Chapter 4:

Modeling Redox Signaling between DNA Repair Proteins for Efficient Lesion Detection[†§]

[†] partially adapted from Boal, A.K.; Genereux, J.C.; Sontz, P.A.; Gralnick, J.A.; Newman, D.K.; Barton, J.K. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15237.

[§] Dr. Amie Boal planned, performed, and analyzed all *in vivo* and mutation experiments and provided EndoIII and Pamela Sontz planned, performed, and analyzed the AFM experiments. Prof. Jeff Gralnick performed strain construction and assisted in designing *E. coli* experiments.

4.1. INTRODUCTION

Base excision repair (BER) proteins, from bacteria to humans, are challenged with combing the genome for DNA base lesions to maintain the integrity of our genetic material.[1,2] This challenge is remarkable given the low copy number of these proteins and that they must discriminate among small differences between modified and natural bases. For MutY, a BER protein in *E.coli* with a human homologue, there are 20 proteins in the *E.coli* cell[3] to interrogate 4.6 million bases; the ratio of binding affinities for the target lesion, an 8-oxoguanine:adenine mismatch, versus well-matched native base pairs is ≤ 1000.[4] Endonuclease III (EndoIII) recognizes a less prevalent lesion, hydroxylated pyrimidines, with equally low specificity; the copy number of EndoIII within *E.coli* is ~500.[1] How these glycosylases fix their substrate lesions, once found, has been well characterized[1,2], as are the structures of MutY and EndoIII bound to DNA.[5,6] Yet how these lesions are efficiently detected before excision is not established.

Location of damaged bases in the genome is likely the rate-limiting step in BER within the cell and, hence, a critical step in maintaining genomic integrity.[7] Current models for genome scanning to detect lesions involve protein sliding along the DNA, squeezing the backbone, slipping bases out to allow for interrogation, or finding transiently opened sites.[8,9] However, given the low copy number of these proteins and their need to sift through the genome to find often subtle base lesions, the time required for this search is long.

Many of these BER proteins contain [4Fe4S] clusters, common redox cofactors in proteins.[1,2] Increasingly, iron-sulfur clusters are found associated with varied DNA-binding proteins and located far from the enzymatic active site with no apparent function. For BER proteins, [4Fe4S] clusters were first thought to play a structural role. When not bound to DNA, these proteins are found in the

$[4Fe4S]^{2+}$ state and are not easily oxidized or reduced under physiological conditions.[10] However, for MutY and EndoIII, we have demonstrated using DNA-modified electrodes that DNA binding shifts the 3+/2+ cluster potential into a physiological range, ~100 mV vs. NHE for each BER enzyme;[11,12] DNA binding stabilizes the protein in the +3 form.

Given the sensitivity of DNA-mediated charge transport (CT) to mismatched and damaged bases, we have proposed that DNA repair glycosylases containing a redox-active [4Fe4S] cluster including EndoIII and MutY, use DNA CT as the first step in substrate detection by signaling one another to search cooperatively for damage in the genome.[11,12] DNA-mediated CT can proceed over long molecular distances on a short timescale.[13] Oxidative damage to DNA has been demonstrated with oxidants covalently tethered and spatially separated from damage sites at distances of >200 Å with negligible loss in efficiency.[14] Reductive CT has been shown to have an equally shallow distance dependence both in electrochemical studies[15] and in assemblies in solution.[16] Previous studies established that CT through DNA is possible in biological environments that include nucleosomes[17] and cell nuclei.[18] DNA CT is, however, extremely sensitive to perturbations in the intervening base-pair stack, such as DNA mismatches and lesions.[19,20] As an example, a single molecule of DNA covalently attached within a nanotube device can conduct charge perpendicular to the π-stack similarly to graphite, but the resistance increases 300-fold with a single base mismatch.[21] DNA-mediated electrochemistry has therefore been utilized in the development of sensors for mutational analysis[20] and protein binding.[22]

Given that this chemistry occurs at a distance and is modulated by the structural integrity of the base-pair stack, these reactions may be useful within the cell for long-range signaling to proteins. In that context, we have previously established the long-range oxidation of the DNA-bound BER enzymes in spectroscopic studies monitoring oxidation of the [4Fe4S] clusters by

guanine radicals in the duplex.[23] Importantly, we have also shown the injection of an electron into the base pair stack from the DNA-bound BER enzymes, with the electron trapped by a well-coupled modified base in the duplex.[24] Both with respect to hole injection into the DNA-bound proteins and electron injection into the DNA from the DNA-bound proteins, EndoIII and MutY behave equivalently, as expected given their similar DNA-bound redox potentials and structures. Here we explore whether it is reasonable to expect DNA-mediated CT to provide a means to facilitate the detection of damage *in vivo*, and then compare these predictions to some experimental data from *E. coli* and single-molecule AFM experiments.

4.2. METHODS AND RESULTS

4.2.1. DNA-BINDING AFFINITY OF ENDOIII

There are three properties that determine a protein's diffusive properties in a DNA environment: the one- and three- dimensional diffusion constants, and the nonspecific binding affinity for DNA.[8,25] Measurements have been performed for the nonspecific affinity of MutY and of EndoIII,[9,26] but these preceded the discovery that the electrochemical potential of iron-sulfur clusters in these proteins varies by 200 mV between the DNA-bound and free forms of the proteins.[27] Thus, measured affinities were for a mixture of oxidized and reduced protein. Hence, we measured binding affinities using EndoIII in the presence and absence of 3 mM DTT, to better estimate the individual nonspecific binding affinities of the two forms of the protein.

The binding buffer was 20 mM $NaPO_4$, 100 mM NaCl, 10% glycerol, 0.1 ug/uL bovine serum albumin. The gels were BioRad 10% TBE gels ran in 0.5X TBE at 4 ^oC. Protein was prepared at a concentration of 80 μM or 8 μM, and serial dilutions in factors of three were used to prepare the other protein concentrations. The labeled strand was of sequence

5'-CTGTAACGGGAGCTCGTGGCTCCATGATCG-3'. This strand and its complement were synthesized on an ABI DNA synthesizer using standard phosphoramidite chemistry, purified twice by reversed-phase HPLC and characterized by mass-spectrometry and UV-vis. Labeling was performed at the 5'-end with [^{32}P] γ-ATP using polynucleotide kinase, followed by isolation with Micro Bio-Spin 6 columns (BioRad). EndoIII was acquired from Dr. Amie Boal and stored at -80 ^{o}C. Samples were eluted through a 20% denaturing polyacrylamide gel for 1.5 hours at 90 W and imaged on a Storm 820 phosphoimager (Molecular Dynamics/ GE Healthcare). DNA bands were quantified by phosphoimagery using Image Quant 5.2 (Molecular Dynamics). In the absence of DTT, the dissociation constant was found to be 60 ± 10 nM. The presence of DTT increased the dissociation constant by about a factor of 5; the binding was weak enough such that greater than 80% DNA bound to protein would have required a higher concentration of EndoIII than was available. Interestingly, these results bracket the published[26] dissociation constant of 250 ± 100 nM for MutY with this DNA in 1 mM DTT.

The electrochemistry tells us that the dissociation constant (K_D) is between 2000 and 50000 times greater for the reduced versus the fully oxidized protein. Although we do not know the fraction of oxidized protein under ambient conditions, 60 nM serves as an upper bound for K_D^{ox}. We will use this highest possible value for K_D^{ox} and the corresponding lowest value for K_D^{red}, 120 μM, as they are those that are most favorable to the models that do not invoke CT signaling, and least favorable to the model for CT signaling. More accurate values will only increase the predicted improvement of CT signaling versus non-CT signaling models.

4.2.2. GENOME SCANNING CALCULATIONS WITHOUT CT SIGNALING

Protein diffusion to a cognate site on dilute, short DNA strands *in vitro* occurs faster than predicted by the Debye-Smoluchowski equation.[8] This is due to facilitation by the non-cognate DNA: protein weakly associates with non-specific regions of the DNA, and then slides in one dimension to the recognition sequence. *In vivo*, where the concentration of DNA is in the millimolar regime, the challenge of site recognition is different.[25,28] Rather than the non-specific DNA serving as a means to funnel the protein to the recognition site, it acts as a competitor that slows protein translocation. To overcome this challenge, the protein must frequently dissociate from non-specific DNA so that it can sample other portions of the genome, and hence avoid highly redundant sliding on sequences far from the target.

