DNA synthesized. For activation, 0.022 mmol N-(carboxypropyl)methylene blue (MB′) was added to a scintillation vial and mixed with 0.045 mmol N,N'-dicyclohexylcarbodiimide (DCC) and 0.045 mol N-hydroxysuccinimide (NHS) in DMF. This reaction was stirred overnight at RT in the dark. The resultant NHS-ester was then dried under vacuum and re-dissolved in a minimal amount of DMSO, after which it was added in 10-fold excess to fully purified 5′-amino-modified DNA in $NaHCO_3$ and placed on a shaker overnight. The resultant MB′-modified DNA was isolated from unreacted MB′ on a NAP5 column and HPLC purified as described above.

Protein Expression and Purification

Yeast DNA polymerase δ 01 (exo^-) and PCNA and *E. coli* endonuclease III (EndoIII) were expressed and purified according to previously published protocols (*10, 12*).

^{3}H-Incorporation Assay for DNA Polymerase Activity

For independent confirmation of Pol δ activity, we used a standard DNA polymerase assay based on incorporation of ^{3}H-labeled dTTP into calf thymus DNA; the same assay was also used as to test activity on the MB′-modified DNA substrate used for electrochemistry (*13-15*). Reactions consisted of 200 nM Pol δ in the presence or absence of 2 μM PCNA was added to a reaction mix (50 μL final volume) consisting of 8 mM $MgCl_2$, 1 mM DTT, 25 mg/mL BSA, 2 mg/mL calf thymus DNA, 400 μM dATP, dCTP, and dGTP, 100 μM dTTP, and 800 cpm/pmol [5-methyl ^{3}H]-dTTP and 10% glycerol v/v in 50 mM Tris-HCl (pH 8.0). Reactions with Klenow

fragment exo$^-$ (New England Biolabs) were run alongside as a control. dNTPs and Klenow fragment were purchased from New England Biolabs (NEB), calf thymus DNA was purchased from Sigma Aldrich, and radiolabeled dTTP was from Perkin Elmer.

Reactions were run at 37 ℃ for 30 minutes, and quenched by addition of 100 μL chilled 100 mM sodium pyrophosphate and 1 mL 10% TCA in EtOH. Quenched reactions were left on ice for 10 minutes, and DNA was isolated by filtration on Whatman GFC fiberglass filters and rinsed twice with 2 mL 1 M HCl, once with 1 mL 50 mM sodium pyrophosphate and once with EtOH. Acid insoluble radioactivity (DNA)was then quantified using liquid scintillation counting in 0.4% 2,5-bis(2-(5-tert-butylbenzoxazolyl)) thiophene (BROT) in toluene.

Electrochemistry-Based DNA Polymerase Assay

All DNA SAMs were formed overnight at RT on integrated multiplex chips, and backfilled for 45 minutes with 1 mM 6-mercato-1-hexanol in 5 mM phosphate buffer with 50 mM NaCl and 5% glycerol v/v. Surfaces were then washed ten times in the same buffer. Initially, surfaces were scanned immediately after backfilling, but, in order to obtain more comparable monolayer structure before and after reaction, later surfaces were incubated at 37 ℃ for 30 minutes. This incubation led to more directly comparable signals before and after reaction, likely due to a final surface reorganization and removal of any unbound MB′ DNA. In all cases, initial scans were performed in the same buffer, then rinsed 5 times and scanned in TBP buffer (5 mM sodium phosphate, pH 7.0,50 mM NaCl, 4 mM spermidine, 4 mM $MgCl_2$, 0.5 mM EDTA, and 10% glycerol v/v). After scan rate dependence was taken in TBP buffer, the surface was rinsed 5 times in reaction buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10% glycerol v/v), scanned, and the process was repeated.

Reaction mixes (20 nM Pol δ 01, 2 μM PCNA, 8.0 mM $MgCl_2$, and 400 μM dNTPs in reaction buffer) were prepared on ice, and 20 μL was added to each quadrant of an integrated multiplexed device. Initially, surfaces were scanned with enzyme immediately after addition and after reaction to observe differences, but, as nothing critical occurred at these points, these scans were later eliminated to achieve more accurate reaction times. The surface was then placed in a 37 ℃ incubator or on a bench at RT for a set time, and the reaction mix was removed and the surface rinsed in activity buffer. Scan rate dependence was taken again, and the surface was finally rinsed and scanned in TBP buffer. For oxidation experiments, reactions were run at RT in a glove bag under a 5% H_2, 95% N_2 atmosphere. In all cases, DNA synthesis was quantified by subtraction of the charge under the MB′ reductive peak at 500 mV/s before reaction from the same peak after reaction.

Electrochemical Replication Product Analysis by HPLC

After the completion of all scans, the surface was incubated at 90 ℃ and copiously washed with TBP buffer to remove DNA. DNA from multiple surfaces was pooled and run through a PLRPS column using the . As controls, stock DNA representing unreplicated and full-length products were run prior to the surface samples to confirm that these substrates elute at distinct times.

Bulk electrolysis

To oxidize the polymerase, 100 nM Pol δ and 1000 nM PCNA were added to a gold rod electrode, and a potential of 0.412 V vs NHE was applied until the curve leveled off. The amount of protein oxidized was determined by integration of the curve, and subtraction of the same curve obtained with buffer alone.

Results and Discussion

Characterization of a Multiplexed Chip with Internal Reference/Counter Electrodes

In order to minimize solution volume while maintaining ease of scanning and multiplexing capacity, we designed and fabricated a multiplexed chip containing internally integrated Pt reference and auxiliary electrodes (Figure 7.1). To make room for the reference/auxiliary, this new device had only 8 electrodes versus the 16 of the standard multiplexed chip. To test the functionality of this new design, we performed CV experiments with a 17-mer MB′-modified DNA duplex. The sequence used is identical to one previously employed in general characterization of the standard device (*16*). To ensure that electrochemistry was similar under all conditions, DNA monolayers of two types were examined: high density (assembled with 100 mM $MgCl_2$) and low density (no $MgCl_2$) (*17, 18*). On LD monolayers, covalent MB′ in particular possesses more flexibility and signals primarily through the electrode surface, while HD monolayers enforce more rigid packing and facilitate DNA-mediated signaling (*19*).

Upon scanning, the same pattern seen previously is observed on integrated chip electrodes, and electron transfer kinetics as determined from Laviron analysis were equivalent (5 - 10 per second in both cases); this result is consistent with CT rates limited by tunneling through the alkane-thiol monolayer and agrees with previous results (*19*). DNA surface coverage was calculated by $Ru(NH_3)_6$ titration, which yielded about 20 pmol/cm^2 for low-density monolayers and 30 pmol/cm^2 for high density monolayer, as expected (*17, 18, 20*). As determined using a MB′ midpoint potential value of -75 mV vs NHE, potentials were shifted 100 mV negatively relative to Ag/AgCl, although this was consistent across all electrodes and multiple devices and

can be attributed solely to the use of a Pt pseudoreference. Critically, variability is remarkably low both within and between separate quadrants (Figure 7.2).

To confirm that protein electrochemistry was also unperturbed on these devices, we next repeated the CV experiments using 50 μM EndoIII in storage buffer (20 mM sodium phosphate, pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 20% glycerol v/v). For protein experiments, we used the same 17-mer DNA duplex from initial characterization, but with an unmodified complement to the thiol strand. As expected, EndoIII showed a reversible signal with the same electrochemical properties on both devices (Figure 7.3), confirming their utility for protein experiments (*12*).

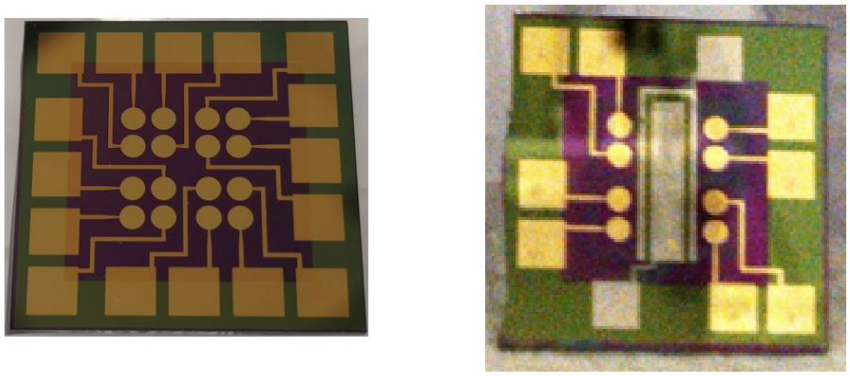


Figure 7.1 A standard 16-electrode multiplexed chip (left) and an 8-electrode chip with internal Pt reference and auxiliary electrodes (integrated chip; right). Both devices use a gold working electrode, although the area of each electrode is slightly lower on the integrated device in order to make room for the reference/auxiliary electrodes.

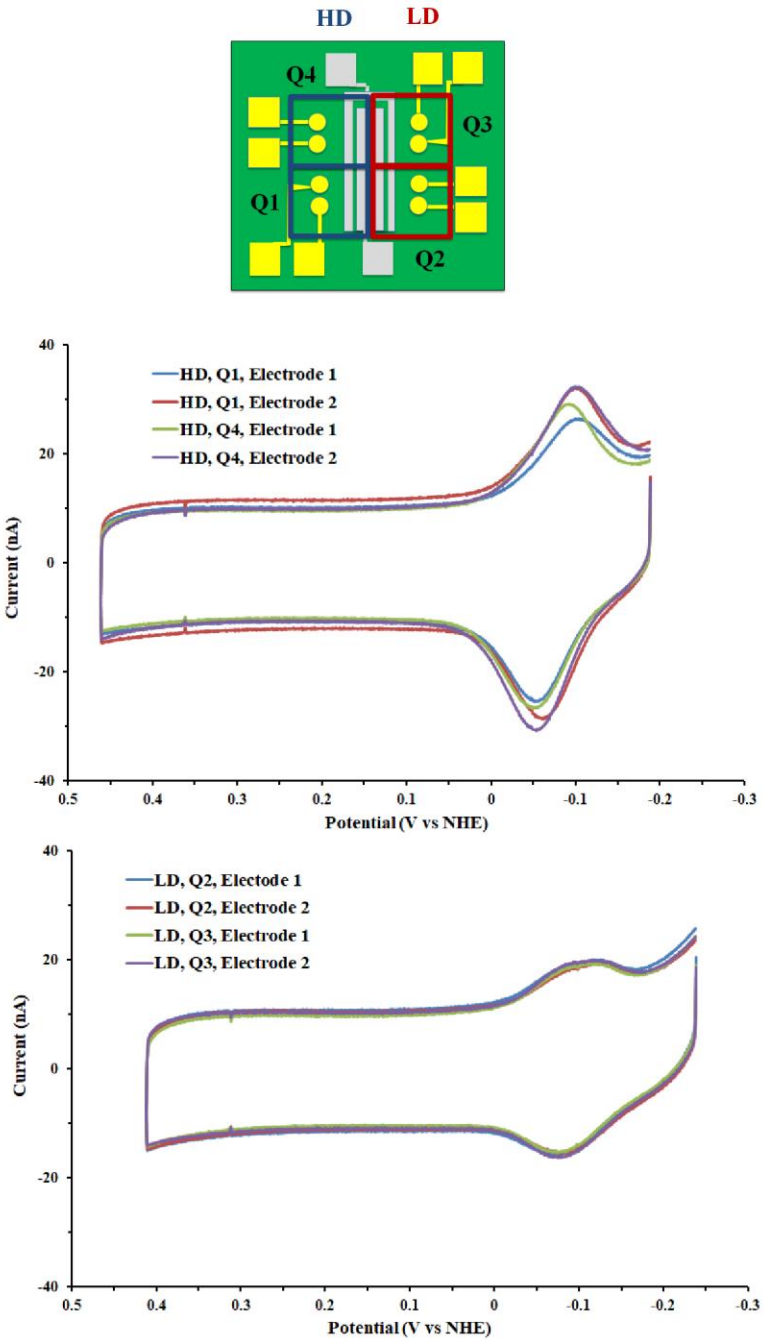


Figure 7.2 Electrochemistry of a 17-mer DNA duplex on an integrated chip. (top) DNA films on one half of the device were prepared in the presence of 100 mM $MgCl_2$ to generate high density (HD) surface packing, while the others were assembled with no $MgCl_2$ for standard low-density (LD) packing. (center) Cyclic voltammetry (CV) of MB′-modified DNA in quadrants one and two (both HD). (bottom) CV of MB′-modified DNA in quadrants one and two (both LD). As expected, signals are larger on HD films, and variability is minimal both within and between quadrants of a given type of DNA film. CV scans were taken at 100 mV/s scan rate, and experiments were performed in TBP buffer (5 mM sodium phosphate, pH 7.0, 50 mM NaCl, 4 mM spermidine, 4 mM $MgCl_2$, 0.05 mM EDTA, and 10% glycerol v/v).

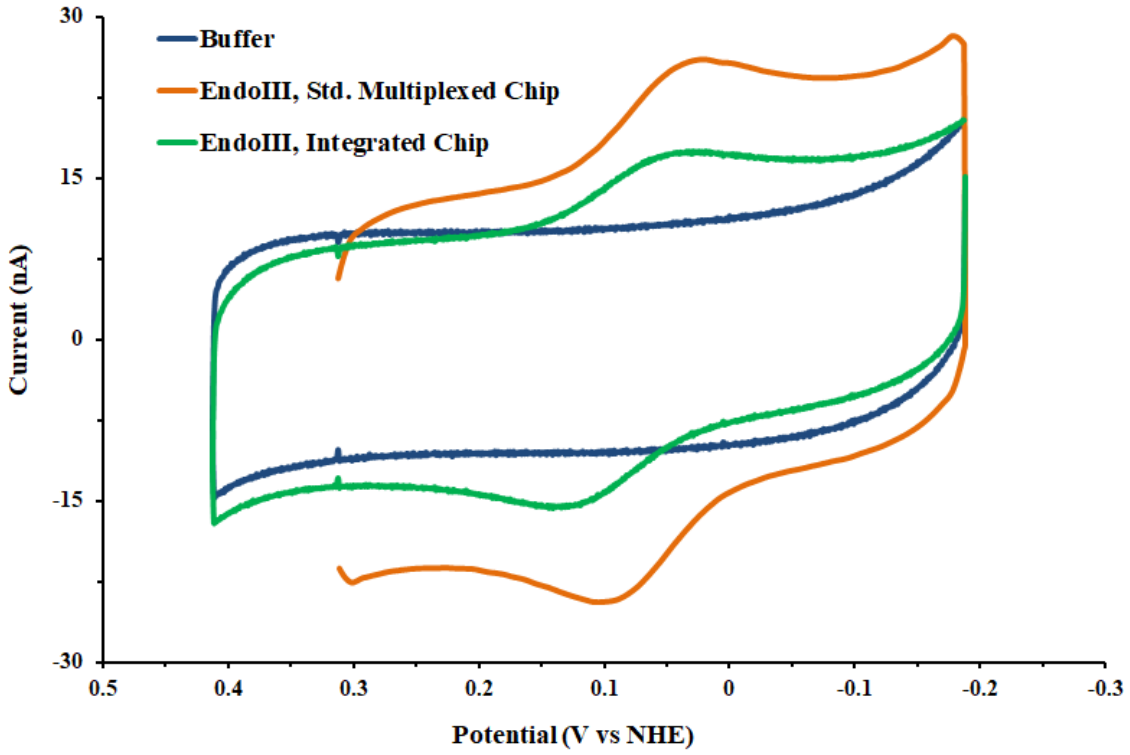


Figure 7.3 50 μM EndoIII on a standard multiplexed chip (orange) and an integrated chip (green) at 100 mV/s scan rate. Experiments were all performed in storage buffer (20 mM sodium phosphate, 150 mM NaCl, 0.5 mM EDTA, 20% glycerol v/v).

Electrochemical DNA Polymerase Assays Using a MB′ Probe

Once the integrated chip was fully tested, we proceeded to design an assay to report on DNA polymerase activity at the electrode surface. The general design of this assay relies on the making the distinction between surface-mediated and DNA-mediated MB′ signals (*19*). In our design, covalent MB′ on the end of a single stranded overhang would be expected to have high flexibility and give a surface-mediated signal characterized by rapid electron transfer rates. Once duplex extension occurred following incubation with Pol δ under replication conditions, this flexibility would be lost and intercalation into the base stack would be favored, resulting in slower, DNA-mediated signals (Scheme 7.1).