In this context, fast target location requires weak non-specific interactions, fast sliding along the DNA when association occurs, and strong specificity for protein recognition of cognate versus non-cognate DNA.[29] The physical challenge of achieving all of these conditions for the same protein has been widely discussed,[29-31] and many studies have demonstrated that fastest target location is achieved when the non-specific dissociation constant is equal to the concentration of base pairs.[28,32] For a transcription factor that does have the above properties, LacI, real-time single-molecule visualization of its translocation and target recognition within the cell is consistent with this model of efficient facilitated diffusion.[33]

DNA base excision repair proteins do not meet all of the conditions for fast target detection. The non-specific binding constants of oxidized and reduced EndoIII are about 40 nM and 40 μM respectively; the values for MutY appear similar.[26] The specificities are reported as ≤ 1000 for both EndoIII and for MutY.[4,26] There is evidence that the human oxoG:C-targeting base-repair enzyme hOGG1 can slide rapidly on stretched DNA.[34] Surprisingly, the one-

dimensional diffusion constant[35] is at the hydrodynamic limit. It is unclear whether this is consistent with the necessity of displacing bound water and ions, and how such rapid motion can allow specific recognition between the protein and its cognate lesion. We will not neglect the problem of extensive bound protein present in the cell,[36] which should substantially retard the maximum allowable sliding length.[31] Although this effect may help transcription factors with a defined cognate site and weak non-specific binding,[37] the presence of roadblocks should substantially slow search by repair proteins.

The simplest approach to calculating the rate of target recognition for MutY using facilitated diffusion is that described in the original derivation:[8]

$$T_{Search} = \ln\left(\frac{1}{1-P}\right)\left(\frac{2C_p\sqrt{D_1 K_D k_a}}{L}\right)^{-1}$$

$$k_a = 4\pi D_3 r_{DNA}$$

The protein diffusion constant, D_3, is determined from the Stokes-Einstein equation using the 10 cP viscosity of *E. coli* cytoplasm[38] and the measured Stokes radius of EndoIII (r_p).[39] A further adjustment must be made to consider the specificity; rapid dissociation will compete with catalysis. This is incorporated as

$$T_{repair} = T_{Search}\left(\frac{k_{cat}}{k_{cat} + \frac{k_d}{s}\left(\frac{s + l_s}{s}\right)}\right)^{-1}$$

$$l_s \equiv \sqrt{\frac{D_1}{k_d}}$$

where k_{cat} is the catalytic rate,[4] s is the specificity,[26] l_s is the sliding length, and D_1 is the one-dimensional diffusion constant.

This approach yields a respectable time of 96 seconds for a 90% chance of discovery and repair of a lesion by oxidized MutY. The challenge of this approach, however, is that it assumes that protein dissociation is followed by immediate reassociation to any other site on the genome *with equal probability*.[8,25,40] This derivation, while acceptable for short DNA strands in dilute solution, each containing a target site, is not appropriate in the context of genomic DNA, unless the protein has extremely weak non-specific affinity. It has been demonstrated by simulation that, for the genomic density present inside *E. coli*, a protein that dissociates from a single site is 87% likely to return to within a single sliding length of the original site.[25] When this factor, which represents only the most proximal dissociation-reassociation events, is considered, the search time increases to 230 seconds for the oxidized proteins. On a slightly larger scale, bacterial DNA adopts a solenoid-like structure, and the genome of *E. coli* is organized into discrete structural domains of 10 to 100 kb.[41] It is reasonable to expect that dissociation-reassociation events will feature substantial autocorrelation with respect to individual domains at each size-scale.

Furthermore, it has been shown that for *E. coli*, the average gap between bound structural proteins is about 10 to 80 bp.[42] This is adequate for the highly specific transcription factors such as LacI, which has a sliding length of < 85 bp,[33] and that might be able to ratchet other proteins off the edges of its cognate site, driven by recognition of part of the cognate site. For repair proteins, recognition is of a single base, and a lesion beneath a structural roadblock will only be found if the protein moves aside. Taking the most generous value of 80 bp, which is still much smaller than the 4400 bp sliding length without roadblocks, the search time for the oxidized protein becomes 430 minutes, much longer than the *E. coli* doubling rate. Hence, it is clear that target location by the base excision repair protein MutY is not explained by a straightforward diffusive mechanism.

4.2.3. GENOME SCANNING CALCULATIONS WITH CT SIGNALING

We have proposed that BER proteins bearing [4Fe4S] clusters exploit DNA-mediated CT as a fast, sensitive method to detect damage (**Figure 4.1**). This redox signaling model is initiated when one 2+ protein (donor) binds DNA (b, e), promoting electron transfer from the donor protein to a distal protein (acceptor) (c, f), already bound in the 3+ state. The newly oxidized donor protein remains DNA-bound while the reduced acceptor diffuses away (d, f). Integral to this model is a differential DNA affinity for the $[4Fe4S]^{3+}$ and $[4Fe4S]^{2+}$ forms of the protein. We have demonstrated this differential affinity by measuring a -200 mV potential shift associated with DNA binding that corresponds thermodynamically to ≥1000-fold difference in DNA affinity between the oxidized and reduced proteins.[27]

Importantly, the DNA-mediated CT reaction between two repair proteins can be considered a scan of the integrity of the intervening DNA, since DNA-mediated CT can only proceed through a well-stacked duplex. As illustrated in **Figure 4.1** (g), when the repair protein, already oxidized, is bound near a base lesion, DNA-mediated CT does not provide a pathway for reduction and subsequent protein dissociation. The protein instead remains bound to the duplex so that, on a slower timescale, the protein can processively diffuse to the target site; now, however, sliding is needed only across a small region and the low target specificity of the protein is sufficient for recognition.[4,37,43,44] Essentially, then, our proposal for base lesion detection utilizing DNA CT yields a redistribution of the BER enzymes onto local regions of the genome that contain lesions. Critical to this mechanism is DNA-mediated signaling among proteins bound at long range so that the proteins, despite their low abundance, cooperate with one another in localizing onto target sites. Simulations of MutY search using a similar mechanism have shown encouraging accumulation at lesion sites when CT is allowed.[45,46] These simulations allow direct electron

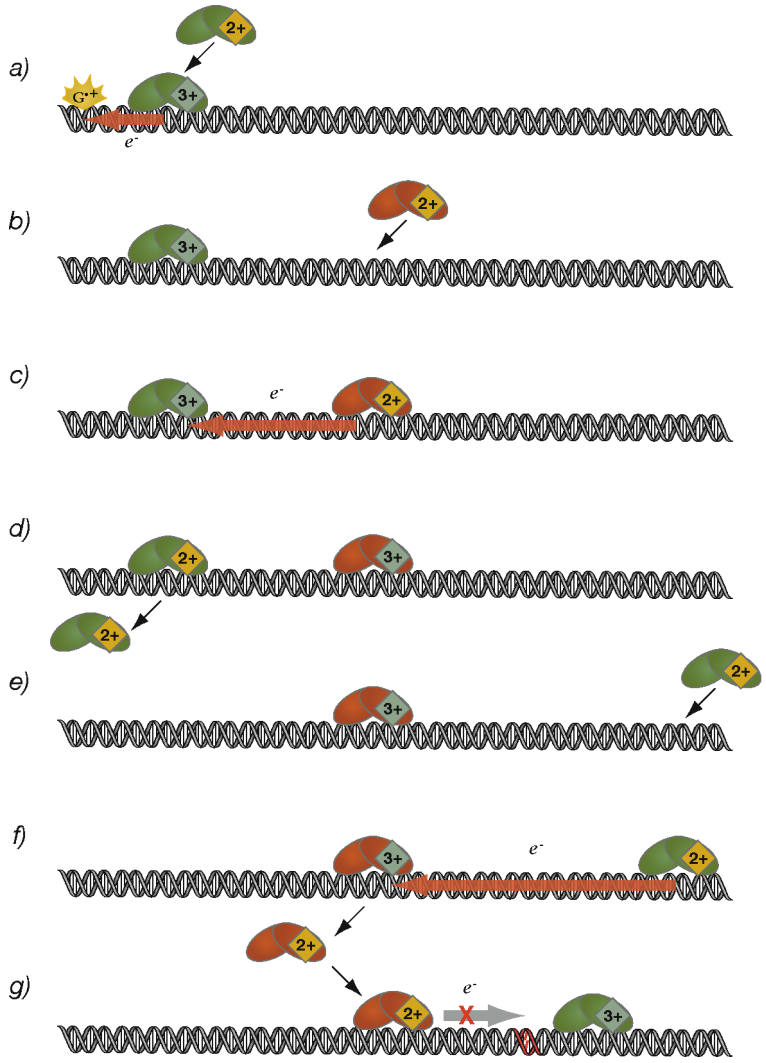

Figure 4.1. A model for DNA-mediated CT in DNA repair. In this model, DNA repair proteins, containing $[4Fe4S]^{2+}$ clusters, for example EndoIII (green) and MutY (orange), bind DNA, activating them towards oxidation to the $[4Fe4S]^{3+}$ state. The sequence of events is as follows: Guanine radical formation can oxidize a repair protein in a DNA-mediated reaction, stabilizing the oxidized protein bound to DNA (*a*). A second protein binds in the vicinity of the first protein (*b, e*). CT to a distally bound protein can occur through the DNA π-stack if the intervening DNA is undamaged (*c, f*). The newly reduced protein has a diminished affinity for DNA and diffuses away (*d*). If, instead, a lesion site is present between the proteins (*g*), the DNA-mediated CT step is inhibited and the oxidized protein remains bound to DNA. In this search mechanism the sum of the DNA-mediated electron transfer steps between proteins constitutes a full search of the genome yielding a redistribution of low-abundance DNA repair proteins in the vicinity of lesions.

injection over very long distances with DNA occupation, and also rely on hole absorption from oxidized 8-oxoguanine, a species known to be highly unstable to oxidation. Since it is not clear whether this is mechanistically allowed, we will take a substantially different approach in this work, limiting our study to CT signaling between proteins, with the DNA only acting as a mediator.

To exploit DNA-mediated CT, some proteins must exist in the oxidized state. There are many oxidants in the cellular milieu, and the level of oxidative stress will govern the proportion of oxidized protein. Indeed, we have shown that these proteins[23] and others[47] can be oxidized by guanine radicals, the first genomic signal of oxidative stress,[48] via DNA-mediated CT. There is also computational support for this activation being facile.[49]

We can calculate the step time for three-dimensional diffusion of the reduced protein to the DNA from

$$T_{Search} = \ln\left(\frac{1}{1-P}\right)\left(k_{CT}\Theta C_p\right)^{-1}$$

where k_{CT} is the rate of search using CT signaling and the other parameters are as defined above. Each oxidized protein provides a separate nucleation site for CT scanning, but draws from the same reservoir of reduced protein to scan different portions of the genome; hence the number of proteins is multiplied by Θ, the fraction of protein that is oxidized. For diffusive search,

$$k_{CT} = 3D_{CT}\left(\frac{2}{L}\right)^2$$

$$D_{CT} = \left(\frac{N}{4}\right)^2 t_s^{-1}$$

where D_{CT} is the effective diffusion of holes using proteins as steps, L is the length of the genome, N is the maximum distance for CT signaling, and t_s is the time for colocalization between the

oxidized and reduced proteins (effectively, the step time). The step length is a quarter of N due to a factor of two from the average yield of self-exchange between the proteins, and another factor of two since the average step length in this case will be half of the maximum step length.

The step time can be found from a modified Debye-Smoluchowski equation for protein collision with a rod of DNA within the cell volume, where the length of the rod is twice the number of bases, N, over which DNA-mediated CT can proceed, since reduced protein can transfer an electron from either side:

$$k_{assoc} = \frac{V}{C_p(1-\Theta)t_s} = 4\pi\kappa f D_3\left(r_{DNA} + r_p\right) + D_a a^3\sqrt{\frac{r_{DNA} + r_p}{a}}$$

$$r_{DNA} = \frac{2(N + l_{s,red})}{2\ln\left(\frac{2(N + l_{s,red})}{R_{DNA}}\right)}$$

$$D_a = \frac{kT\ln\left(\frac{a}{R_{DNA}}\right)}{3\pi\eta a}$$

where V is the cell volume, $C_p(1-\Theta)$ is the number of reduced repair proteins in the volume, r_{DNA} is the DNA radius, r_p is the Stokes radius of the protein, D_a is the segmental diffusion constant, and k_{assoc} is the bimolecular rate constant for protein association with the DNA target within the cellular volume. Any contact of the reduced protein within the DNA rod allows electron transfer to the DNA-bound oxidized protein. We also allow the reduced protein to slide to within this region, although the sliding length for the weakly associated reduced protein is negligible. The Smoluchowski equation is constructed with two terms: one describes the ballistic 3-dimensional diffusion of the reduced protein to the DNA and the second[50] considers the gyrations of a rod with a persistence length of 150 base pairs and the ends fixed as part of the chromosome. The translational DNA diffusion is considered to be negligible. The electrostatic (f) and orientational

(κ) constants are taken as unity,[51] in keeping with the high ionic strength of the *in vivo* environment. The dissociation rate of the protein is not included in our model because charge equilibration should occur on a much faster timescale than dissociation of the reduced protein. This time, for reasonable parameters ranges, varies from 0.5 ms to 2 ms, which corresponds to a sliding length of between 50 and 100 base pairs. This is similar to the distance between bound structural protein on the DNA; CT signaling minimizes the redundancy of search by the oxidized protein between redox events. Note that we make no distinction between 5' to 3' versus 3' to 5' transport, although subtle differences have been observed.[52]

Scanning through sliding/jumping without a CT search represents a boundary condition, so that the total time is

$$T = (T_{CT}^{-1} + T_D^{-1})^{-1}$$

where T_D is the diffusion scanning time, calculated as in Section 4.2.2.

In our model, the DNA is essentially scanned by the electron with the repair proteins facilitating electron migration. Thus we calculate a genome scanning time for MutY in *E.coli* that is significantly more efficient through DNA CT. Since an injected charge equilibrates on the nanosecond timescale,[51] and protein diffusion occurs in micro- to milliseconds, the rate-limiting step in this process is the 3D diffusion of this reduced protein to within CT range of the oxidized DNA-bound protein.

Importantly, since this model involves cooperation among the repair proteins, we can utilize the *total* concentration of these proteins within the cell, rather than copy numbers for MutY or EndoIII individually. Thus MutY, present in 20 copies, benefits from 500 copies of EndoIII.[1] We do, however, neglect contributions from any other proteins that might participate in DNA-mediated signaling; other DNA-bound proteins containing iron-sulfur clusters exhibit similar

potentials, and CT reactions involving these proteins too would substantially speed the search process.

Our model relies on the fact that DNA-mediated interprotein CT is much faster than protein diffusion, and that the oxidized repair proteins have higher nonspecific DNA affinity than the reduced proteins; both assumptions have experimental support.[13,27] One advantage of DNA CT over other search mechanisms is that the electron travels *through* the DNA base pairs and no proteins need to be displaced.[17,53]

Figure 4.2 shows how the interrogation time varies as a function of *N*, the maximum distance over which DNA-mediated CT proceeds, and ox, the percentage of proteins oxidized. Remarkably, permitting DNA CT over 500 bp with 10% oxidized protein yields a conservative interrogation time of 30 minutes, while DNA CT over 500 bp with 20% oxidized protein yields an interrogation time of 17 minutes; permitting DNA CT over 1500 bp yields scan times of about a minute. These values are well within the one hour doubling time of normally growing *E. coli*. While we have not yet established the distance limits for DNA CT, we have demonstrated substantial oxidative damage in tethered DNA assemblies *in vitro* over 60 bp and in DNA within mitochondria over ~100 bp.[14,54]

The dependence of interrogation time on the percentage of proteins oxidized is also noteworthy (**Figure 4.3**). The scanning efficiency resembles a switch that is turned on at low levels of oxidation, when DNA repair is needed. Activation of this switch depends upon the redox buffering capacity of the cell and the level of oxidative stress. This local activation of MutY by oxidized DNA has been supported by theoretical calculation.[48] Furthermore, there might be other redox-active DNA repair proteins in *E. coli*. If other proteins can participate in