Before working with protein, it was first necessary to verify the occurrence of a difference in the signals from DNA with an overhang and fully duplexed DNA. To address this issue, we prepared a suitably large DNA polymerase substrate with MB′ appended to a modified dT by either a 6-carbon or 3-carbon linker. In addition, we tested both HD and LD monolayers to identify the optimal conditions, and ran electrochemistry in TBP buffer and polymerase reaction buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl). Intriguingly, a noticeable difference was observed, although only on HD monolayers in TBP buffer (Figure 7.4). Under these conditions, MB′ signals on overhang DNA overlap closely and have electron transfer rates of ~30 s-1, consistent with the probe ending up partially buried in the monolayer. In contrast, fully duplexed DNA shows a split reductive peak, with the lower potential peak being much larger and showing slower kinetics (~5-10 electrons s^{-1}), in agreement with previous studies of duplexed DNA with covalent MB′ (*19*).

Having confirmed that duplexed and overhang DNA can be distinguished, we next confirmed polymerase activity using a ^{3}H-dTTP incorporation assay. This assay was performed

with activated (partially DNase digested) calf thymus DNA using 200 nM Pol δ 01 (exo$^-$) in the presence or absence of 2 μM PCNA. Pol δ 01 was used to prevent exonucleolytic degradation of surface-bound DNA by the 3′ - 5′ proofreading site of the polymerase. Since PCNA stimulates processive DNA synthesis, which is more relevant to the *in vivo* situation, we hoped to see much greater ^{3}H incorporation with PCNA present. As expected, all polymerase reactions gave counts significantly above baseline, confirming that the enzyme was active (Figure 7.5). In addition, PCNA did stimulate synthesis to a quantifiable extent, although the differences were not as substantial as those seen on gels using much larger single-stranded substrates (*10, 11*). After verifying enzymatic activity, we repeated this assay with our MB′ DNA substrate (lacking a thiol in this case) to ensure that the probe does not interfere with polymerase activity. These assays were run in parallel with those containing calf thymus DNA for a more direct comparison. Fortuitously, Pol δ activity was equivalent or greater on MB′ DNA relative to calf thymus DNA (Figure 7.5).

Having performed all of the necessary characterization, we proceeded to try an electrochemical assay with Pol δ. Because the amount of DNA present on an electrode is much less than in our solution assays, we lowered the polymerase concentration to 20 nM for these assays to prevent saturation. Initial reactions were run for 30 minutes at 37 °C, after which sodium pyrophosphate and EDTA were added and the surface washed to remove enzyme. No polymerase was added to half of the device, although the treatment was otherwise identical. As MB′ coupling is most efficient in TBP buffer, we scanned in TBP prior to reaction, washed in polymerase reaction buffer prior to enzyme incubation, and performed a final set of washes in TBP buffer prior to a set of post-reaction scans. Remarkably, surfaces with polymerase added showed the same split reductive signal seen in fully duplexed DNA, while those which lacked

polymerase did not change after the reaction (Figure 7.6). It should be noted that signals were smaller in all cases following the reaction and extensive washing, but losses were not great. In any case, the no-enzyme control and reaction surfaces changed by about the same extent, so the effect of losses could be easily accounted for. To confirm that the differences were caused by polymerase reaction rather than other factors such as surface passivation by quenched protein mixes, we repeated reactions in the absence of dNTPs. For these experiments, polymerase was added to each quadrant, while half of the chip lacked dNTPs. Following reaction, complete mixes showed the same MB′ signal change observed before, while no appreciable differences occurred in the absence of dNTPs (Figure 7.6).

As a final control, we removed DNA from post-reaction surfaces by heating to 90 ℃ and running the resultant DNA through an HPLC column. For comparison, we separately ran an "unreplicated" 38-mer/49-mer duplex and a full-length 38-mer/58-mer duplex (not from electrodes) to confirm that these strands eluted at different times. On a PLRPS column, the unreplicated control substrate elutes around 24 minutes, while the full-length duplex elutes around 32 minutes (Figure 7.7). DNA removed from electrodes elutes around 30 minutes, with minor peaks at 24 and 32 minutes, consistent with a significant degree of replication (Figure 7.7). The slightly earlier elution relative to fully duplexed DNA may be due to failure to incorporate dNTPs at or near the MB′-modified dT; nonetheless, all of our controls are consistent with the occurrence of DNA synthesis at the electrode surface.

Once we were confident that activity could be observed, we attempted to see if we could use this assay to distinguish basal polymerase activity from the more biologically relevant PCNA-mediated activity. In standard gel assays, PCNA is required for the effective synthesis of substrates larger than several nucleotides. When we tested both conditions on our

electrochemical platform, however, no differences were apparent (Figure 7.8). Although disappointing, this was not entirely surprising. Even in solution assays, the effect of PCNA was not as dramatic as that seen on gels, most likely as a consequence of the single-stranded DNA region being so short relative to a typical substrate (*10, 11*). That no difference occurs on an electrode may be due to the fact that HD DNA on the electrode is less accessible to PCNA-Pol δ complexes. Furthermore, with only one DNA end available for the clamp to slide on, PCNA may have been even less stimulatory without the ability to rapidly slide from substrate to substrate. Overall, basal DNA synthesis appears to dominate on electrodes, which is not ideal but still does not preclude further analysis of polymerase activity.

With these initial results in hand, we next aimed to generate a time course for reaction to enable quantification of reduced vs oxidized Pol δ activity in future experiments. Because MB′ signals are most readily quantified in TBP buffer, reaction progress could not be observed in real time. To generate a time course, we thus prepared chips with three quadrants for reaction and one control, and incubated at 37 ℃ for 0, 5, 10, 15, 25, or 30 minutes prior to quenching. Quantification was performed by taking the area under half of the second reductive peak and multiplying by 2; background was then subtracted using a control quadrant with no polymerase added. Finally, the % increase in secondary peak charge was plotted over time as an indicator of reaction progress. After all time points had been assessed, it became clear that reactions were largely complete by 10 minutes at 37 ℃ (Figure 7.9).

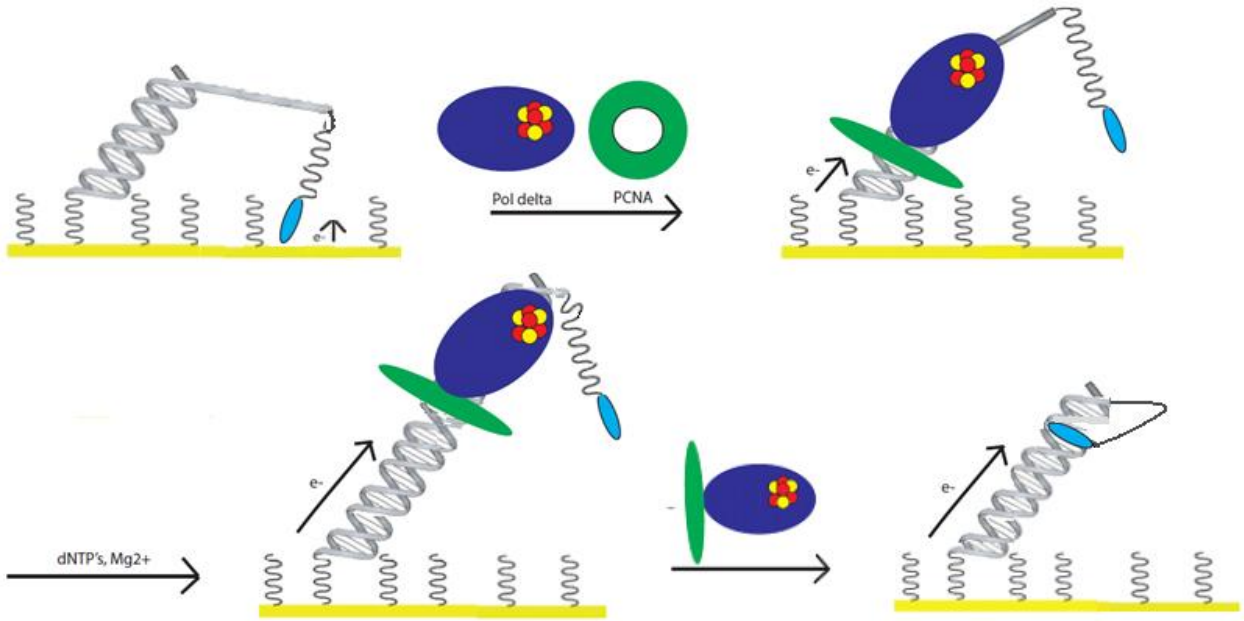
Convinced that we were seeing polymerase activity on an electrode that could be quantified over time, we attempted to compare oxidized and reduced Pol δ activity levels. Initially, oxidation by bulk electrolysis was attempted in one quadrant containing unmodified DNA on the same integrated device as reactions were performed; however, electrolysis yields

were inconsistent and often low (as low as 5% in some cases), likely due to low chances of oxidizing an entire sample from one of two electrodes. A further problem arose when it became clear that under anaerobic conditions, the Pt reference potential wandered, moving positive hundreds of mV after 1-2 hours; with MB′ present, this could be corrected, but it was nonetheless an undesirable result. Overall, it was clear that integrated devices (or any multiplexed electrode) are unsuitable for bulk electrolysis, so we moved forward using electrolysis on a separate, single gold rod electrode as described in *11*. Yields were much improved in this case, and DNA synthesis by oxidized Pol δ was found to be lower than untreated samples by the electrochemical assay, as expected. However, a time course failed to reveal a consistent pattern in activity differences (Figure 7.10).

The discrepancy between these results and those seen by conventional gel-based assays is almost certainly attributable to the limitations of the electrochemical system. Specifically, the infeasibility of using longer single-stranded overhangs precludes observations of biologically relevant, PCNA-mediated processive DNA synthesis. This is particularly problematic, as oxidation is more limiting to processivity than the ability to synthesize DNA at all, and 9 nucleotides is likely near the limit at which a difference can be detected. Second, DNA surface coverage can vary between chips and is not easily controlled, meaning that the amount of DNA in each assay differs, unlike in solution reactions which use known concentrations of all components. Third, what DNA is present is spread between two electrodes that do not entirely cover a quadrant, making protein contact more random than would be the case in solution. Finally, multiplexed chips cannot readily be incubated at higher temperatures in a glove bag, so reactions are limited by slow diffusion. Overall, the fact that a difference was consistently seen is

remarkable, but the unfavorable features of this platform far outweigh its potential benefits for monitoring activity differences of this type.

In summary, we have designed a multiplexed device suitable for working with low volumes and demonstrated that it is possible to observe polymerase activity on this surface. However, the Pt pseudoreference requires close monitoring, as the potential is prone to wander especially under anaerobic conditions. Furthermore, limits in the size and amount of DNA that can be distributed on a surface are not optimal for the study of highly processive polymerases, as illustrated by the ineffectiveness of PCNA in stimulating Pol δ on the electrode while having the opposite effect on the same substrate in solution. These limitations combine to make this assay design far less suitable for studying the effects of oxidation on [4Fe4S] DNA polymerases than we initially hoped. However, while this assay did not meet expectations, we were able to monitor non-processive DNA synthesis over time; only efforts to see finer details caused problems. Thus, while processive polymerases are best studied by traditional assays, a similar electrochemical assay might be useful for other proteins like telomerase, which do not synthesize hundreds of nucleotides at a time. Finally, it is also possible that this or a similar assay could be adapted to study DNA repair proteins, which require only a target site and are not limited by a need for extensive quantities of single-stranded DNA.



Scheme 7.1 Electrochemical DNA polymerase activity assay. Before replication, DNA with a MB′ reporter tethered to the 5′ end of a single-stranded overhang signals primarily through the electrode surface, which is characterized by rapid electron transfer kinetics. The addition of polymerase and/or the sliding clamp PCNA followed by dNTPs is expected to allow extension of the primed DNA. Once the proteins are removed, the fully duplexed DNA now favors MB′ intercalation into the base stack, leading to slower electron transfer kinetics than those seen through the surface.

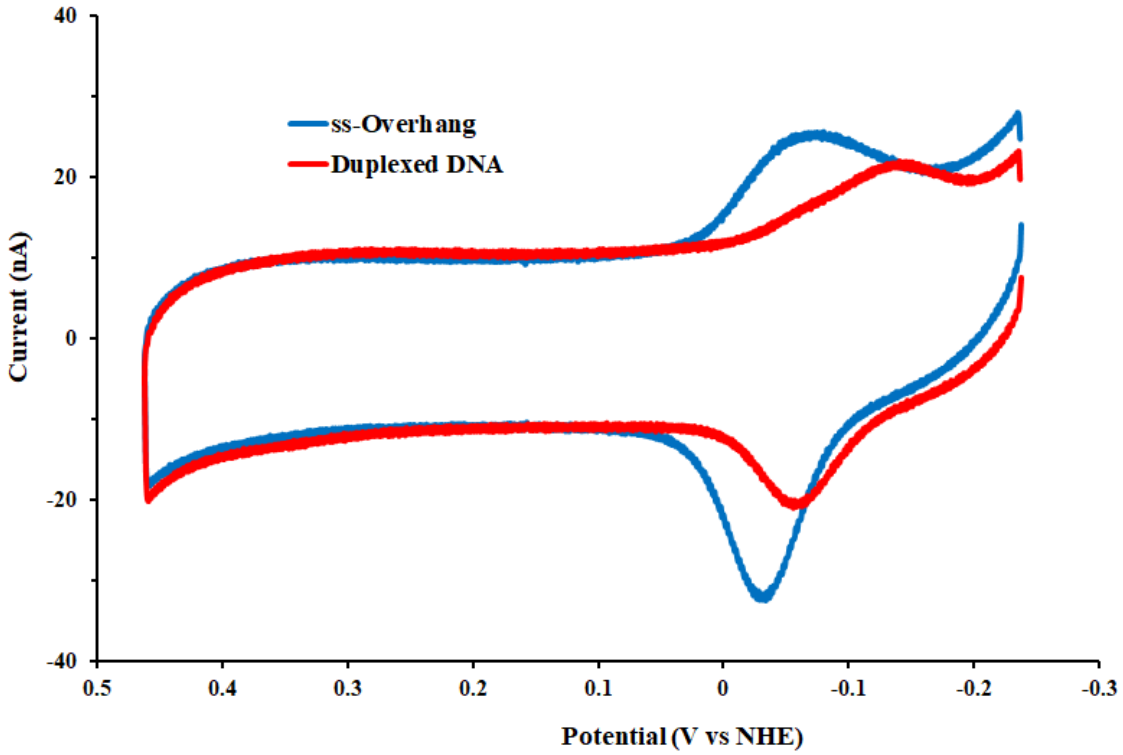


Figure 7.4 CV of MB′-modified Pol δ substrates. DNA with a single-stranded overhang (blue) exhibits a single reductive peak with low peak splitting and rapid electron transfer rates. In contrast, fully duplexed DNA (red) shows a split reductive peak, with the primary peak showing marked splitting and slower electron transfer rates. Scans were taken at a scan rate of 100 mV/s in TBP buffer (5 mM sodium phosphate, pH 7.0, 50 mM NaCl, 4 mM spermidine, 4 mM $MgCl_2$, 0.5 mM EDTA, 10% glycerol v/v).

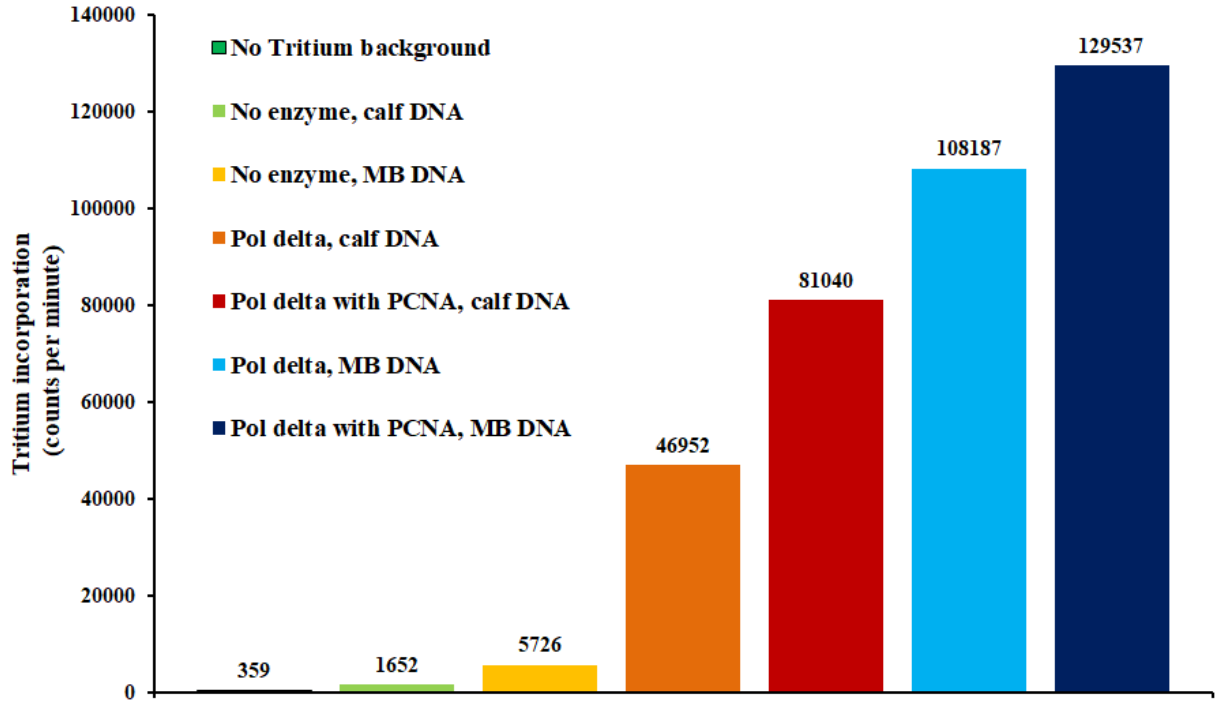


Figure 7.5 Data from a representative ^{3}H incorporation assay. ^{3}H-labeled dTTP was used to record DNA synthesis, and Pol δ activity with (red) and without PCNA (orange) on a standard calf thymus DNA substrate was compared with the same MB′ substrate used for electrochemical activity assays, but lacking a thiol (blue). Significant activity was observed on both, showing that the polymerase was active and unimpeded by the presence of MB′ under any conditions. Assays included 200 nM Pol δ 01, 2.0 μM PCNA, 8.0 mM $MgCl_2$, 25 mg/mL BSA, 2.0 mg/mL calf thymus DNA or 6.0 μM MB′ DNA, 400 μM dATP, dCTP, and dGTP, 100 μM dTTP, 800 cpm/pmol [5-methyl ^{3}H]-dTTP and 10% glycerol v/v in 50 mM Tris-HCl (pH 8.0)

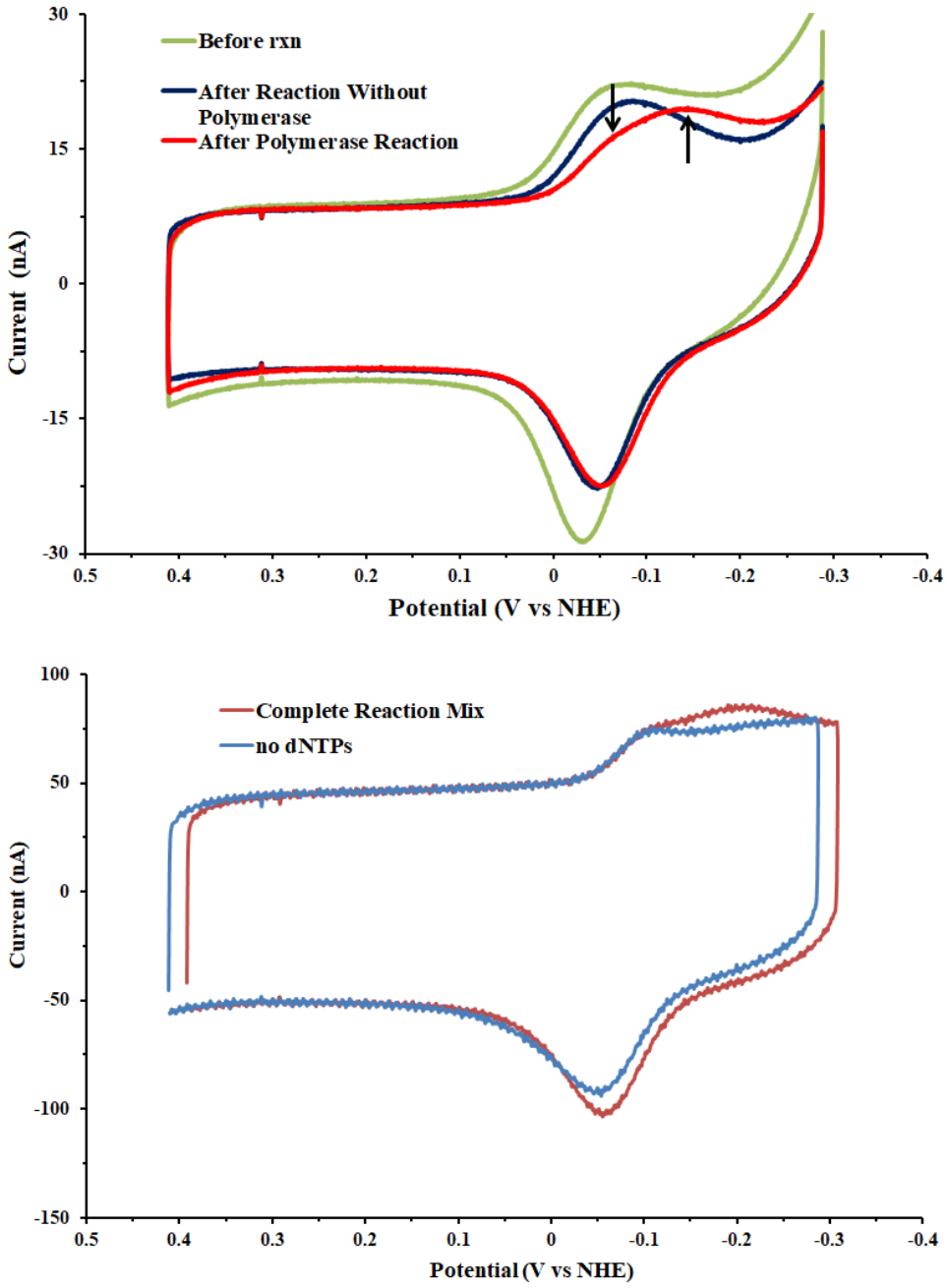


Figure 7.6 Polymerase reaction at an electrode. (top) Before reaction, CVs all match those observed previously for DNA with a single-stranded overhang (green), while a second reductive peak develops after incubation with 20 nM Pol δ (red), consistent with the formation of fully duplexed DNA. If no polymerase is added, the overall CV signal form does not change (blue). (bottom) Excluding dNTPs from the reaction mix does not produce a secondary peak, supporting the assertion that DNA synthesis is being observed at the electrode surface. Reactions were performed with 20 nM Pol δ 01, 2.0 μM PCNA, 8.0 mM $MgCl_2$, 400 μM dNTPs, and 10% glycerol v/v in 50 mM Tris-HCl, pH 8.0. Scans shown were taken at 100 mV/s (top) or 500 mV/s (bottom) in TBP buffer (5 mM sodium phosphate, pH 7.0, 50 mM NaCl, 4 mM spermidine, 4 mM $MgCl_2$, 0.5 mM EDTA, 10% glycerol v/v).

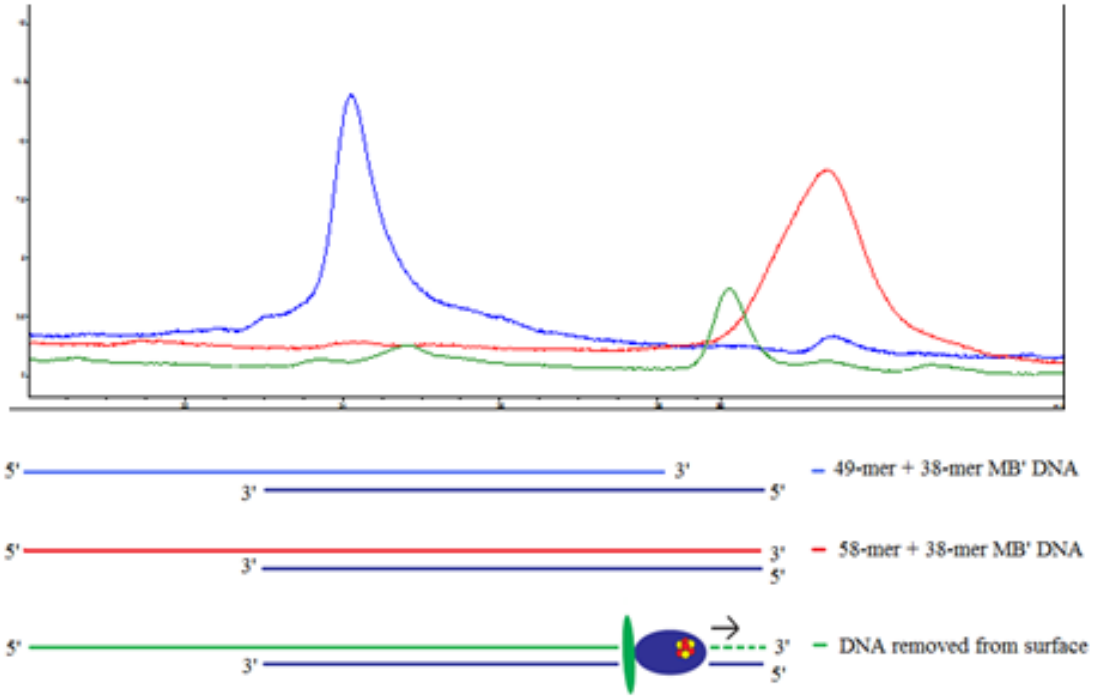


Figure 7.7 HPLC traces from DNA removed from an electrode by heating. DNA with an intact overhang, representing unreplicated DNA (blue trace), and fully-duplexed DNA (red trace) were run as controls. The DNA from the electrode (green trace) eluted at about the same time as the full-length DNA, suggesting that it had been extended by the polymerase. HPLC gradient was 5 – 15% ACN/95 – 85% 50 mM NH_4Ac over 35 minutes at a flow rate of 2 mL/min.

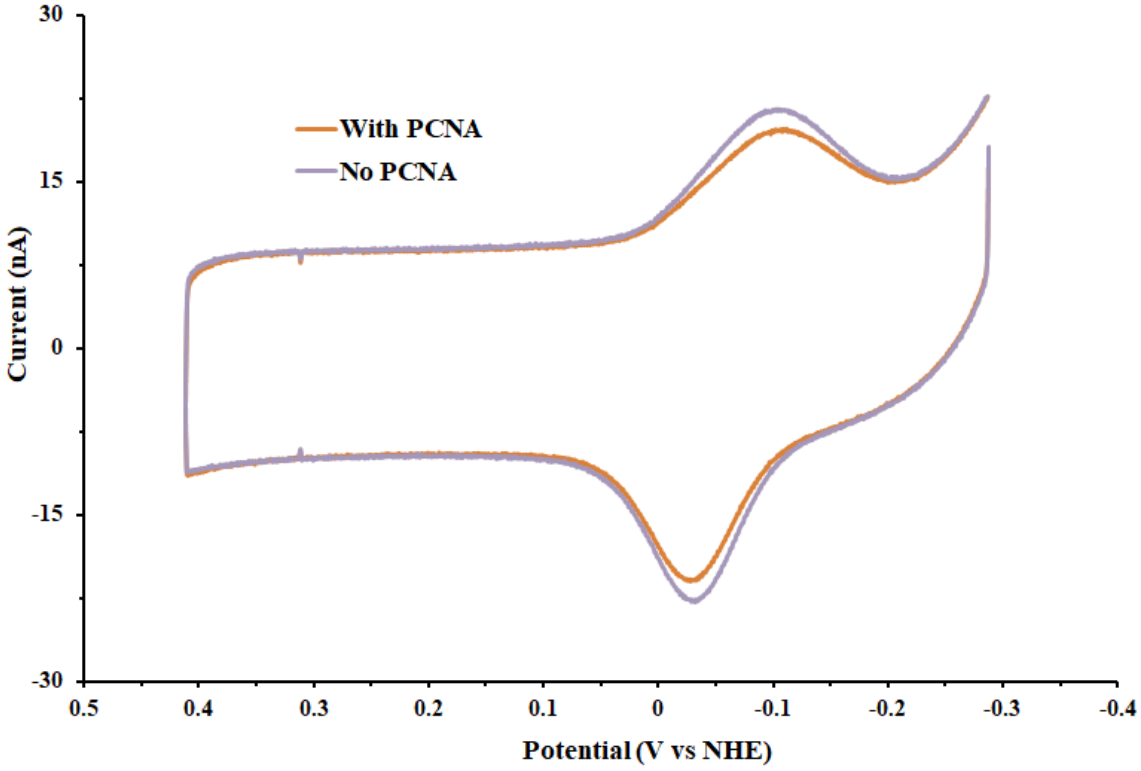


Figure 7.8 Polymerase reactions at a DNA-modified electrode are unaffected by PCNA. Conditions were as follows: 20 nM Pol δ 01, 0 or 2.0 μM PCNA, 8.0 mM $MgCl_2$, 400 μM dATP, dCTP, dGTP and dTTP, and 10% glycerol v/v in 50 mM Tris-HCl (pH 8.0)

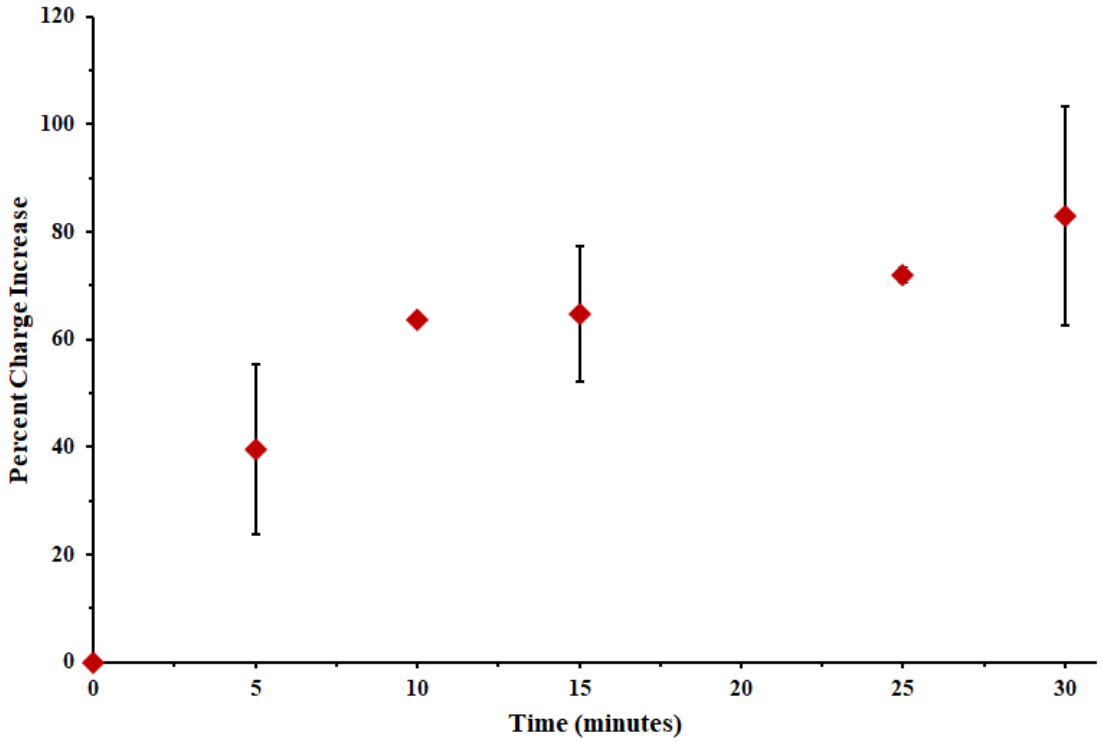


Figure 7.9 Time course for Pol δ replication reaction at 37 ℃ on an electrode. The percent of MB′ secondary reductive peak charge increase at 500 mV/s scan rate is plotted and normalized for surface coverage. Reaction is largely complete by 10 minutes in the presence of 20 nM Pol δ 01. Reaction conditions were as follows: 20 nM Pol δ 01, 2.0 μM PCNA, 8.0 mM $MgCl_2$, 400 μM dATP, dCTP, dGTP, and dTTP and 10% glycerol v/v in 50 mM Tris-HCl (pH 8.0).