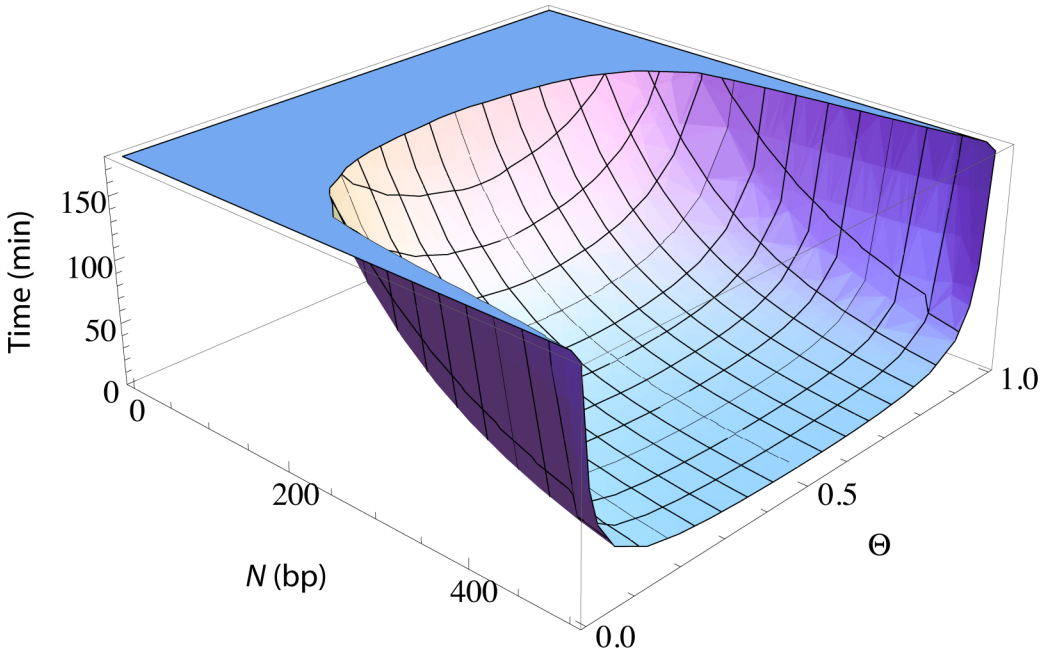

Figure 4.2. Scanning time as a function of maximum distance of DNA-mediated interprotein CT (N) and the fraction of repair proteins that are in the 3+ state (Θ) using the CT scanning model. At 10% oxidized protein with a maximum CT distance of 500 bp, the time required to interrogate the genome is ~5 minutes.

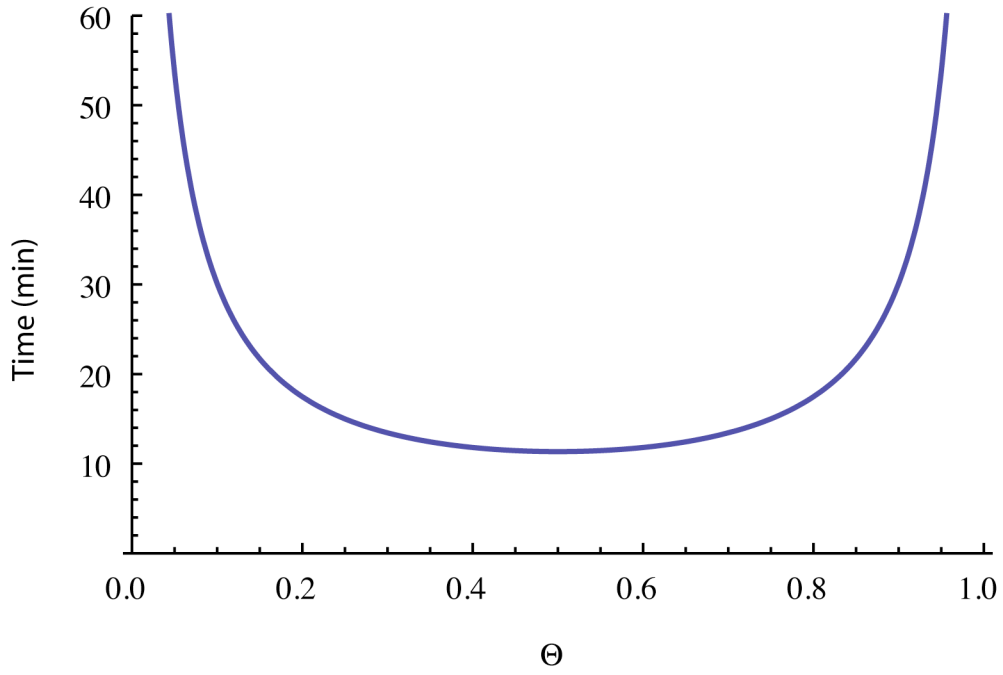

Figure 4.3. For N = 500, the search time using CT shows a sharp drop when a few proteins become oxidized. If all the protein is oxidized or reduced, then CT signaling is no longer possible.

helper function, search times would also rapidly decrease; the search rate increases with the square of searching protein.

We have not taken into account the effect of the reduced protein directly repairing damage. Besides our uncertainty in the reduced protein diffusion rate by an order of magnitude, a further concern is the fact that the specificity of the reduced protein has never been measured. For the weak binding of the reduced species, changes in specificity can change search time substantially. Experimentally determined specificities[4,26] have been measured using mutated protein and base analogues, and under conditions where even a small proportion of oxidized protein will determine the measurement. It has not been demonstrated whether the reduced protein has catalytic activity.

4.2.4. DISTRIBUTION OF ENDOIII BETWEEN TWO STRANDS OF DNA

The above mechanism suggests protein accumulation along a single DNA at a lesion site, which is the genomically relevant case. Recently, the Barton group has begun to explore EndoIII distribution between two different types of long strands: those with fully matched DNA, and those containing a single non-cognate lesion. The relative protein affinities are measured by AFM, with matched (2.2 kb, 1.6 kb) distinguished from mismatch-containing DNA (3.8 kb) on the basis of length. EndoIII accumulates on the mismatch-containing strand, in a manner dependent on the CT competence of the protein. In the context of the above model, one might not expect EndoIII to distinguish between these two types of DNA. After all, CT signaling does not affect the number of reduced and oxidized proteins on a given strand of DNA, it only changes their positions along the strand. How can a change in distribution along strands affect the distribution between strands?

A protein bound to genomic DNA has only one pathway for dissociation. On linear DNA, however, protein can dissociate by falling off, or by sliding off the ends.[55] How important these mechanisms are depends on the nature of the protein-DNA interaction, and the relative size of the strand length versus the sliding length. There is experimental evidence to support end-sliding as a dissociative mechanism for some proteins, but not for others. If end-sliding is an important component of the dissociation of EndoIII from linear DNA, then redistribution of the protein to a central lesion will increase the overall binding affinity of the protein to the DNA relative to the fully matched control.

The ratio of protein between the long and short strands is:

$$R_{L,S} = \frac{\#_L}{\#_S} = \frac{K_L}{K_S} = \frac{k_{a,L}}{k_{a,S}} \frac{k'_{d,S}}{k'_{d,L}}$$

where # is the number of proteins counted on a given strand of DNA, the subscripts *L* and *S* identify long and short DNA respectively, and k'_d represents the composite dissociation rate of protein from the DNA (including end-sliding). The rate of end-sliding is:

$$k_e(n) = \frac{D_1}{n^2} = \frac{k_d l_s^2}{n^2}$$

where *n* is the distance to the nearest end, if we assume that reflection at the ends is negligible and that the DNA is long enough that end-sliding is negligible for protein in the center of the DNA. The latter will clearly be the case for several thousand base-pair long DNA. In this case,

$$k'_d(n) = k_d\left(1 + \frac{l_s^2}{n^2}\right)$$

$$k'_d = 2k_d \sum_{1 \le n \le \frac{N}{2}} \phi_n \left(1 + \frac{l_s^2}{n^2}\right); \; \sum_n \phi_n = 1$$

where N is the length of the strand and ϕ_n is the protein distribution on the DNA. We can approximate ϕ_n as uniform along the DNA.

For the association constants,[8]

$$k_a = 4\pi D_3 r_g \left(1 - \frac{\tanh(\eta r_g)}{\eta r_g}\right)$$

$$r_g = \sqrt{\frac{Na}{3}}$$

where r_g is the radius of gyration, D_3 is the three-dimensional diffusion constant, a is the persistence length, and η is a parameter describing the geometry of the strand. That leaves a single parameter in the expression for $R_{L,S}$: the sliding length l_s.

For the case where all the protein is matched, and the sliding length is 22 bp, $R_{L,S}$ is expected to be about 1.8, which corresponds to a binding density ratio ($r \equiv R_L * N_S/N_L$) of 0.9. However, if a fraction φ of the protein is segregated in the middle of a long strand due to accumulation at a mismatch, then

$$k'_{d,MM} = 2k_d \sum_{1 \le n \le \frac{N}{2}} \phi_n \left(1 + \varphi \frac{l_s^2}{n^2}\right).$$

This provides a basis for justifying the preferential affinity of EndoIII for DNA containing a non-cognate mismatch over a strand that is fully matched.

4.3. EXPERIMENTAL VALIDATION OF CT SIGNALING

4.3.1. AFM MEASUREMENTS OF DISTRIBUTION OF ENDOIII BETWEEN DNA STRANDS

While we have earlier carried out studies establishing hole and electron injection across the protein/DNA interface,[22-24] our model also predicts that DNA/protein CT would promote the redistribution of repair proteins in the vicinity of base lesions or mismatches. We can assay for

this redistribution by AFM. A mixture of DNAs, both long (3.8 kilobase) DNA duplexes containing a single CA mismatch and short (2.2 and 1.6 kilobase) well-matched duplexes of the same total sequence were prepared;[56] the longer sequence was obtained by ligation of the two shorter sequences. This mixture of matched and mismatched DNA strands was incubated with EndoIII and examined using established AFM techniques[57] (**Figure 4.4**). Only clearly identifiable long or short strands were counted. Protein assignments were verified through analysis of their 4 nm heights in the images; without protein, features of this dimension are not observed and still larger heights indicate salt precipitates. Although a CA mismatch effectively inhibits DNA CT,[13] it is not a lesion that is preferentially bound by EndoIII; a gel shift assay on 21-mers with and without a central CA mismatch shows no detectable difference in EndoIII binding. Thus without DNA CT between bound EndoIII molecules, one might expect an equal density of proteins on the short and long strands.

We find that EndoIII shows a significant preference for the longer strands containing the CA mismatch. Examination of the number of proteins bound to 300 long strands and 465 short strands reveals a greater density of proteins bound to the long strand; r(long/short) is 1.6. If instead we examine the distribution of EndoIII on long versus short strands, where all strands are matched, we see a small preference for the short strands; the ratio of protein densities, r(long/short), is 0.9. When we calculate the strand preference based on DNA CT, this protein density ratio depends upon the DNA CT length and/or the length of the DNA over which protein can diffuse before dissociating. Using a signaling/sliding length of 90 base pairs, we calculate a protein density ratio of 1.6, that which we find, where half of the protein population is near the mismatch.

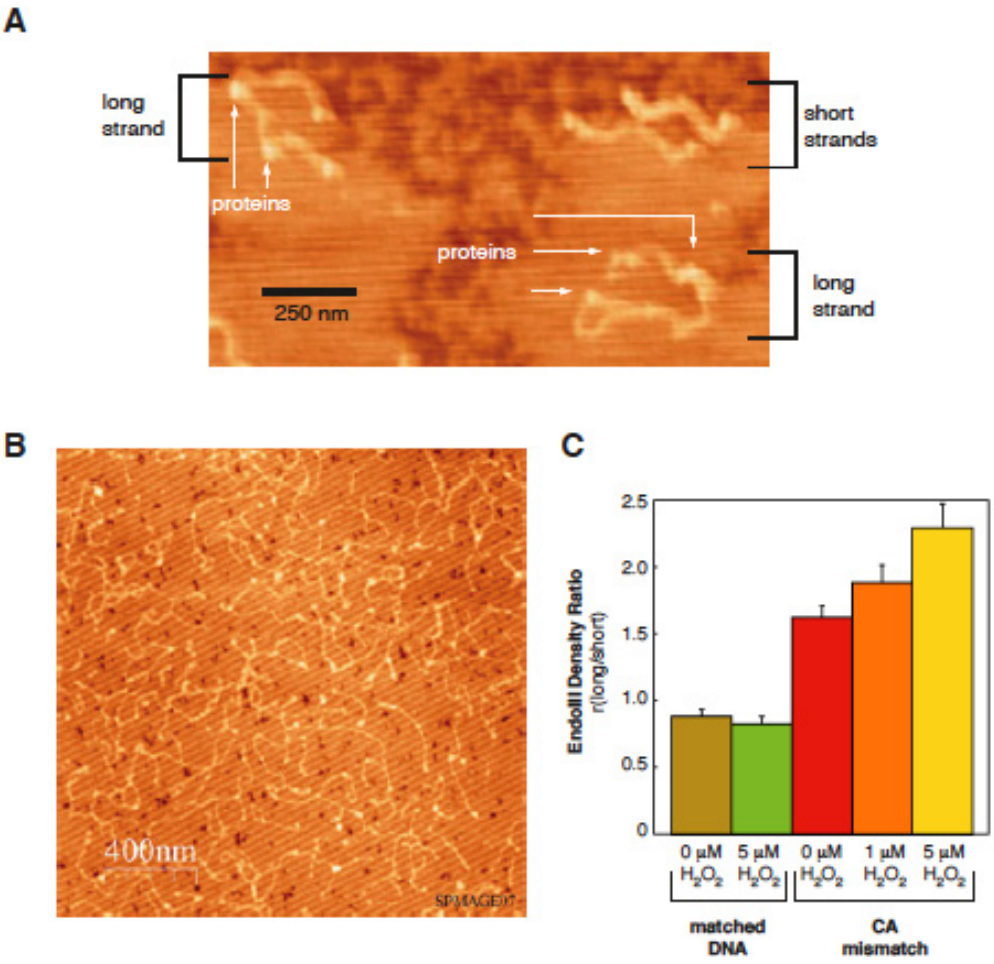

Figure 4.4. Measurements of repair protein distributions on DNA by AFM. A zoomed-in view (**A**) and a zoomed-out view (**B**) of representative AFM images of DNA strands incubated overnight with wild-type EndoIII. A higher density of proteins is apparent on the longer DNA strands containing the single base CA mismatch. Densities of 0.19 proteins bound per kilobase long strand and 0.12 proteins bound per kilobase short strand are observed, giving a density ratio (r(long/short)) of 1.6; the uncertantity is ≤ 10%. (**C**) Quantitation of protein density ratios. A CA mismatch is contained on the long strand except for the sample indicated by matched DNA, where both the long and the short strands are fully matched. EndoIII redistributes onto the strand with the CA mismatch and that preference is increased with increasing concentrations of peroxide.

AFM measurements as a function of oxidation of proteins bound to DNA, using H_2O_2 as oxidant, reveal an additional increase in the ratio of EndoIII bound to mismatch-containing strands. Examination of more than 250 long CA mismatch-containing strands and 300 shorter matched strands incubated with EndoIII and treated with 5 μM peroxide reveals a ratio of bound protein densities, r(long/short), of 2.4; when both long and short strands are matched, the ratio is 0.83.

These results are consistent with our model. DNA-mediated CT will drive the redistribution of repair proteins away from undamaged regions such that the proteins will cluster near damaged sites. As a result, we see the proteins redistribute preferentially onto the DNA strands containing the mismatch even though a CA mismatch is not a substrate for EndoIII. Moreover, as predicted by the model, the redistribution of EndoIII is more pronounced in the presence of oxidative stress.

4.3.2. TRANSVERSION ASSAYS IN *E. COLI* AND CT SIGNALING

This CT scanning model was tested *in vivo* by assaying for the cooperation among repair proteins facilitated by DNA-mediated signaling. If these proteins are able to help each other in their search for damage using DNA CT, knocking out the gene for EndoIII or reducing its capability to carry out CT should lead to a decrease in MutY activity *in vivo*. Assays for MutY and EndoIII activity inside *E. coli* cells have already been developed.[58] The assay for "helper function" used here employs engineered mutations in the *lacZ* gene to report the frequency of a particular base-pair substitution. The strain that serves as an assay for MutY activity, CC104, substitutes a cytosine for an adenine in the *lacZ* Glu 461 codon, which is essential for β-galactosidase activity. Since MutY prevents GC to TA transversions,[59] reversion of this original

mutation back to wild-type (wt) *lacZ* reflects a deficiency in MutY activity. Analogously, the CC102 strain[57] serves as an assay for EndoIII activity by monitoring GC to AT transitions.[60,61]

In the CC104 MutY activity reporter strain (**Table 4.1**), 20 ± 9 *lac*+ revertants are observed per 10^9 cells, while inactivation of *mutY* in CC104 (CC104 *mutY*-) causes the number of *lac*+ revertants to increase 15x (300 ± 33) as expected.[58,59] When the gene encoding EndoIII (*nth*) is inactivated in CC104 (CC104 *nth*-), the *lac*+ reversion frequency observed is 54 ± 5, more than a factor of two increase over CC104. Thus, loss of EndoIII does have a small but significant effect on MutY activity *in vivo*. This loss in activity is consistent with a loss in helper function by EndoIII, as predicted; the lower activity of MutY without EndoIII could reflect the lack of cooperative searching via DNA CT. An alternative explanation, however, is that MutY and EndoIII share some overlapping ability to repair lesions. In this case, the *lac*+ reversion frequency of the CC104 *mutY-/nth*- strain (270 ± 29) should be greater than that of CC104 *mutY*, but they are, within error, equivalent.