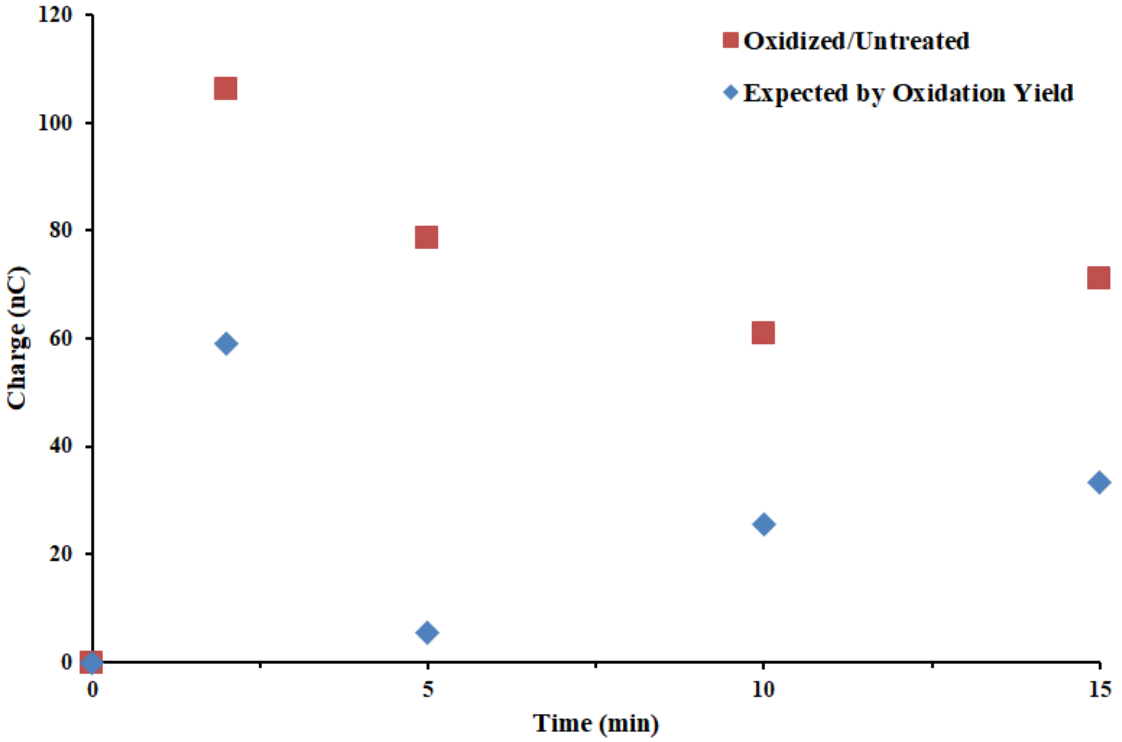


Figure 7.10 Activity following Pol δ oxidation by bulk electrolysis. The percent of untreated activity is shown in red, while blue denotes the amount of synthesis expected based on oxidation yield. While a difference is apparent at later time points, it is not as large as that seen in gel-based assays in which processive, PCNA-mediated synthesis is dominant. Overall, the use of a short single-stranded region and likely saturation by Pol δ make this type of assay unsuited to quantification on an electrode. Reaction conditions were as follows: 20 nM Pol δ 01, 2.0 μM PCNA, 8.0 mM $MgCl_2$, 400 μM dATP, dCTP, dGTP, and dTTP and 10% glycerol v/v in 50 mM Tris-HCl (pH 8.0).

References

1) Drummond, T.G..; Hill, M.G.; Barton, J.K. Electrochemical DNA Sensors. *Nat.Biotech.* **2003**, *21*, 1192 – 1199.

2) Furst, A.L.; Hill, M.G.; Barton, J.K. DNA-Modified Electrodes Fabricated Using Copper-Free Click Chemistry for Enhance Protein Detection. *Langmuir*, **2013**, *29*, 16141 – 16149.

3) Muren, N.B.; Barton, J.K. Electochemical Assay for the Signal-On Detection of Human DNA Methyltransferase Activity. *J. Am. Chem. Soc.* **2013**, *135*, 16632 – 16640.

4) Furst, A.L.; Barton, J.K. DNA Electrochemistry shows DNMT1 Hyperactivity in Colorectal Tumors. *Chem. & Biol.* **2015**, *22*, 938 – 945.

5) Arnold, A.R.; Grodick, M.A.; Barton, J.K. DNA-Mediated Charge Transport: from Chemical Principles to the Cell. *Cell Chem. Biol.* **2016**, *23*, 183 – 197.

6) DeRosa, M.C.; Sancar, A.; Barton, J.K. Electrochemically Monitoring DNA Repair by Photolyase. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10788 – 10792.

7) Mui, T.P.; Fuss, J.O.; Ishida, J.P.; Tainer, J.A.; Barton, J.K. ATP-Stimulated, DNA-Mediated Redox Signaling by XPD, a Repair and Transcription Helicase. *J. Am. Chem. Soc.* **2011**, *133*, 16378 – 16381.

8) Grodick, M.A.; Segal, H.M.; Zwang, T.J.; Barton, J.K. DNA-Mediated Signaling is Necessary for Genomic Integrity. *J. Am. Chem. Soc.* **2014**, *136*, 6470 – 6478.

9) Bartels, P.L.; O'Brien, E.; Barton, J.K. DNA Signaling by Iron-Sulfur Cluster Proteins. pp. 405 - 423 in *Iron-Sulfur Clusters in Chemistry and Biology, 2nd ed.* (ed. Rouault, T.), **2017**, *DeGruyer, Berlin/Boston.*

10) Netz, D.J.A; Stith, C.M.; Stumpfig, M.; Kopf, G.; Vogel, D.; Genau, H.M.; Stodola, J.L.; Lill, R.; Burgers, P.M.J.; Pierik, A.J. Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat. Chem. Biol*. **2012**, *8*, 125 – 132.

11) Bartels, P.L.; Stodola, J.L.; Burgers, P.M.J.; Barton, J.K. A Redox Role for the [4Fe4S] Cluster of Yeast DNA Polymerase δ. *J. Am. Chem. Soc.* **2017**, *139*, 18339 – 18348.

12) Pheeney, C.G.; Arnold, A.R.; Grodick, M.A.; Barton, J.K. Multiplexed Electrochemistry of DNA-Bound Metalloproteins. *J. Am. Chem. Soc.* **2013**, *135*, 11869 – 11878.

13) Chang, L.M.S. DNA Polymerases from Baker's Yeast. *J. Biol. Chem.* **1977**, *252*, 1873 – 1880.

14) Plevani, P.; Chang, L.M.S. Initiation of Enzymatic DNA Synthesis by Yeast RNA Polymerase I. *Biochemistry* **1978**, *17*, 2530 – 2536.

15) Bauer, G.A.; Heller, H.M.; Burgers, P.M.J. DNA Polymerase III from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **1988**, *263*, 917 – 924.

16) Slinker, J.D.; Muren, N.B.; Renfrew, S.E.; Barton, J.K. DNA charge transport over 34 nm. *Nat. Chem.* **2011**, *3*, 230 – 235.

17) Boon, E.M.; Salas, J. E.; Barton, J. K. An Electrical Probe of Protein-DNA Interactions On DNA-Modified Surfaces. *Nat. Biotechnol.* **2002**, *20*, 282–286.

18) Kelley, S. O.; Barton, J. K.; Jackson, N. M.; Hill, M. G. Electrochemistry of Methylene Blue bound to a DNA-Modified Electrode. *Bioconjug. Chem.* **1997**, *8*, 31–37.

19) Pheeney, C.G.; Barton, J.K. DNA Electrochemistry with Tethered Methylene Blue. *Langmuir*, **2012**, *28*, 7063 – 7070.

20) Barton, J.K.; Bartels, P.L.; Deng, Y.; O'Brien, E. Electrical Probes of DNA-Binding Proteins. *Methods in Enzymology* **2017**, *591*, 355 – 414.

Chapter 8

Investigating Charge Transport Through RPA-Bound Single-Stranded DNA Using Covalent Photooxidants

This work was part of a collaboration with A. Ehlinger and W. Chazin at Vanderbilt University.

P. Bartels performed all DNA preparation and purification, and carried out all photooxidation experiments and associated analysis. A. Ehlinger expressed and purified RPA.

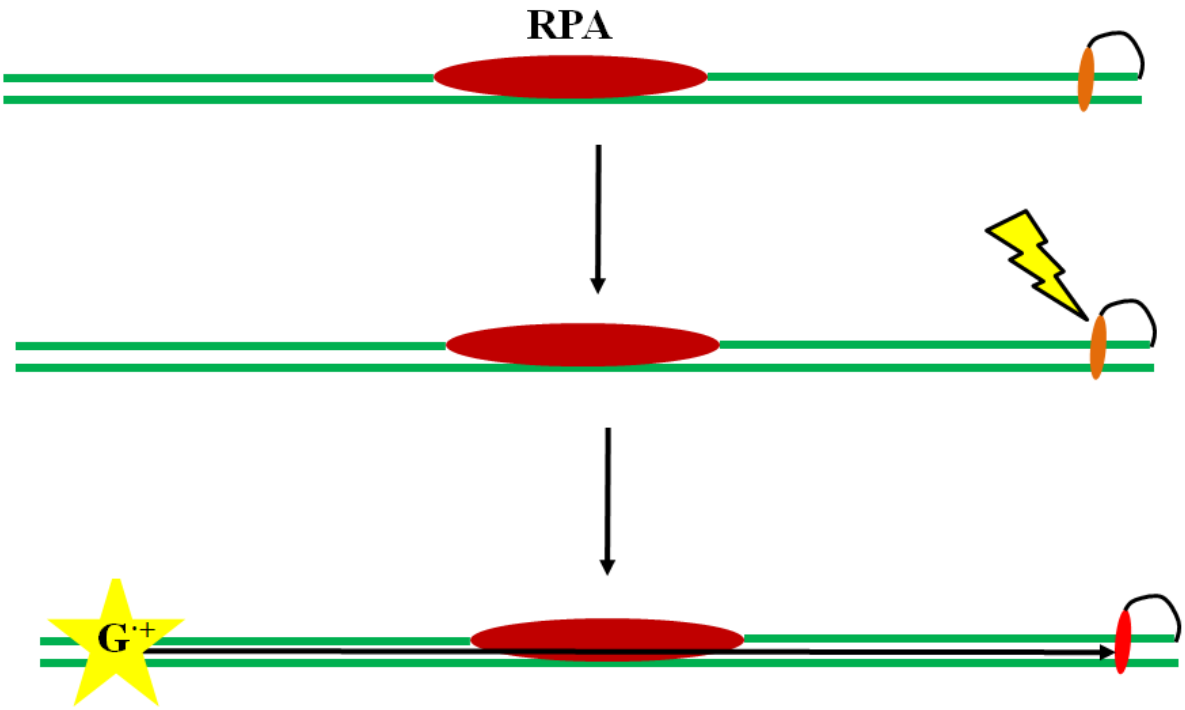
Introduction

Eukaryotic DNA replication involves a host of proteins that must coordinate their activity, both to accomplish their task in general and to respond to DNA damage and other setbacks collectively referred to as replication stress (*1, 2*). Of interest with respect to the stress response, several eukaryotic DNA replication and repair proteins, specifically DNA primase, DNA polymerase (Pol) δ, and MUTYH, have been shown to contain redox-active [4Fe4S] clusters (*3-5*). While redox-activity has not yet been demonstrated, [4Fe4S] clusters are also present in Pol ε, the helicase-nuclease Dna2, the translesion synthesis polymerase Pol ζ, and the endonuclease III homologues Ntg2/NTHL1 (yeast and human names, respectively) (*6-10*). Among those proteins previously studied, all were shown to be capable of long-range DNA-mediated charge transport (DNA CT) (*3, 4*). DNA CT, in which electrons and holes travel rapidly over long molecular distances through the π-stacked DNA base pairs, exhibits an acute sensitivity to even slight perturbations in base pair stacking and is thus an excellent reporter of damage (*11*). Indeed, electrochemical studies have shown dramatic CT attenuation by a wide variety of different lesions (*12*). DNA CT, however, is a phenomenon exclusive to duplexed DNA, while DNA replication involves a significant amount of single-stranded DNA, limiting the range of communication between redox-active proteins in this context.

At the replication fork, unwound single stranded DNA is coated with the single-stranded binding protein replication protein A (RPA) (*1*). RPA is a heterotrimeric complex consisting of three subunits: RPA70, RPA32, and RPA14 (*13*). Among these subunits are four DNA binding domains, which act sequentially in three binding modes to enhance binding affinity. These are an 8-10 nucleotide (nt) mode, a 25 nt mode, and an intermediate 12-22 nt binding mode (*13*). Crystal structures of the DNA binding domains show stacking interactions between aromatic

residues in RPA and the bases of single-stranded DNA, raising the intriguing possibility of CT through RPA-coated DNA during replication (*14*). Indeed, aromatic residues have been shown to couple with duplexed DNA in HhaI, ensuring that such a prospect is feasible (*15*). Notably, a structure of the entire complex bound to DNA shows a U-shaped bend that may be problematic for CT (*16*). However, the bend is not sharp, and CT can certainly occur through wrapped DNA as long as local base stacking remains intact (*17*). Furthermore, crystal structures show only a single conformation, and the situation in solution is likely to be far more dynamic. In any case, even if CT were not possible in this binding mode, the shorter modes remain of interest. If it can occur, CT through RPA-bound DNA would dramatically expand the possible communication pathways in DNA replication, and is thus worth investigating.

To test the possibility of CT through RPA-bound DNA, we have designed a system consisting of a DNA duplex with a covalent photooxidant on one end and containing an RPA-sized gap in the middle (Scheme 8.1). Taking advantage of the propensity of damage to localize on guanine multiplets (guanine has the highest oxidation potential of the DNA bases) (*18*), we aimed to evaluate damage on DNA irradiated in the presence and absence of RPA. These experiments are ongoing and represent a first step toward addressing the possibility of CT through single-stranded DNA at the replication fork, and future efforts should shed light on this intriguing possibility.



Scheme 8.1 Photooxidation scheme to test CT through RPA-coated ssDNA. In these experiments, RPA binds a single-stranded gap in a DNA substrate containing a photooxidant (orange) at one end. Irradiation will excite the photooxidant and provide a driving force for guanine oxidation. If CT can occur through RPA-bound DNA, guanine radical will form on the other end of the gap, resulting in damage products that can be cleaved by piperidine. Appending a 5′ ^{32}P label to this strand will allow cleavage products to be visualized.

Materials and Methods

Protein Expression and Purification

Heterotrimeric yeast RPA was overexpressed and purified as previously described (*19*).

DNA Synthesis and Purification

Unmodified DNA oligomers were purchased from IDT and purified by HPLC. DNA for conjugation to anthraquinone (AQ) or rhodium photooxidants was prepared on an Applied Biosystems automated DNA synthesizer. For AQ conjugation, 3′ C3 amino-modified CPG beads were purchased from Glen Research (3′-PT-Amino-Modifier C3 CPG). After synthesis, DNA was deprotected with 28-30% ammonium hydroxide (Sigma-Aldrich) (8 hrs, 65 ℃), filtered, and purified by HPLC (Agilent PLRPS column, 5-75% ACN/95-25% 50 mM ammonium acetate gradient over 30 minutes at a flow rate of 2 mL/min). The terminal DMT group was removed by 30′ RT incubation in 80% aqueous acetic acid, and the DNA was HPLC purified a second time (PLRPS column, 5-15% ACN/95-85% 50 mM ammonium acetate over 35 minutes at a 2 mL/min flow rate). DNA was then desalted by standard EtOH precipitation.

5′ ^{32}P labeling was achieved by incubating 50 pmol freshly prepared single-stranded DNA with T4 polynucleotide kinase (PNK; New England Biolabs) and [γ-^{32}P] ATP (Perkin Elmer) for 30 minutes at 37 ℃. PNK was heat inactivated (10 minutes, 85 ℃) and DNA was isolated from unreacted radiolabeled ATP by running samples through MicroBioSpin6 columns (BioRad) according to product instructions.