This *in vivo* relationship between EndoIII and MutY has been observed previously, although in different experimental contexts. Small increases in mutational frequency have been detected when *mutY* is inactivated in CC102,[58] as was also observed here, or when *nth* is inactivated in CC104.[61] In the latter case, it was proposed that this could be due to some intrinsic ability of EndoIII to repair oxidatively damaged guanine residues. Reported EndoIII repair activities do not prevent GC to TA transversion mutations[62] and, thus, are not relevant to the CC104 assay.

We can furthermore test directly whether the loss of MutY activity in the CC104 assay is the result of overlapping glycosylase activities by determining whether the number of *lac*+ revertants is still suppressed by an EndoIII mutant that is biochemically incompetent to carry out

Table 4.1. Assay for DNA repair in *E. coli* by MutY (CC104)

Strain	*Lac*+ Revertants[a,b]	Increase
	(*lac*$^+$ colonies/ 10^9 cells)	(x/CC104)
CC104[c]	20 ± 9	---
CC104 *nth-*	54 ± 5	2.7
CC104 *mutY-*	300 ± 33	15
CC104 *mutY-/nth-*	270 ± 29	13.5

a. *Lac*$^+$ revertants are reported as the average number of *lac*$^+$ colonies that arise per 10^9 cells plated on minimal lactose media.

b. These data represent a single set of experiments with 10 replicates per strain assayed concurrently. Values reported as the mean ± s.d.

c. CC104 strains reflect the rate of GC to TA transversion mutations and serve as a reporter for MutY activity in *E. coli*.

the glycosylase reaction. A mutant of EndoIII (D138A) that is known to be deficient in glycosylase activity[63] was introduced on a plasmid into both the CC102 and CC104 strains along with appropriate vector controls. Because this mutant cannot perform the base excision reaction, D138A fails to reduce the high reversion frequency observed with CC102 *nth*- . However, D138A is able to complement the CC104 *nth*- strain. Thus, the glycosylase activity of EndoIII is not required for its helper function to aid MutY in repairing lesions inside the cell. Nonetheless, it appears that EndoIII lacking D138 can bind DNA and contains an intact [4Fe4S] cluster.[63] Based upon our model, D138A should be competent to carry out DNA-mediated CT and thus serve as a helper to MutY, as we observe.

In our model, it is the ability to carry out DNA-mediated CT, not the glycosylase activity of EndoIII, that is critical to its helper function. Thus, perturbing the path for electron transfer to the DNA would interfere with this helper function. Aromatic tyrosine and tryptophan residues often facilitate long-range electron transfers in proteins,[64,65] and EndoIII contains many of these residues. In particular, Y82 is conserved in most EndoIII and MutY homologues,[66] and an analogous mutation (Y166S) in the human homologue of MutY is associated with cancer.[67] In the crystal structure, Y82 is located very close to the DNA backbone.[5] Y82A EndoIII was thus introduced on a plasmid into both reporter strains (CC102 and CC104) and their *nth* knockouts to explore whether this mutation attenuates helper function. Significantly, Y82A in the CC104 *nth*- strain shows an increase in mutation rate versus the CC104/Y82A and CC104/p controls (**Figure 4.5**). The number of *lac*+ revertants is found to increase by 53 ± 16% when comparing CC104 *nth*-/ Y82A to CC104/p. When comparing CC104 *nth*-/Y82A to CC104/Y82A, the number of *lac*+ revertants increases by 68 ± 13%. Similarly, for these trials, the ratio of the

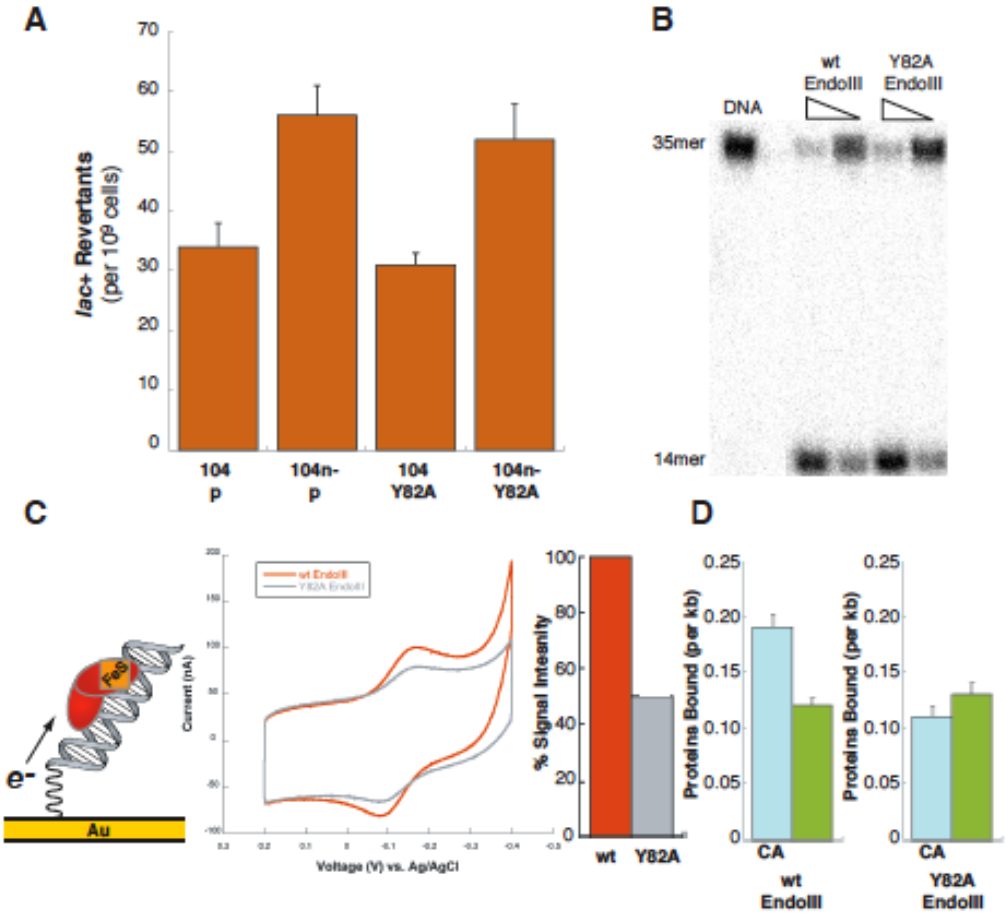

Figure 4.5. Y82A EndoIII, a mutant in DNA-mediated CT capability. (**A**) Bar graph showing *lac*+ revertants for CC104/p, CC104 *nth*-/p, CC104/Y82A and CC104 *nth*-/Y82A strains, where p denotes inclusion of an empty vector. *Lac*+ revertants are reported as the average number *lac*+ colonies that arise per 10^9 cells plated on minimal lactose media containing ampicillin. Data for the CC104 strains are shown based upon five sets of independent experiments, each containing 10 replicates per strain. (**B**) Autoradiogram after denaturing PAGE of ^{32}P-5'-TGTCAATAGCAAG**X**GGAGAAGT-CAATCGTGAGTCT-3' + complementary strand where X = 5-OH-dU base-paired with G. Protein samples (100 or 10 nM) were incubated with duplexes for 15 min at 37 °C and quenched with 1 M NaOH. Cleavage of the ^{32}P-labeled strand at the lesion site (**X**) by EndoIII results in formation of a 14mer. No significant difference in glycosylase activity (10% uncertainty) is observed between Y82A and wt EndoIII. (**C**) Cyclic voltammetry of Y82A EndoIII at a Au electrode modified with $SH(CH_2)_2CONH(CH_2)_6NHOCO$-5'-AGTACAGTCATCGCG-3' + complementary strand showing the reduction and reoxidation of the DNA-bound protein. DNA-modified surfaces were prepared, backfilled with mercaptohexanol, and wt or Y82A EndoIII was tested. Surfaces were then rinsed and the other protein analyzed on the same surface. Over several trials, the electrochemical signal associated with Y82A is 50 ± 13% smaller per [4Fe4S] cluster compared with wt EndoIII, reflecting poor electronic coupling of the mutant to the DNA-modified electrode. (**D**) Comparative densities for wt (left) and Y82A (right) EndoIII bound to matched versus mismatched (CA) strands measured by AFM. Although wt EndoIII preferentially redistributes onto the mismatched strand, Y82A shows no preference.

number of *lac*+ revertants for CC104 *nth*-/ p versus CC104/p is 165 ± 13%. These results clearly indicate that Y82A does not restore helper function.

It is noteworthy that inclusion of Y82A EndoIII in CC102 *nth*- leads to a diminished mutation rate, indicating that this mutant is competent for EndoIII activity inside the cell. Interestingly, the observation that Y82A complements CC102 *nth*-, but not CC104 *nth*-, is consistent with the conclusion that the glycosylase activity of EndoIII is not a source of helper function. Moreover the fact that Y82A complements CC102 *nth*- is understandable in the context of our model, because of the higher copy number of EndoIII in *E. coli* cells than MutY. In our model, without oxidative stress, we would predict that DNA CT is not essential for EndoIII repair activity inside the cell. We would therefore anticipate that the role of EndoIII in helping MutY search for lesions may be more important than the ability of EndoIII to find its own lesions. This distinction becomes more complex when considering that other DNA-binding proteins with iron-sulfur clusters might also participate in the signaling process.

To establish the biochemical characteristics of Y82A EndoIII, the protein was purified and its redox and glycosylase activities examined. Importantly, the mutant enzyme does contain the [4Fe4S] cluster, characterized by its distinctive absorbance spectrum. Y82A EndoIII also maintains glycosylase activity against a 5-OH-dU lesion in a ^{32}P-5'-endlabeled 35-mer duplex (**Figure 4.5**); the activity of the mutant in this assay is equal to that of wild type. Note that this experiment on a 35-mer duplex measures only the base excision reaction, not the search process. Similarly, in the *E. coli* EndoIII activity assay, where we expect that the search process is not rate-limiting, Y82A EndoIII activity is comparable to that of wild-type EndoIII. In contrast, D138A EndoIII, which instead inhibits the base excision reaction, fails to complement the *nth*

knockout in the EndoIII activity reporter strain but does complement the *nth* knockout in the MutY activity reporter strain, where lesion detection is limiting.

To test for DNA-bound redox activity, Y82A was examined on a Au electrode modified with thiol-terminated DNA duplexes. Significantly, in the cyclic voltammogram, the potential for the DNA-bound mutant resembles that of the wild type,[12] but the signal intensity is diminished (**Figure 4.5**). The protein concentrations are determined based on the 410 nm absorbance of the [4Fe4S] cluster; the smaller electrochemical signal observed with Y82A does not reflect a lower concentration of [4Fe4S] clusters. Over several trials, Y82A EndoIII exhibits a signal that is 50 ± 13% smaller than that for wt EndoIII (per [4Fe4S] cluster). This signal intensity provides a reliable measurement of reduction/oxidation of the DNA-bound protein. Since the glycosylase activity on the 35-mer is equal for the mutant and wild type, this diminished signal cannot reflect diminished binding of the mutant to the DNA. Instead this lowered signal intensity would be expected with an attenuated efficiency of CT from the cluster to DNA and reflects poor electronic coupling of the mutant with the DNA duplex. These results therefore indicate that Y82A EndoIII is defective in DNA-mediated signaling.

Significantly, and consistent with these results, examination of the distribution of Y82A on mismatched and matched strands by AFM shows no preference for the mismatched strand; we observe 0.11 proteins per kilobase long strand and 0.13 proteins per kilobase short strand (**Figure 4.5**). In fact, the ratio of protein densities on mismatched versus matched strands with Y82A, r(long/short) is 0.9, essentially equal to that of wild-type EndoIII bound to fully matched long versus short strands. Since the Y82A mutant, biochemically defective only in DNA CT, cannot redistribute to the vicinity of the lesion, DNA CT must play a role in finding the lesion both in the AFM experiment and in the helper function assay. These results together demonstrate

a distinct connection between DNA-mediated CT to the [4Fe4S] cluster, the detection of DNA defects, and the *in vivo* relationship observed between MutY and EndoIII.

4.4. DISCUSSION

These experiments indicate that MutY and EndoIII cooperate in their search for damage in the genome and redistribute in the vicinity of lesions consistent with CT scanning. This cooperation, or helper function, does not involve the glycosylase reaction. Based on their chromosomal arrangement, the expression of MutY and EndoIII, furthermore, do not appear to be linked.[68] There is also no chemical evidence that the proteins physically bind to one another, and their low abundance within the cell makes random associations improbable. This cooperation thus arises from a distance. Importantly, what does appear to be required for helper function is an intact [4Fe4S] cluster as well as an electroactive protein-DNA interface. Mutation of an aromatic amino acid residue near the DNA binding site, Y82A, leads to a decrease in CT efficiency *in vitro*, the inability of the protein to redistribute near lesions by AFM, and diminished helper function *in vivo*. These experiments thus establish a link between DNA-mediated CT and the cooperative search for damage by these repair proteins both *in vitro* and *in vivo*.

BER glycosylases are known to prevent mutations inside the cell, yet in most organisms, these enzymes are not required for normal growth and development.[2] Recently it was discovered that germline mutations in human BER homologues result in a genetic predisposition to cancer.[67] Specifically, the human homologue of *mutY* (*MUTYH*) is found mutated in a subset of patients predisposed to colorectal cancer. Many of the cancer-associated mutations in *MUTYH* are missense, or single amino acid, mutations. Though several of the most common mutants have

been characterized biochemically, it remains unclear exactly how these variants lead to disease. Given that initial detection of lesions is likely the rate-limiting step in BER,[7] it is possible that mutants with defects in protein-DNA CT would be associated with cancer. Indeed, many of these *MUTYH* missense mutations found in colorectal cancer patients result in loss or gain of aromatic residues near predicted protein-DNA interfaces.[67] Significantly, *MUTYH* contains two adjacent tyrosine residues (Y165 and Y166) that closely align with Y82 in *E. coli* EndoIII and inherited mutations in these *MUTYH* residues (Y165C and, less commonly, Y166S) are clinically relevant in cancer. These results thus provide tantalizing evidence for association between defects in lesion detection via DNA-mediated CT by BER enzymes and human disease.

Iron-sulfur clusters are becoming increasingly ubiquitous to proteins that repair, replicate, and transcribe DNA.[69,70] Recent characterizations of archaeal DNA primase, RNA polymerase, and nucleotide excision repair helicase (XPD) homologues reveal an iron-sulfur cluster required for normal enzyme function. Though the precise role of the cluster in these proteins is unclear, the cysteine residues ligating the cluster are conserved in eukaryotic homologues of these proteins. In archaeal XPD, moreover, the iron-sulfur cluster occupies a site far from the ATP hydrolysis domain but implicated in DNA binding.[69] It is interesting to consider whether in these proteins, as in BER enzymes, the iron-sulfur cluster is poised to send and receive redox signals mediated by the DNA helix, which may, in turn, modulate DNA binding affinity, enzyme activity, or protein structure. Such long-range signaling among proteins bound to DNA would make searching for lesions much more efficient and may generally provide a means of genome-wide communication to monitor cellular stresses.

DNA-mediated CT serves as a fast and efficient reaction that is exquisitely sensitive to lesions in the base-pair stack. This chemistry helps to explain how these repair glycosylases locate

their lesions efficiently in the cell, a key function since mutations in these enzymes in humans are implicated in colorectal cancer.[67] This mechanism furthermore provides a rationale for iron-sulfur clusters in DNA repair proteins. More generally, these experiments illustrate the importance of DNA-mediated CT in long-range signaling among proteins in low abundance that are bound to DNA. Other roles for DNA-mediated CT in biological signaling must now be considered.

4.5. REFERENCES

1. Demple B.; Harrison L. *Annu. Rev. Biochem.* **1994**, *63*, 915.

2. David S.S.; O'Shea V.L.; Kundu S. *Nature* **2007**, *447*, 941.

3. Bai, H.; Lu, A.-L., *J. Bacter.* **2007**, *189*, 902.

4. Francis A.W.; Helquist S.A.; Kool E.T.; David S.S. *J. Am. Chem. Soc.* **2003**, *125*, 16235.

5. Fromme J.C.; Verdine G.L. *EMBO J.* **2003**, *22*, 3461.

6. Fromme J.C.; Banerjee A.; Huang S.J.; Verdine G.L. *Nature* **2004**, *427*, 652.

7. Livingston A.L.; O'Shea V.L.; Kim T.; Kool E.T.; David S.S. *Nat. Chem. Biol.* **2008**, *4*, 51.

8. Berg O.G.; Winter R.B.; von Hippel P.H. *Biochemistry* **1981**, *20*, 6929.

9. Parker, J.B.; Bianchet, M.A.; Krosky, D.J.; Friedman, J.I.; Amzel, L.M.; Stivers, J. *Nature* **2007**, *449*, 433.

10. 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. *Biochemistry* **1989**, *28*, 4450

11. Boon E.M.; Livingston A.L.; Chmiel N.H.; David S.S.; Barton J.K. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12543.