Once all DNA oligomers were prepared, the appropriate strands were quantified by UV-visible spectroscopy using molar extinction coefficients from IDT and annealed in equimolar ratios (5 minutes at 95 ℃ followed by cooling to RT over 1.5-2 hours) in storage buffer (5 mM sodium phosphate, pH 7.0, 50 mM NaCl). DNA substrate design consisted of 1 long strand

annealed to two complements with an RPA-appropriate gap in between. The DNA strand on one side of the gap contained a 5′ ^{32}P label and a guanine doublet to serve as a hole trap, while the strand on the other side contained either an AQ or rhodium photooxidant covalently tethered to the 5′ end. For comparison with RPA-bound DNA, a DNA sequence complementary to the gap was included as a control. The DNA sequences for photooxidation experiments were as follows:

AQ DNA

^{32}P 29-mer: ^{32}P - 5′ - TGT ACG ACA GGT GTC ATG CTA GCA TCA TA - 3′

16-mer AQ: 5′ - ACA CTA TAG CTC AGA G - 3′ - AQ

55-mer: 5′ - CTC TGA GCT ATA GTG TGA CGT TCG AGA TCA CGT CAT ATG ATG CTA GCA T - 3′

15-mer: 5′ - GAC ACC TGT CGT ACA - 3′

40-mer: 5′ - CTC TGA GCT ATA GTG TAC GTC GCA CAT ATG ATG CTA GCA T - 3′

10-mer gap: 5′ - TGT GCG ACG T - 3′

25-mer gap: 5′ - TGA CGT CAA CGT GAT CTC GAA CGT C - 3′

Rh DNA

Rh 16-mer: Rh - 5′ - ACA CTA TAG CTC AGA C - 3′

^{32}P 15-mer: ^{32}P - 5′ - TGT AGG ATG CAG TCG - 3′

56-mer: 5′ - CGA CTG CAT CCT ACA ACT GCA GTT GCA CTA GAG CTT GCA GCT CTG AGC TAT AGT GT - 3′

25-mer gap: 5′ - CTG CAA GCT CTA GTG CAA CTG CAG T - 3′

Coupling Anthraquinone 2-Carboxylic Acid (AQ) to DNA

To couple AQ to DNA, anthraquinone 2-carboxylic acid (Sigma-Aldrich, 0.022 mmol) was activated to the NHS ester by dissolving in DMF (~500 μL) and stirring overnight in the

presence of N,N′-Dicyclohexylcarbodiimide (DCC; Sigma-Aldrich, 0.045 mmol) and N-hydroxysuccinimide (NHS; Sigma-Aldrich, 0.045 mmol). The reaction was then dried and the components redissolved in a minimal amount of DMSO (~600 μL). Dried amino-modified DNA was dissolved in 200 μL 100 mM sodium bicarbonate and 50 μL of activated AQ NHS ester was added to each tube and gently shaken in the dark overnight. DNA was isolated by NAP-5 column (GE Healthcare) and HPLC purified a final time using the same method as the second purification. MALDI-TOF mass spectrometry was used to confirm the mass of the oligomer.

Coupling $[Rh(phi)_2bpy']^{3+}$ to DNA

For rhodium conjugation, DNA was prepared on the synthesizer using CPG beads with 2000 Å pore size (*20*). The terminal DMT group was cleaved on the column as the final step of the synthesis to expose the 3′ hydroxyl group for further reactions. Once synthesized, the CPG beads were added directly to a 2.5 cm x 10 cm glass cylinder with a glass frit and stopcock. The cylinder was sealed from above and connected to an aspirator, and the beads were washed 3x with 3 mL dry 1,4 dioxane (Sigma-Aldrich). 0.308 mmol (50 mg) carbonyldiimidazole (CDI; Sigma-Aldrich) was dissolved in dry dioxane and added directly to the beads; this mixture was gently shaken for 30 minutes. The beads were then dried and washed 5x with 3 mL dry dioxane, and a 1 mL solution of 1,9 diaminononane (0.202 mmol; 32 mg) in 9:1 dioxane:water was added and the reaction gently stirred at RT for 25 minutes (Scheme 8.2). Following reaction, the beads were rinsed 3x each in dry dioxane followed by MeOH.

$[Rh(phi)_2bpy']Cl_3$ was prepared as previously described (*21*), and the structure confirmed with ^1H NMR in deuterated water (Figure 8.2). To prepare for 5′ coupling, $[Rh(phi)_2bpy']Cl_3$ (0.013 mmol; 11.4 mg) was placed in a scintillation vial and dissolved in a minimal volume of 1:1:1 dry $CH_3OH/CH_2Cl_2/CH_3CN$ along with 0.027 mmol (8.1 mg) N,N,N′,N′-Tetramethyl-O-

(N-succinimidyl)uronium tetrafluoroborate (TSTU). The reaction was started by adding 0.034 mmol (5.92 μL) N,N-Diisopropylethylamine (DIEA) and stirred at RT for one hour (Scheme 8.3). Immediately after addition of DIEA, the solution went from orange to dark red. After one hour, a small sample was analyzed by ESI mass spectrometry, which yielded two major peaks at 866 Da and 994 Da (Figure 8.3). The peak at 866 Da consistent with NHS ester formation accompanied by loss of two protons from the phi ligand imine nitrogen atoms in DIEA, while that at 994 Da is attributable to TSTU coupled directly to the bpy′ carboxylic acid through the imine group.

After TSTU activation, the DNA beads were added directly to the rhodium reaction mix and stirred overnight (Scheme 8.4). Alternatively, the rhodium solution was added to the beads on the glass cylinder and shaken overnight; both methods produced equivalent yields. In both cases, the beads were placed on the glass cylinder, dried, and washed successively (3x each) with dry MeOH, CH_2Cl_2, and ACN. Even after drying, the beads remained orange, confirming successful coupling (Figure 8.4). Notably, the beads are strongly solvatochromic, appearing palest in MeOH and darkest in CH_2Cl_2, with intermediate intensity in ACN (Figure 8.4). Once rinsed, DNA was removed from the beads by incubation with 800 μL 28-30% ammonium hydroxide (Sigma-Aldrich) for 6 hours at 60 ℃.

Isolated DNA was dried on a speed vacuum, dissolved in 600 μL buffer (5 mM sodium phosphate, pH 7.0, 50 mM NaCl), filtered, and purified by HPLC (PLRPS column, 5-15% ACN/95-85% 50 mM ammonium acetate gradient over 35 minutes, 2 mL/min flow rate). To ensure collection of coupled DNA, absorbance was monitored at 260, 280, and 360 nm; two major peaks (consistent with previous reports and corresponding to separate enantiomers) eluted around 22 and 28 minutes (Figure 8.5). MALDI-TOF mass spectrometry was used to confirm the

identity of the DNA; a mass of 5819 Da was obtained, very close to the expected mass of 5814 Da. DNA was quantified by UV-visible spectroscopy using a molar extinction coefficient at 390 nm of 19000 $M^{-1}cm^{-1}$ for Rh-tethered DNA (Figure 8.6) (*20*).

Photooxidation Reactions with RPA

Reaction mixes (30 – 50 μL) consisted of 133 nM DNA (2 pmol) with a 0, 1, 5, or 10-fold molar excess of RPA, 0.2 mg/mL BSA, 150 mM NaCl, 50 mM Tris-HCl, pH 7.8, and 10% glycerol v/v. Samples were prepared in Eppendorf tubes, incubated at 37 ℃ for one minute to facilitate RPA binding, and transferred to glass vials for photooxidation. Photooxidation was accomplished by irradiation (30 minutes or 1 hour) under a solar simulator with a UVB/C filter applied, after which reactions were transferred back into Eppendorf tubes and dried on a speed vacuum. To cleave the DNA at sites of damage, dried DNA was dissolved in 100 μL buffer (50 mM Tris-HCl, pH 7.8, 150 mM NaCl) containing 10% piperidine by volume. This solution was then incubated 30 minutes at 90 ℃, then precipitated in 1 mL chilled EtOH, left on dry ice 30 minutes, and spun down at 14000x g. The supernatant was removed, taking care to ensure that the bulk of the radioactivity remained in the tube, and samples were dried on the speed vacuum. Samples were then washed 3x in water and counted on a liquid scintillation counter prior to drying a final time. Dried samples were dissolved in 2 μL formamide loading dye, heated to 90 ℃ for 5 minutes, and equivalent amounts of radioactivity were loaded directly onto a 20% denaturing polyacrylamide gel. Electrophoresis was carried out at 90 W power for ~2 hours, or until the dye had migrated ~3/4 the length of the gel. Gels were then exposed (1 hour for every 300000 cpm) and imaged on a Typhoon phosphorimager; bands were quantified using ImageQuant software.

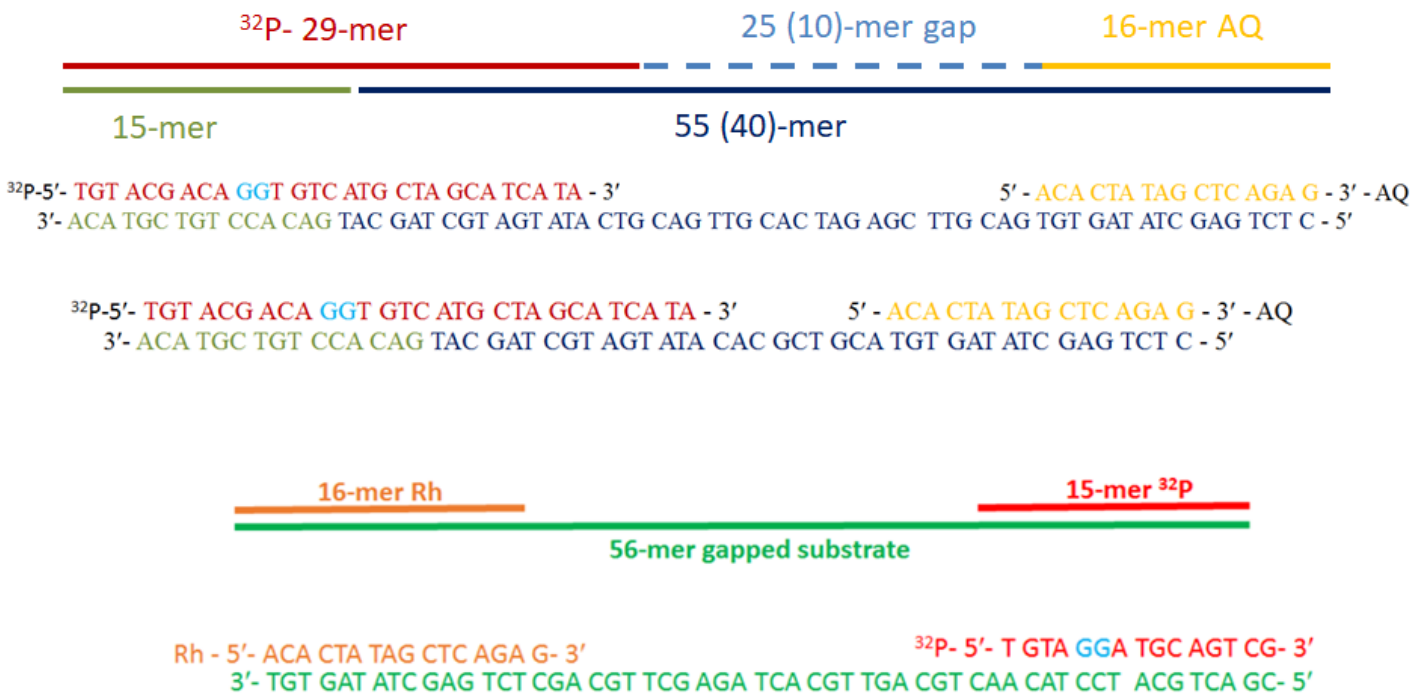


Figure 8.1 Design of AQ- and Rh-modified DNA substrate for RPA (top and bottom, respectively). The 25 or 10 nucleotide gap in the center fits two distinct binding modes of RPA, and the main differences between the substrates are the location and identity of the photooxidant. The decrease in size of the radiolabeled oligomer in the Rh-DNA substrate represented an effort to obtain higher yields in automated synthesis.

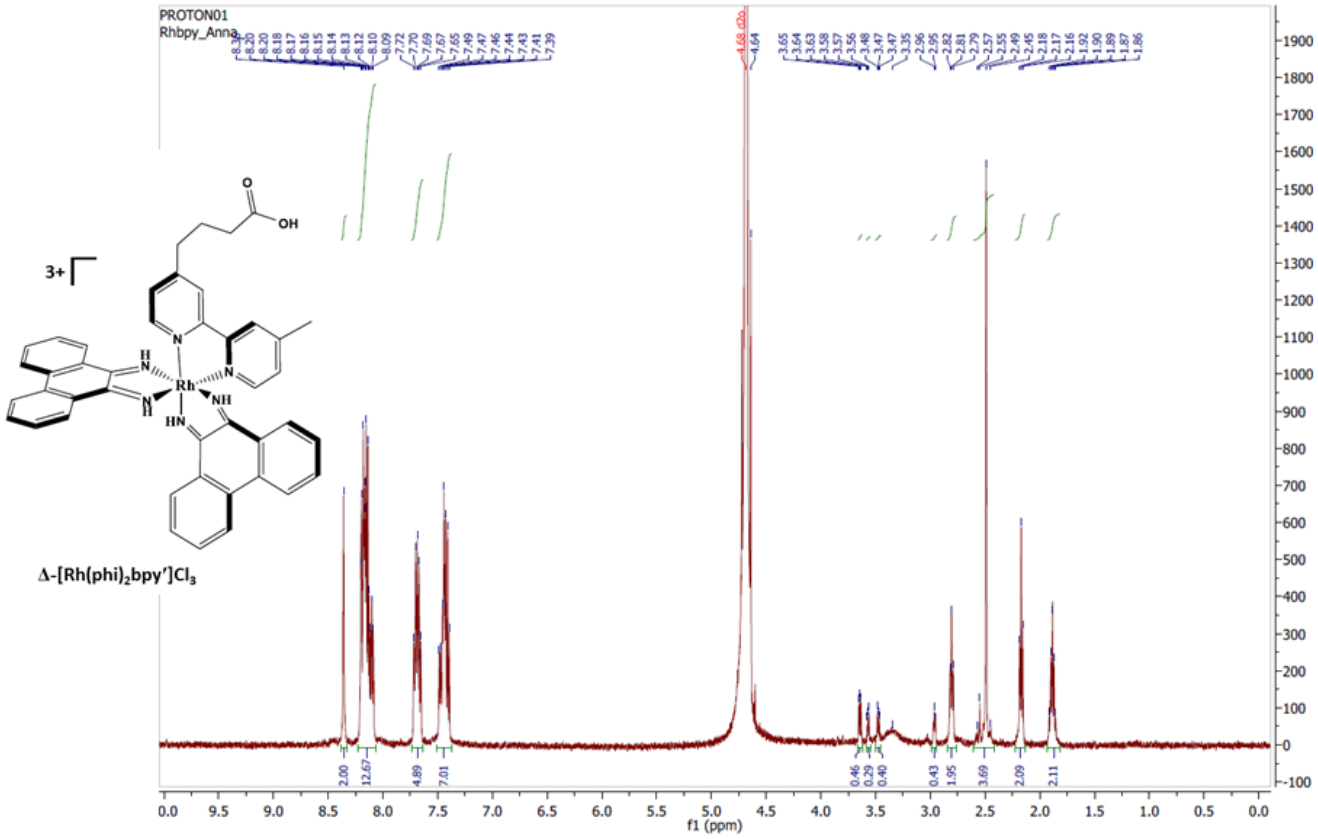
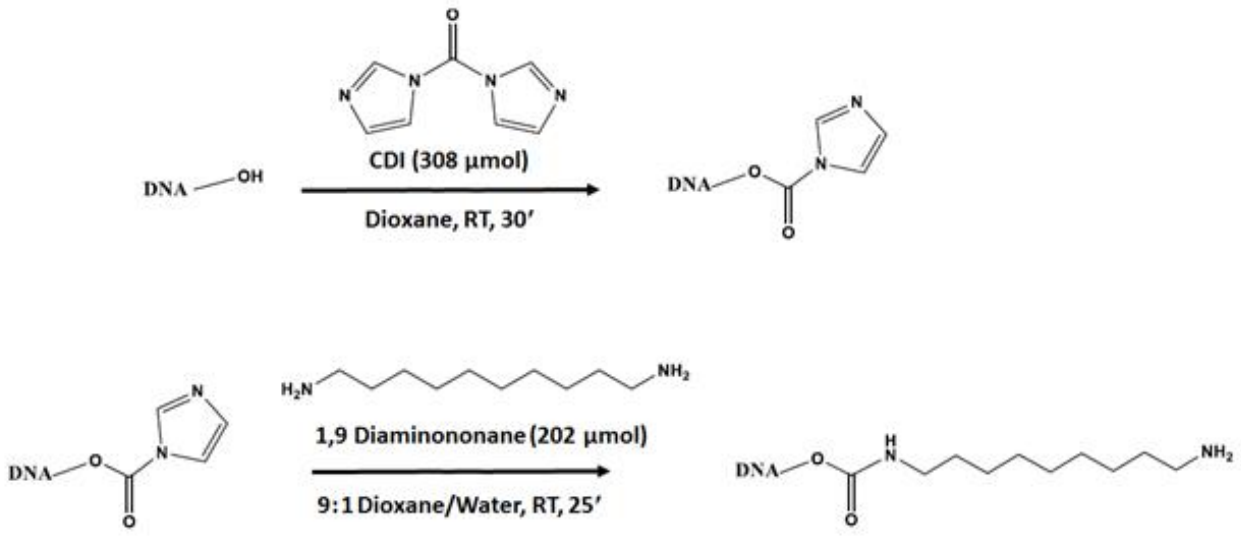
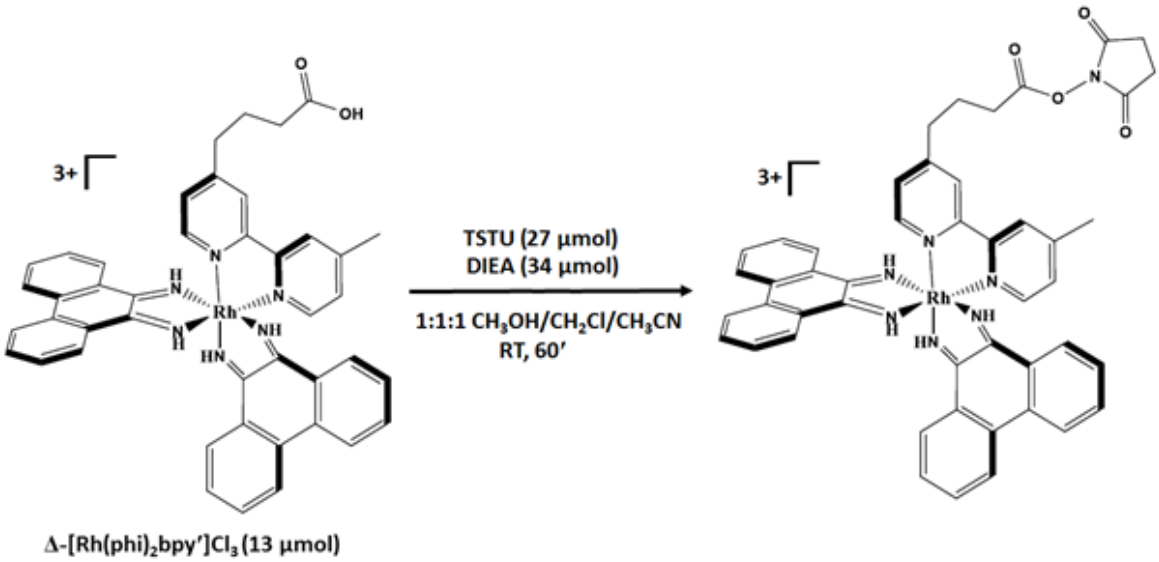