12. Boal, A.K.; Yavin, E.; Lukianova, O.A.; O'Shea, V.L.; David, S.S.; Barton, J.K. *Biochemistry* **2005**, *44*, 8397.

13. Wagenknecht, H. A., Ed. *Charge Transfer in DNA*; Wiley-VCH: Weinheim, **2005**.

14. Núñez M.E.; Hall D.B.; Barton J.K. *Chem. Biol.* **1999**, *6*, 85.

15. Kelley S.O.; Jackson N.M.; Hill, M.G.; Barton, J.K. *Angew. Chem. Int. Ed.* **1999**, *38*, 941.

16. Elias B.; Shao F.; Barton J.K. *.J Am. Chem. Soc.* **2008**, *130*, 1152.

17. Núñez M.E.; Noyes K.T.; Barton J.K. *Chem. Biol.* **2002**, *9*, 403.

18. Núñez M.E.; Holmquist G.P. Barton J.K. *Biochemistry* **2001**, *40*, 12465.

19. Boal A.K.; Barton J.K. *Bioconjugate Chem.* **2005**, *16*, 312.

20. Boon E.M.; Ceres D.M.; Drummond T.G.; Hill M.G.; Barton J.K. *Nat. Biotechnol.* **2000**,*18*, 1096.

21. Guo X.F.; Gorodetsky A.A.; Hone J.; Barton J.K.; Nuckolls C. *Nat. Nanotechnol.* **2008**, *3*, 163.

22. Boon E.M.; Salas J.E.; Barton, J.K. *Nat. Biotechnol.* **2002**, *20*, 282.

23. 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. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3546.

24. Yavin, E.; Stemp, E.D.A.; O'Shea, V.L.; David, S.S.; Barton, J.K. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *3610.*

25. Wunderlich Z.; Mirny L.A. *Nuc. Acids. Res.* **2008**, *36*, 3570.

26. Porello, S.L.; Williams, S.D.; Kuhn, H.; Michaels, M.L.; David, S.S. *J. Am. Chem. Soc.* **1996**, *118*, 10684.

27. Gorodetsky A.A.; Boal A.K.; Barton J.K. *J. Am. Chem. Soc.* **2006**, *128*, 12082.

28. Halford, S.E.; Marko, J.F. *Nuc. Acids Res.* **2004**, *32*, 3040.

29. Slutsky M.; Mirny L.A. *Biophys. J.* **2004**, *87*, 4021.

30. Bénichou, O.; Kafri, Y.; Sheinman, M.; Voituriez, R. *Phys. Rev. Lett.* **2009**, *103*, 138102.

31. Murugan, R. *Phys. Rev. E* **2009**, *79*, 061920.

32. Kolesov, G.; Wunderlich, Z.; Laikova, O.N.; Gelfand, M.S.; Mirny, L.A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 13948.

33. Elf, J.; Li, G.-W.; Xie, X.S. *Science*, **2007**, *316*, 1191.

34. Blainey P.C.; van Oijen A.M.; Banerjee A.; Verdine G.L.; Xie X.S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 5752.

35. This sliding is only pseudo-one dimensional, consisting of rotational motion around the twisted double helix. See Bagchi, B.; Blainey, P.C.; Xie, X.S. *J. Phys. Chem. B* **2008**, *112*, 6282.

36. Ali Azam, T.; Iwata, A.; Nishimura, A.; Ueda, S.; Ishihama, A. *J. Bacteriol.* **1999**, *181*, 6361.

37. Li, G.-W.; Berg, O.G.; Elf, J. *Nat. Phys.* **2009**, *5*, 294.

38. Mullineaux C.W.; Nenninger A.; Ray N.; Robinson C. *J. Bacteriol.* **2006**, *188*, 3442.

39. Asahara H.; Wistort P.M.; Bank J.F.; Bakerian R.H.; Cunningham R.P. *Biochemistry* **1989**, *28*, 4444.

40. Hu, T.; Grosberg, A.; Shklovskii, B. *Biophys. J.* **2006**, *90*, 2731.

41. Thanbichler M, Wang SC, Shapiro L (2005) The bacterial nucleoid: a highly organized and dynamic structure. *J Cell Biochem* 96:506-521.

42. Flyvbjerg, H.; Keatch, S.A.; Dryden, D.T.F *Nuc. Acids Res.* **2006**, *34* 1550.

43. O'Handley S.; Scholes C.P.; Cunningham R.P. *Biochemistry* **1995**, *34*, 2528.

44. Slutsky M.; Mirny L.A. *Biophys. J.* **2004**, *87*, 4021.

45. Fok, P.-W.; Guo, C.-L.; Chou, T. *J. Chem. Phys.* **2008**, *129*, 235101.

46. Fok, P.-W.; Chou, T. *Biophys. J.* **2009**, *96*, 2949.

47. Augustyn K.E.; Merino E.J.; Barton J.K. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18907.

48. Burrows C.J.; Muller J.G. *Chem. Rev.* **1998**, *98*, 1109.

49. Lin, J.C.; Singh, R.R.; Cox, D.L. *Biophys. J.* **2008**, *95*, 3259.

50. Berg O.G. *Biopolymers* **1986**, *25*, 811.

51. Berg O.G.; von Hippel P.H. *Annu. Rev. Biophys. Chem* **1985**, *14*, 131.

52. O'Neill M.A.; Barton J.K. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16543.

53. Núñez, M.E.; Noyes, K.T.; Gianolio, D.A.; McLaughlin, L.W.; Barton, J.K. *Biochemistry* **2000**, *39*, 6190.

54. Merino E.J.; Barton J.K. *Biochemistry* **2007**, *46*, 2805.

55. Surby, M.A.; Reich, N.O. *Biochemistry* **1996**, *35*, 2209.

56. Donahue W.F.; Turczyk B.M.; Jarrell, K.A. *Nucleic Acids Res.* **2002**, *30*, 95.

57. Chen L.; Haushalter K.A.; Lieber C.M.; Verdine G.L. *Chem. Biol.* **2002**, *9*, 345.

58. Cupples C.G.; Miller J.H. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5345.

59. Nghiem Y.; Cabrera M.; Cupples C.G.; Miller J.H. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2709.

60. Kreutzer D.A.; Essigmann J.M. *Proc Natl Acad Sci U.S.A.* **1998**, *95*, 3578.

61. Tano K.;, Iwamatsu Y.; Yasuhira S.; Utsumi H.; Takimoto K *J Radiat Res* **2001**, *42*, 409.

62. Thayer M.M.; Ahern H.; Xing D.; Cunningham R.P.; Tainer J.A. *EMBO J.* **1995**, *14*, 4108.

63. Kuo C.F.; McRee D.E.; Fisher C.L.; O'Handley S.F.; Cunningham R.P.; Tainer J.A. *Science* **1995**, *258*, 434.

64. Stubbe J.; Nocera D.G.; Yee C.S.; Chang M.C. *Chem. Rev.* **2003**, *103*, 2167.

65. Shih, C.; Museth, A.K.; Abrahamsson, M.; Blanco-Rodriguez, A.M.; Di Bilio, A.J.; Sudhamsu, J.; Crane, B.R.; Ronayne, K.L.; Towrie, M.; Vicek, A.; Richards, J.H.; Winkler, J.R.; Gray, H.B. *Science* **2008**, *320*, 1760.

66. Watanabe T.; Blaisdell J.O.; Wallace S.S.; Bond J.P. *J. Biol. Chem.* **2005**, *280*, 34378.

67. Cheadle J.P.; Sampson J.R. *DNA Repair* **2007**, *6*, 274.

68. Gifford C.M.; Wallace S.S. *Nuc. Acids. Res.* **2000**, *28*,762.

69. Fan, L.; Fuss, J.O.; Cheng, Q.J.; Arvai, A.S.; Hammel, M.; Roberts, V.A.; Cooper, P.K.; Tainer, J.A. *Cell* **2008**, *133*, 789.

70. Hirata, A.; Klein, B.J.; Murakami, K.S. *Nature* **2008**, *451*, 851.