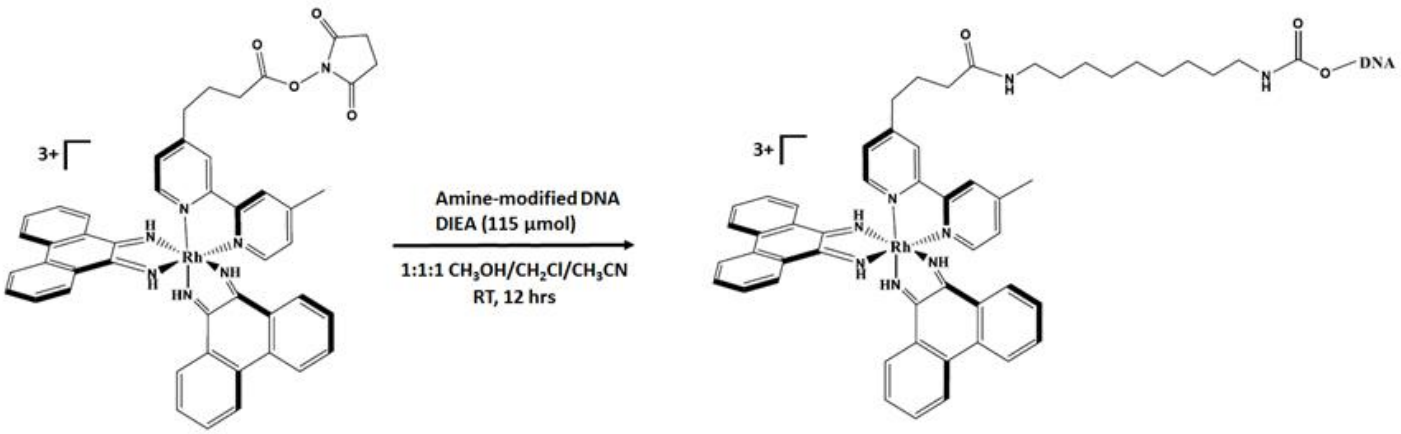
Figure 8.2 ^{1}H NMR spectrum of [Rh(phi)2bpy′]$^{3+}$. All peaks integrate to the expected proportions, and are in agreement with previous studies (*20*).



Scheme 8.2 Amino modification of DNA on CPG beads with the DMT group removed.



Scheme 8.3 TSTU activation of $[Rh(phi)_2bpy']^{3+}$ for coupling to amino-modified DNA. NHS ester formation is shown, but ESI-MS (Figure 8.3) confirms that this is one of two products along with a direct conjugation of TSTU at the imine group.



Scheme 8.4 Coupling of amino-modified DNA to activated $[Rh(phi)_2bpy']^{3+}$.

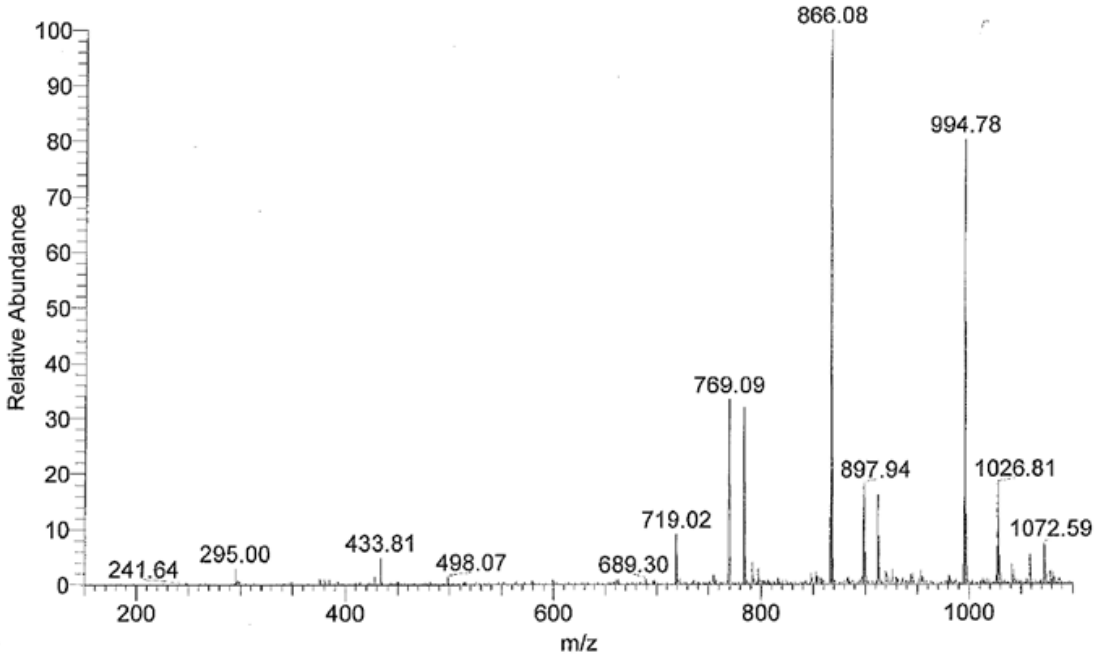


Figure 8.3 ESI-MS spectrum of $[Rh(phi)_2bpy']^+$ after 1 hour of TSTU activation in 1:1:1 $CH_3OH/CH_2Cl_2/CH_3CN$ with DIEA present. The major peaks at 866 Da and 994 Da are consistent with NHS ester formation and direct [Rh(phi)2bpy′] TSTU coupling, respectively.

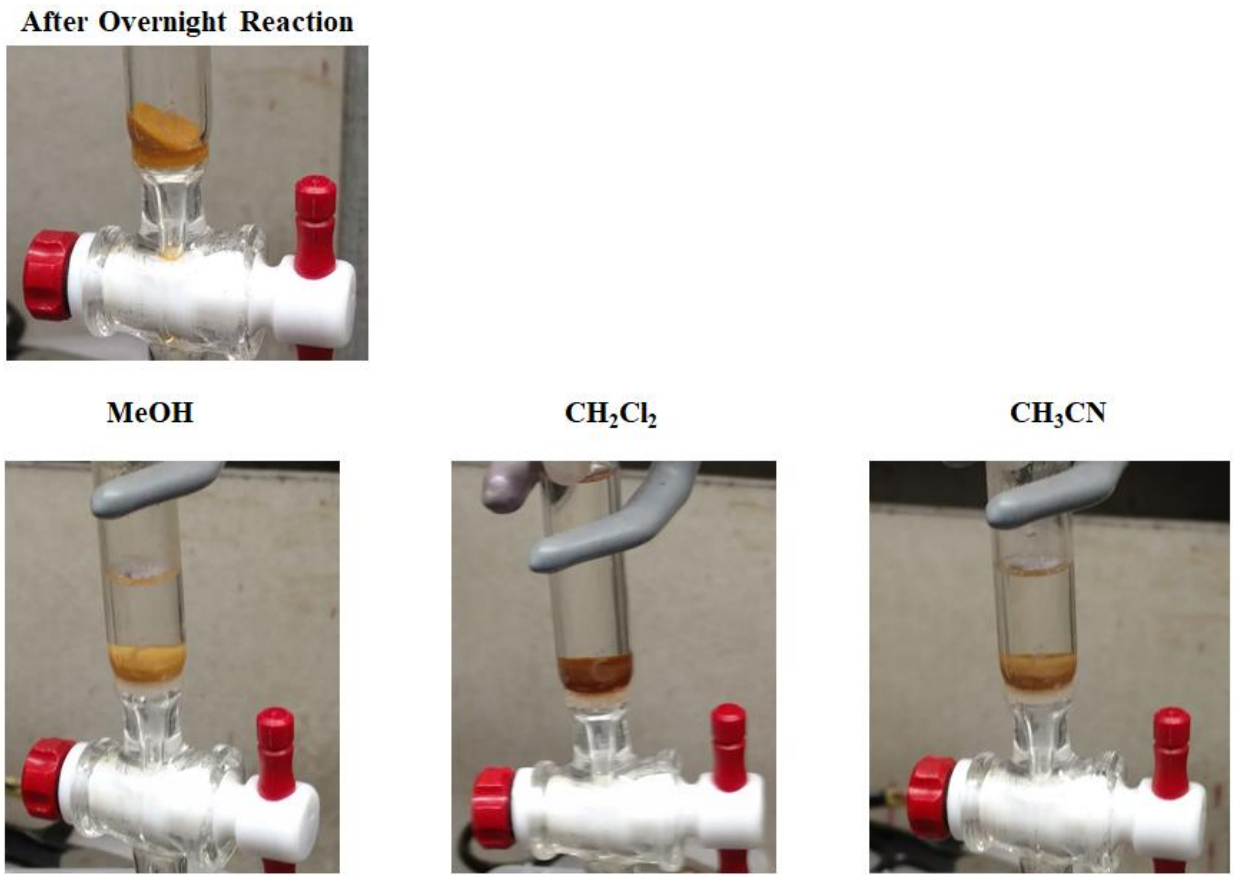


Figure 8.4 DNA on CPG beads post-Rh coupling and in a series of solvent washes.

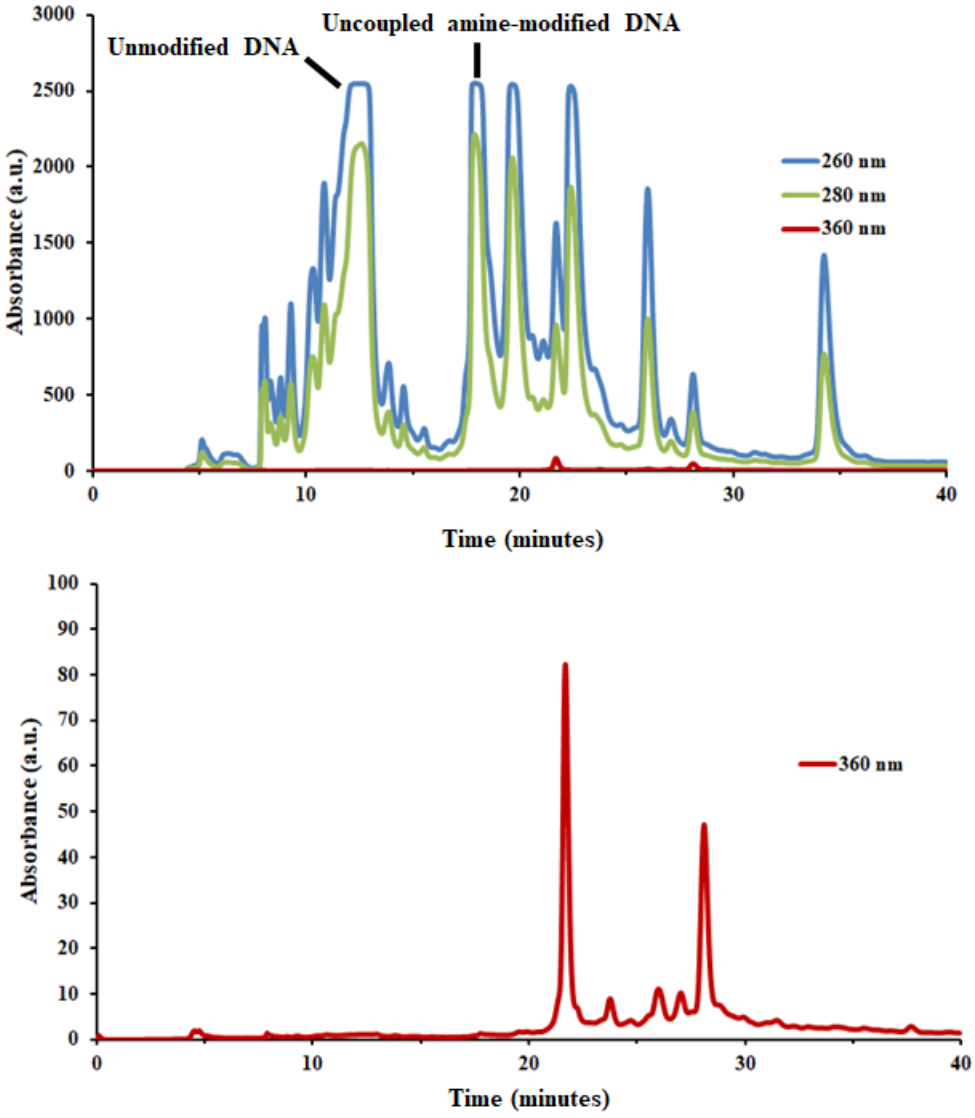


Figure 8.5 HPLC traces for Rh-conjugated DNA with UV-visible absorbance monitored at 260 and 280 nm to track DNA and 360 nm to track Rh. Peaks eluted around 12 minutes (unmodified DNA), 18-22 minutes (amino-modified DNA), and 21 and 28 minutes (Rh-coupled DNA).

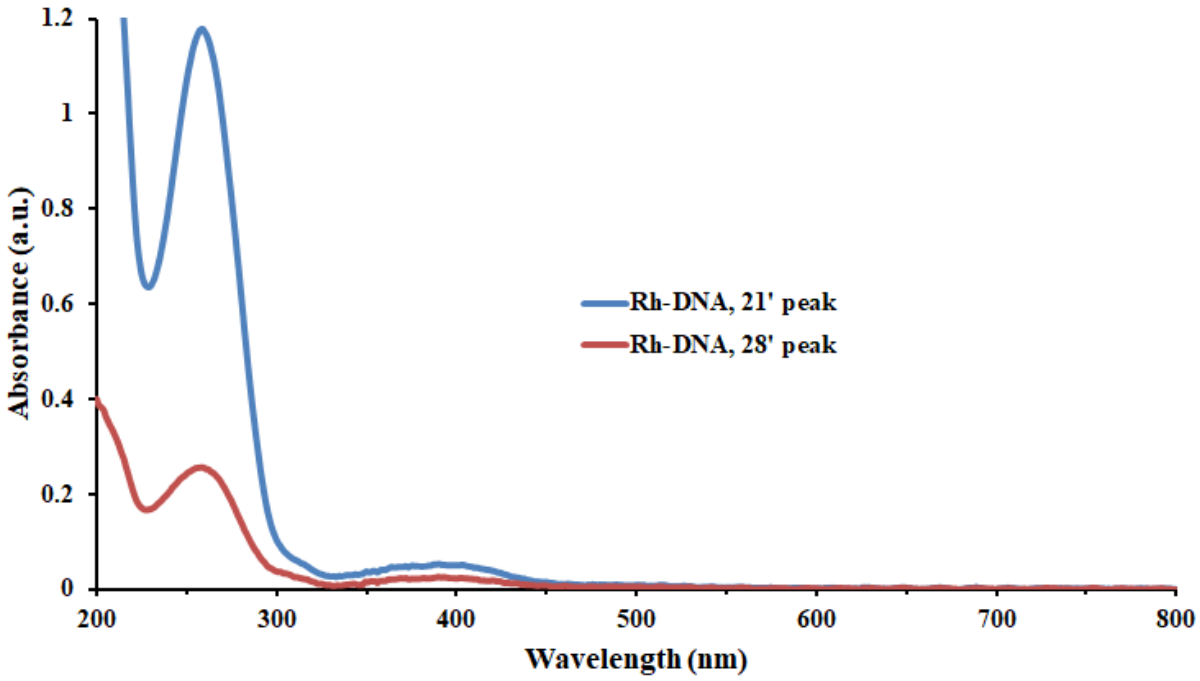


Figure 8.6 UV-visible spectra of 16-mer Rh-modified DNA. The UV-visible spectrum contains the broad peak centered at 390 nm expected for DNA-bound $[Rh(phi)_2bpy']^{3+}$ conjugated to DNA. The two peaks shown are from different collections from HPLC representing different enantiomers.

Results and Discussion

RPA Reactions with Photooxidation by AQ

In AQ reactions, DNA was incubated with increasing concentrations of RPA, and control reactions lacking RPA and containing a fully duplexed substrate were subjected to the same treatment as RPA reactions. DNA containing both a 10-mer and a 25-mer gap was used to assess CT in RPA bound in its shortest and longest binding modes, respectively. From the polyacrylamide gels (Figure 8.7), it is apparent that the majority of DNA substrate (78% or more by band volume) did not cleave in any reaction. However, the reactions with 25 nM RPA on a 10-mer gapped substrate and 10 nM RPA on a 25-mer gapped substrate do show some low molecular weight bands representing 2-5% of the total radioactivity present. Using Maxam-Gilbert sequencing lanes as a guide, DNA cleavage, when present, occurred primarily at the expected GG site, but products are also present at the neighboring G and A positions. Overall, these results provide an intriguing start, but low photooxidation yields make it impossible to say if RPA had any effect, as even what little DNA damage is present does not occur with any particular pattern on either the 10- or 25-mer gapped substrates.

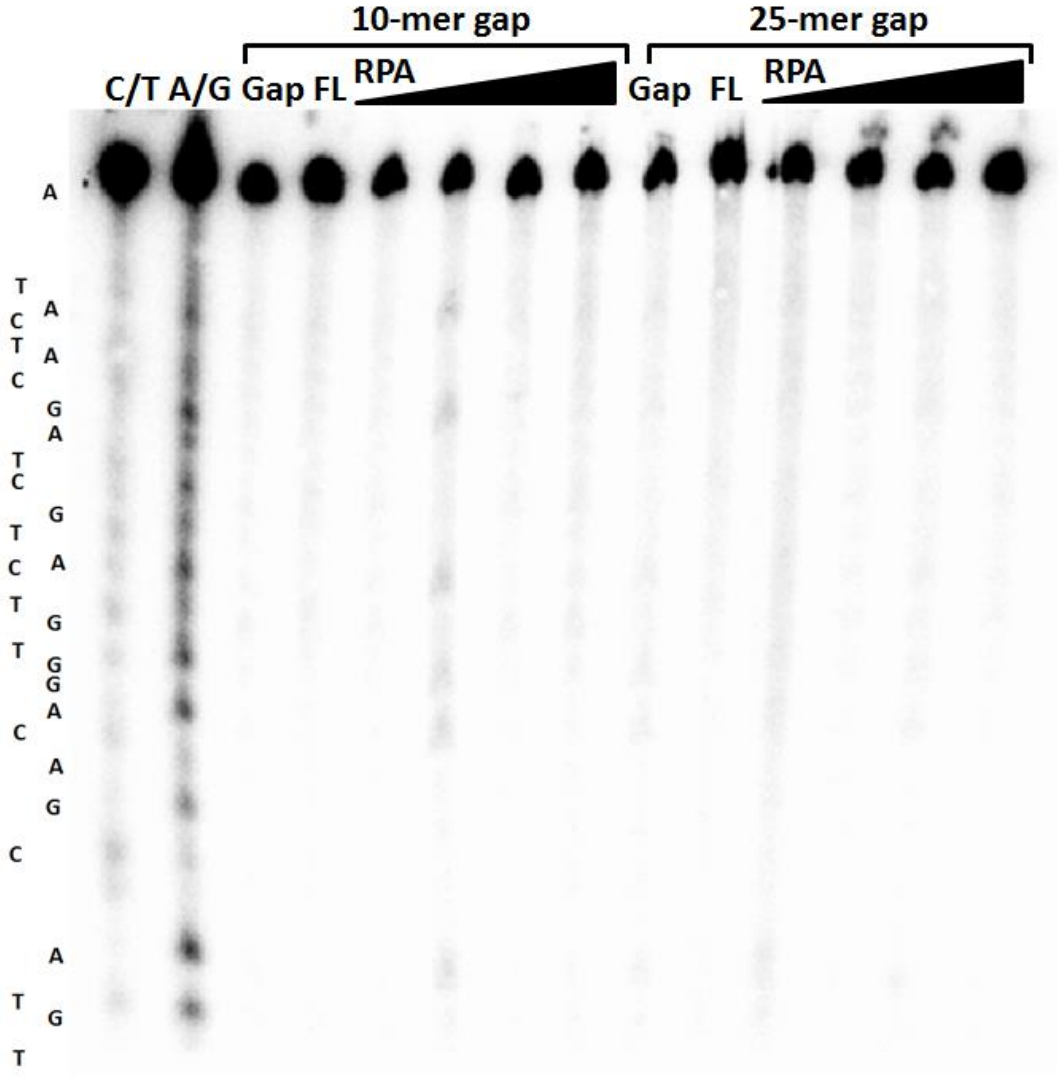


Figure 8.7 RPA reactions visualized on a 20% denaturing polyacrylamide gel. Based on Maxam-Gilbert sequencing lanes, several RPA reactions (25 nM RPA on 10-mer gap DNA and 10 nM RPA on 25-mer gap DNA) show damage at A/G sites representing 2-5% of the sample, but low yields prevent further analysis.

RPA Reactions with a Rh-Based Photooxidant

To improve oxidation yields, the photooxidant in our system was switched to $[Rh(phi)_2bpy']^{3+}$. $[Rh(phi)_2bpy']^{3+}$ yields are reported to be higher when covalently tethered to the 5′ end of single stranded DNA, so the polarity of the modifier was reversed. To simplify substrate preparation, however, the 15-mer sequence used in AQ DNA was eliminated and the sequences were adjusted to leave only three components for RPA reactions (four for fully duplexed controls). Finally, to facilitate troubleshooting, only the 25-mer gapped substrate was investigated in these experiments, and irradiation time was increased to one hour in an effort to further enhance yields.

At this point, data from gels are inconclusive for several reasons (Figure 8.8). In all experiments, indiscriminate DNA cleavage was observed. In early experiments, this was likely due to excessive cutting by piperidine in water rather than buffer. Extra care was taken in later efforts to treat with piperidine in buffer and to remove it through extensive washing and drying steps. Unfortunately, the result of this prolonged treatment (1 week in total) was likely ^{32}P-induced damage; supporting this, Maxam-Gilbert sequencing lanes prepared at the same time and treated in the same way also show indiscriminate cleavage. In summary, the problems encountered in these early efforts are expected to be simple to overcome, and future experiments will likely yield more fruitful results.

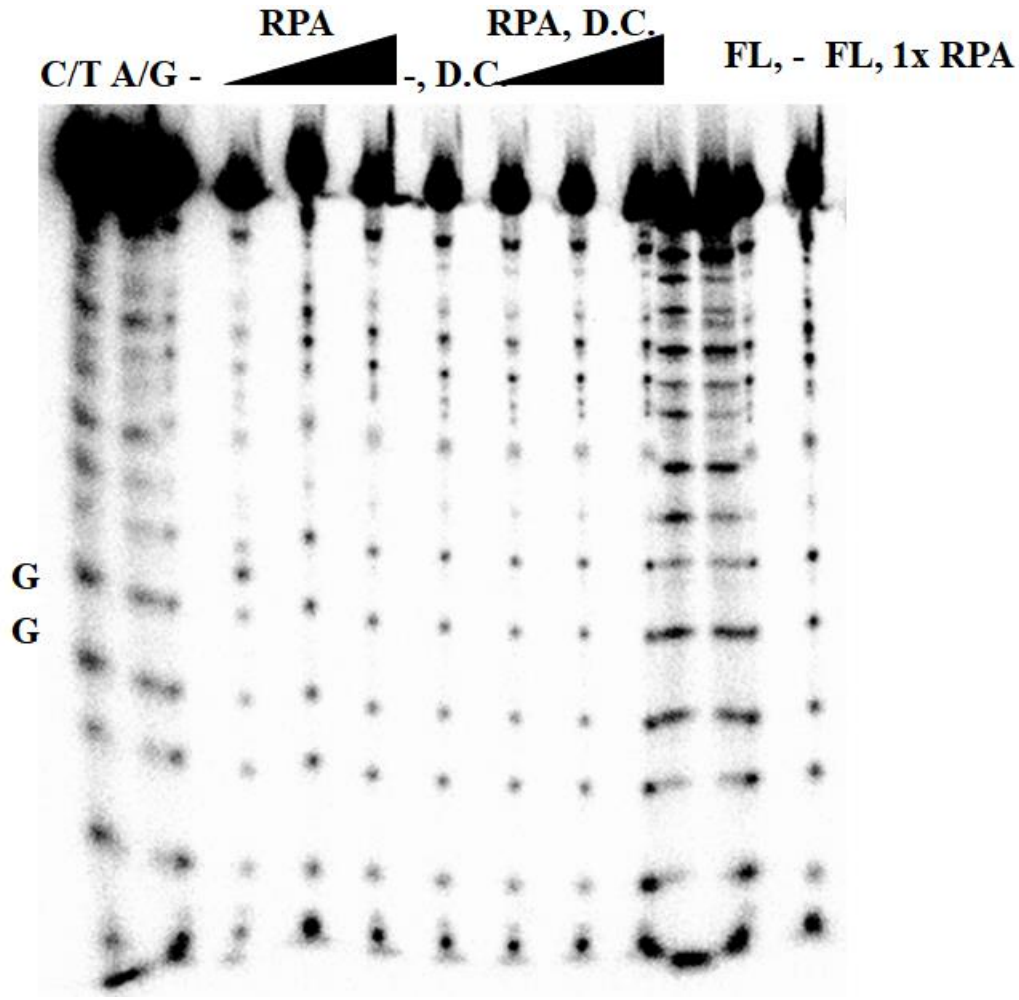


Figure 8.8 20% denaturing polyacrylamide gel containing RPA reactions alongside controls. No apparent differences are present between any reactions. Furthermore, both dark controls (D.C.) and Maxam-Gilbert sequencing lanes also show indiscriminate cleavage products. The damage seen in this gel is likely the result of overly long (several days) storage post-reaction, allow time for ^{32}P-induced damage even after piperidine removal. The DNA substrate contained a 25 nt gap, and reactions were run in the presence of 0.2 mg/mL BSA, 150 mM NaCl, and 10% glycerol v/v in 50 mM Tris-HCl, pH 7.8.

References

1) Burgers, P.M.; Kunkel, T.A. Eukaryotic DNA Replication Fork. *Annu. Rev. Biochem.* **2017**, *86*, 417 – 438.

2) Berti, M.; Vindigni, A. Replication Stress: Getting Back on Track. *Nat. Struct. Mol. Biol.* **2016**, *23*, 103 – 109.

3) Bartels, P.L.; Stodola, J.L.; Burgers, P.M.J.; Barton, J.K. A redox role for the [4Fe4S] cluster of yeast DNA polymerase δ. *J. Am. Chem. Soc.* **2017**, *139*, 18339 – 18448.

4) O'Brien, E.; Holt, M.E.; Thompson, M.K.; Salay, L.E.; Ehlinger, A.C.; Chazin, W.J.; Barton, J.K. The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport. *Science* **2017**, *355*, 6327.

5) McDonnell, K.J.; Chemler, J.A.; Bartels, P.L.; O'Brien, E.; Marvin, M.L.; Ortega, J.; Stern, R.H.; Raskin, L.; Li, G.; Sherman, D.H.; Barton, J.K.; Gruber, S.B. A Human MUTYH Variant Linking Colonic Polyposis to Redox Degradation of the $[4Fe4S]^{2+}$ Cluster. **2018**, *Submitted*.

6) Jain, R.; Vanamee, E.S.; Dzikovski, B.G.; Buku, A.; Johnson, R.E.; Prakash, L.; Prakash, S.; Aggarwal, A.K. An iron-sulfur cluster in the polymerase domain of yeast DNA polymerase ε. *J. Mol. Biol.* **2014**, *426*, 301 – 308.

7) Netz, D.J.A.; Stith, C.M.; Stümpfig, M.; Köpf, G.; Vogel, D.; Genau, H.M.; Stodola, J.L.; Lill, R.; Burgers, P.M.J.; Pierik, A.J. Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat. Chem. Biol.* **2012**, *8*, 125 – 132.

8) Pokharel, S.; Campbell, J.L. Crosstalk between the nuclease and helicase activities of Dna2: role of an essential iron-sulfur cluster domain. *Nucleic Acids Res.* **2012**, *40*, 7821 – 7830.

9) Alseth, I.; Eide, L.; Pirovano, M.; Rognes, T.; Seeberg, E.; Bjørås, M. The *Saccharomyces cerevisiae* homologues of Endonuclease III from *Escherichia coli*, Ntg1 and Ntg2, are both required for efficient repair of spontaneous and induced oxidative damage in yeast. *Mol. Cell. Biol.* **1999**, *19*, 3779 – 3787.

10) Aspinwall, R.; Rothwell, D.G.; Roldan-Arjona, T.; Anselmino, C.; Ward, C.J.; Cheadle, J.P.; Sampson, J.R.; Lindahl, T.; Harris, P.C.; Hickson, I.D. Cloning and characterization of a functional human homologue of *Escherichia coli* endonuclease III. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 109 – 114.

11) Arnold, A.R.; Grodick, M.A.; Barton, J.K. DNA Charge Transport: from Chemical Principles to the Cell. *Cell Chem. Biol.* **2016**, *23*, 183 – 197.

12) Boal, A.K.; Barton, J.K. Electrochemical Detection of Lesions in DNA. *Bioconj. Chem.* **2005**, *16*, 312 – 321.

13) Fanning, E.; Klimovich, V.; Nager, A.R. A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res.* **2006**, *34*, 4126 – 4237.

14) Bochkareva, E.; Korolev, S.; Lees-Miller, S.P.; Bochkarev, A. Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *The EMBO J.* **2002**, *21*, 1855 – 1863.

15) Wagenknecht, H.A.; Rajski, S.R.; Pascaly, M.; Stemp, E.D.; Barton, J.K. Direct observation of radical intermediates in protein-dependent DNA charge transport. *J. Am. Chem. Soc.* **2001**, *123*, 4400 – 4407.

16) Fan, J.; Pavletich, N.P. Structure and conformational change of a replication protein A heterotrimer bound to DNA. *Genes & Dev.* **2012**, *26*, 2337 – 2347.

17) Nuñez, M.E.; Noyes, K.T.; Barton, J.K. Oxidative Charge Transport through DNA in Nucleosome Core Particles. *Chem. & Biol.* **2002**, *9*, 403 – 415.

18) Schaefer, K.N.; Barton, J.K. DNA-Mediated Oxidation of p53. *Biochemistry* **2014**, *53*, 3467 – 3475.

19) Arunkumar, A.I.; Klimovich, V.; Jiang, X.; Ott, R.D.; Mizoue, L.; Fanning, E.; Chazin W.J. Insights into hRPA32 C-terminal domain-mediated assembly of the simian virus 40 replisome. *Nat. Struct. Mol. Biol.* **2005**, *12*, 332 – 339.

20) Homlin, R.E.; Dandliker, P.J.; Barton, J.K. Synthesis of Metallointercalator-DNA Conjugates on a Solid Support. *Bioconj. Chem.* **1999**, *10*, 1122 – 1130.

21) Olmon, E.D.; Hill, M.G.; Barton, J.K. Using Metal Complex Reduced States to Monitor the Oxidation of DNA. *Inorg. Chem.* **2011**, *50*, 12034 – 12044.

Chapter 9

Summary and Perspective

The importance of [4Fe4S] clusters in DNA-processing enzymes is difficult to understate. [4Fe4S] clusters have now been reported in base excision repair (BER), nucleotide excision repair (NER), and DNA replication pathways in organisms from bacteria to man (*1 - 9*). This is underscored by the strict conservation of such a metabolically expensive complex when, in many cases, structural integrity and basic enzymatic activity can be achieved with zinc or even bulky amino acids (*10-12*). In addition, many of these proteins have disease-relevance, in particular human NTHL1 and MUTYH BER glycosylases (*13, 14*). Thus, understanding the function of the [4Fe4S] cluster in these proteins is a matter of great importance.

Although their response to oxidants and reductants in solution varies (*1 - 9*), previous work in the Barton lab has demonstrated that DNA binding activates the cluster for oxidation, making the $[4Fe4S]^{3+/2+}$ couple accessible at biological potentials (*15*). Furthermore, once oxidized, the DNA-binding affinity is increased by ~500-fold at least in *Escherichia coli* endonuclease III (EndoIII) (*15, 16*). In addition to activating these proteins for redox activity, DNA has been shown to serve as a means of long-range communication through the process of DNA-mediated charge transport (DNA CT) (*17*). Bacterial BER proteins use DNA CT as a means of localizing to their target lesions (*17*). In this model, proteins are initially oxidized to the $[4Fe4S]^{3+}$ state by species like the guanine radical cation that are generated under conditions of oxidative stress (*18*). Once oxidized, these individual proteins can then be reduced by a distally protein in the reduced state; at this point, the reduced protein diffuses away from the DNA, while the oxidized protein remains tightly bound. This effectively confirms that the DNA stretch between the two proteins is undamaged. If an intervening lesion is present between two proteins, DNA CT is attenuated and this search process cannot occur. In this case, the oxidized proteins remain bound in the vicinity of the lesion and are poised to resolve the damage. Overall, this

model serves to explain how BER proteins find scarce targets on a biologically relevant time scale.

While the work leading up to this DNA repair model has provided great insight into the function of [4Fe4S] clusters in DNA, several details remained to be elucidated. In particular, the precise cause of the potential shift seen upon DNA binding was not clear, and the relevance of redox activity and DNA CT to human disease had not been addressed. Indeed, eukaryotic repair proteins had not been directly investigated using the electrochemical and spectroscopic techniques typically used in the Barton lab. Moving beyond repair, the recently reported occurrence of [4Fe4S] clusters in eukaryotic DNA replication brought to light a set of proteins whose activity could not be easily understood using the model developed for BER.

In this thesis work, we have addressed each of these questions using a combination of electrochemical, spectroscopic, and biochemical techniques. First, the molecular cause of the potential shift was investigated using a combination of electrochemistry on a pyrolytic graphite edge (PGE) electrode and S K-edge X-ray absorbance spectroscopy (XAS) (*Chapter 2, 3*). Second, we carried out the first complete characterization of a disease-relevant mutant of a human [4Fe4S] protein, the BER glycosylase MUTYH (*Chapter 4*). Third, the role of the [4Fe4S] cluster in yeast DNA polymerase (Pol) δ, the eukaryotic lagging strand DNA polymerase, was assessed using a variety of complementary *in vitro* methods (*Chapter 5-7*).

Experiments with EndoIII and MutY on PGE electrodes in the presence and absence of DNA revealed a potential shift similar to that seen on highly-oriented pyrolytic graphite (HOPG) in earlier studies (*15, Chapter 2*). Simply considering electrode surfaces, this work represented an advance by making the electrochemistry of DNA-free proteins simpler and more reproducible. HOPG, which was used for this type of work in the past, is a very flat, hydrophobic

surface ill-suited to protein adsorption, and is furthermore quite challenging to prepare in a reproducible manner (*15*). PGE, in contrast, is a rough surface with plenty of adsorption sites for proteins; these effects can be enhanced by drying the protein on in a thin film and applying highly conductive carbon nanotubes (*Chapter 2*). With respect to the role of nearby amino acid residues in tuning [4Fe4S] cluster potential, we used PGE electrochemistry to study a series of EndoIII point mutants in which single charges near the cluster were altered. No significant potential shift was observed in the absence of DNA, supporting the notion that the negatively charged DNA backbone is the dominant effect in tuning the [4Fe4S] redox potential.

In parallel with PGE electrochemistry, S K-edge XAS was used to study the effects of DNA binding and solvation on the potential of EndoIII and MutY (*Chapter 3*). S atom absorption in XAS reports on Fe-S bond covalency, which directly reflects redox potential. Notably, a decrease in absorbance corresponding to a -150 mV shift in potential was observed for EndoIII upon DNA binding; similar results occurred for MutY, although spectra were too small and noisy for accurate quantification. In addition, removal of solvent by lyophilization decreased absorbance, although not to the same extent seen with DNA binding; furthermore, the same effect of DNA-binding was seen in lyophilized samples. These results indicate that solvent hydrogen-bonding has a relatively minor role in defining the [4Fe4S] potential, and, together with PGE electrochemistry, emphasize the importance of DNA in this process.

Base excision repair in humans involves homologues of many of the *E. coli* proteins, including NTHL1 and MUTYH, which are, respectively, the EndoIII and MutY homologues (*13, 14*). Notably, recent years have seen an increasing number of catalogued mutations in the Fe-S region of these proteins in cancer databases, although none had been characterized *in vitro* up to this point. The discovery of a novel MUTYH mutation in one of the cysteine residues that ligate

the [4Fe4S] cluster (C306W) brought an opportunity to advance both the understanding of the cluster in human disease and the role of CT in human cells (*Chapter 4*). MUTYH C306W was determined to be pathological, and biochemical assays showed low activity levels and poor DNA binding relative to wild type MUTYH and the mutants Y179C and G396D, which are cancer-relevant but have no defects in the Fe-S domain. UV-visible spectroscopy confirmed that MUTYH C306W could still bind cluster as effectively as wild type, but electrochemistry showed that this mutant is susceptible to cluster degradation in the presence of oxygen. Using EPR spectroscopy, we identified the degradation product as the $[3Fe4S]^{+}$ cluster. Taken as a whole, these results indicate that MUTYH C306W tends to lose an iron upon aerobic oxidation, resulting in a $[3Fe4S]^{+}$ cluster. This species would already be expected to have less electrostatic attraction to the DNA than the $[4Fe4S]^{3+}$, or even the $[4Fe4s]^{2+}$ states, and reduction to the $[3Fe4S]^{0}$ state exacerbates this problem even further. Ultimately, the cluster degrades in the course of the redox search, impairing DNA binding and, as a consequence, enzymatic activity. In the context of work in the Barton lab, this study was the first confirmation that the human homologues of bacterial [4Fe4S] BER proteins have comparable redox activity when DNA-bound, and represents a first step toward assessing the pathology of numerous cancer-relevant MUTYH and NTHL1 mutations.

While CT in eukaryotic DNA repair was reasonably assumed to work similarly to the bacterial version, no explanation existed for the role of [4Fe4S] clusters in DNA replication proteins until recently. A characterization of the cluster in yeast Pol δ over the course of this thesis work shed some light on this situation, and led to a more general expansion of the CT signaling model to incorporate diverse protein activities (*Chapter 5*). Electrochemistry on DNA-modified gold electrodes demonstrated that Pol δ is redox active with a midpoint potential in the

same range as previously studied proteins. Incorporation of an abasic site or CA mismatch into the DNA led to signal attenuation, confirming that Pol δ is capable of DNA-mediated signaling. EPR spectroscopy confirmed the identity of the relevant redox couple as $[4Fe4S]^{3+/2+}$, as expected. Anaerobic oxidation by either electrochemical methods or photooxidation with anthraquinone (AQ) resulted in a stalling of DNA synthesis as seen on alkaline agarose gels. Critically, re-reduction using bulk electrolysis restored activity to near-untreated levels. As a whole, these results show that the oxidation state of the [4Fe4S] cluster acts as a reversible switch for enzymatic activity, stalling Pol δ upon oxidation and reactivating processive DNA synthesis upon reduction. Oxidative stalling could represent a means to respond to replication stress, especially oxidative stress, on a time scale even more rapid than protein-protein signaling cascades in the standard stress response. This work thus led to a generalization upon the CT model, with cluster oxidation state serving as a switch to turn activity on or off in different metabolic contexts.

Efforts to elucidate the role of the cluster in Pol δ *in vivo* by using CT-deficient mutants were also begun in this work (*Chapter 6*). We designed and characterized two mutants, W1053A and Y1078A, that were conserved and located sufficiently close to the [4Fe4S] cluster to be involved in the CT pathway. Unfortunately, the complete set of *in vitro* studies revealed structural instability in both mutants, necessitating a switch to less destabilizing mutations such as W, Y → F. While nothing conclusive came out of these efforts, the failures can be used to guide the next round of experiments, and overall this line of investigation remains worth pursuing.

Long-range CT in the context of replication could certainly occur within newly-synthesized DNA stretches, but much of the DNA present at the fork is single-stranded and

coated with the single-stranded binding protein RPA (*Chapter 8*). Single-stranded DNA cannot support CT, so long-distance signaling during replication appears limited. However, given that aromatic amino acid residues can electronically couple into the base stack and facilitate CT, we reasoned that RPA, with its numerous aromatic stacking interactions with DNA, might be able to open up more pathways to this type of communication. To investigate this possibility, we have begun setting up photooxidation experiments to see if RPA can bridge a Rh photooxidant and a GG site on the other side of a single-stranded gap. No significant results have yet been achieved in this work, but the possibility is intriguing and will be continued.

In summary, the work detailed in this thesis has expanded upon previous work in DNA repair and initiated significant advances in understanding the role of [4Fe4S] proteins in DNA replication. Furthermore, several new methods were brought to the Barton lab in the course of these studies, and the work on DNA replication in particular sparked new approaches to the study of [4Fe4S] proteins in the context of DNA CT. As is readily apparent from the leaps in understanding made in the last decade, this work itself is not the final say on these issues as more proteins involved in DNA processing that contain [4Fe4S] clusters are identified, and it is almost a certainty that more critical roles for long range DNA CT will emerge.

References

1) Cunningham, R.P.; Asahara, H.; Bank, J.F.; Scholes, C.P.; Salerno, J.C.; Surerus, K.; Munck, E.; McCracken, J.; Peisach, J.; Emptage, M.H. Endonuclease III is an iron-sulfur protein. *Biochemistry* **1989**, *28*, 4450 – 4455.

2) Porello, S.L.; Cannon, M.J.; David, S.S. A Substrate Recognition Role for the $[4Fe4S]^{2+}$ Cluster of the DNA Repair Glycosylase MutY. *Biochemistry*, **1998**, *37*, 6465 – 6475.

3) Aspinwall, R.; Rothwell, D.G.; Roldan-Arjona, T.; Anselmino, C.; Ward, C.J.; Cheadle, J.P.; Sampson, J.R.; Lindahl, T.; Harris, P.C.; Hickson, I.D. Cloning and characterization of a functional human homolog of *Escherichia coli* endonuclease III. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 109 – 114.

4) Alseth, I.; Eide, L.; Pirovano, M.; Rognes, T.; Seeberg, E.; Bjørås, M. The *Saccharomyces cerevisiae* Homologues of Endonuclease III from *Escherichia coli*, Ntg1 and Ntg2, are both Required for Efficient Repair of Spontaneous and Induced Oxidative DNA Damage in Yeast. *Mol. Cell Biol.* **1999,** *19*, 3779 – 3787.

5) Ren, B.; Duan, X.; Ding, H. Redox control of the DNA damage-inducible protein DinG helicase activity via its iron-sulfur cluster. *J. Biol. Chem.* **2009**, *284*, 4829 – 4835.

6) Fan, L.; Fuss, J.O.; Cheng, Q.J.; Arvai, A.S.; Hammel, M.; Roberts, V.A.; Cooper, P.K.; Tainer. J.A. XPD Helicase Structure and Activities: Insights into the Cancer and Aging Phenotypes from XPD Mutations. *Cell*, **2008**, *133*, 789 – 800.

7) Pokharel, S.; Campbell, J.S. Cross talk between the nuclease and helicase activities of Dna2: role of an essential iron-sulfur cluster domain. *Nucleic Acids Res.* **2012**, *40*, 7821 – 7830.

8) **a.** Klinge, S.; Hirst, J.; Maman, J.D.; Krude, T.; Pellegrini, L. An iron-sulfur domain of the eukaryotic primase is essential for RNA primer synthesis. *Nat. Struct. Mol. Biol.* **2007**, *14*, 875-877. **b.** Weiner, B.E.; Huang, H.; Dattilo, B.M.; Nilges, M.J.; Fanning, E.; Chazin, W.J. An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase. *J. Biol. Chem.* **2007**, *282*, 33444 – 33451.

9) Netz, D.J.A.; Stith, C.M.; Stümpfig, M.; Köpf, G.; Vogel, D.; Genau, H.M.; Stodola, J.L.; Lill, R.; Burgers, P.M.J.; Pierik, A.J. Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat. Chem. Biol.* **2012**, *8*, 125 – 132.

10) Netz, D. J. A.; Mascarenhas, J.; Stehling, O.; Pierik, A. J.; Lill, R. Maturation of cytosolic and nuclear iron-sulfur proteins. *Trends Cell Biol.* **2014**, *24*, 303 – 312.

11) Klinge, S.; Núñez-Ramírez, R.; Llorca, O.; Pellegrini, L. 3D architecture of DNA Pol α reveals the functional core of multi-subunit replicative polymerases. *EMBO J.* **2009**, *28*, 1978 – 1987.

12) Trasvina-Areñas, C. H.; Lopez-Castillo, L. M.; Sanchez-Sandoval, E.; Brieba, L. G. Dispensability of the [4Fe-4S] cluster in novel homologues of adenine glycosylase MutY. *FEBS J.* **2016**, *283*, 521 – 540.

13) Weren, R.D.A. et al. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nat. Genet.* **2015**, *47*, 668 – 671.

14) D'Agostino, V.G., Minoprio, A., Torreri, P., Marinoni, I., Bossa, C., Petrucci, T.C., Albertini, A.M., Ranzani, G.N., Bignami, M., and Mazzei, F. Functional analysis of MUTYH mutated proteins associated with familial adenomatous polyposis. *DNA Repair (Amst)*. **2010**, *9*, 700 – 707.

15) Gorodetsky, A.A.; Boal, A.K.; Barton, J.K. Direct electrochemistry of Endonuclease III in the presence and absence of DNA. *J. Am. Chem. Soc.* **2006**, *128*, 12082-12083.

16) Tse, E.C.M.; Zwang, T.J.; Barton, J.K. The Oxidation State of [4Fe4S] Clusters Modulates the DNA-Binding Affinity of DNA Repair Proteins. *J. Am. Chem. Soc.* **2017**, *139*, 12784 – 12792.

17) Arnold, A. R.; Grodick, M. A.; Barton, J. K. DNA Charge Transport: from Chemical Principles to the Cell. *Cell Chem. Biol*. **2016**, *23*, 183 – 197.

18) Yavin, E.; Boal, A. K.; Stemp, E. D. A.; Boon, E. M.; Livingston, A. L.; O'Shea, V. L.; David, S. S.; Barton, J. K. Protein-DNA Charge Transport: Redox Activation of a DNA Repair Protein by Guanine Radical. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 3546 – 3551.