REGULATION OF WILD-TYPE AND MUTANT P53 THROUGH DNA-MEDIATED CHARGE TRANSPORT

Thesis by

Wendy Mercer Geil

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2012

(Defended September 29, 2011)

Acknowledgements

I would like to thank my adviser, Professor Jackie Barton, for giving me this project and for the support and encouragement she has given over the years. Also, I thank my committee, Doug Rees, Shu-ou Shan, and Jack Beauchamp for pointing me in the right direction, giving me advice, and brainstorming on various aspects of my project.

Many thanks go out to Maureen Renta for keeping the lab together and encouraging me when things were rough.

I thank Kate Augustyn, Nicolai K. Andersen, and Rebecca Chen (SURF) for p53 camaraderie. Katie Schaefer deserves appreciation for a fresh look at the problem and many hours of brainstorming and experiments. Many thanks go to Eddie Merino, for teaching me how to approach experiments in an organized and meaningful way.

I especially thank Amie Boal and Joey Genereux for teaching me how to write for a scientific audience. I have appreciated the insights that Paul Lee, Marisa Buzzeo, Russ Ernst, John Phillips, Eric Olmon, and Brian Zeglis have given me over the past 6 years. Many thanks go out to Susan Shadle for being the best undergraduate mentor ever. I thank Paul Nelson and the Caltech Toastmasters for help with public speaking.

To my friends Katy Muzikar, Christine Romano, Anna Maltsev, Jenn Stockdill, Pam Sontz, Cindy and Joey Genereux, Melanie Yen, Katie Brenner, Stephanie Johnson, Natalie Muren, and Andy and Lesley Monro: thank you so much friendship and support throughout grad school. This would not have happened without you all.

Most of all, I thank my parents, Tom and Janice Mercer, sister LeeAnn Mercer, and husband Ethan Geil for endless support, boundless encouragement, and many, many phone calls. I love you, more than words could ever express.

ABSTRACT

The global transcription factor p53 controls many cellular processes, including the cellular response to oxidative stress. It had been determined that that dissociation of wild type p53 from its promoter site can occur upon DNA-mediated oxidation. In this work, we use site-directed mutagenesis to construct charge-deficient mutants of p53; the chemistry of DNA-mediated oxidation of p53 was examined using these mutants.

The control point for p53 oxidation through DNA-mediated charge transport (DNA CT) is cysteine 275. Using differential thiol labeling and detection of modified peptides with mass spectrometry, we demonstrated that cysteines 124 and 141 in superstable p53 form a terminal disulfide bond upon DNA-mediated oxidation. This leads to a conformational change that inhibits DNA from binding by p53. The disulfide formed between cysteines 124 and 141 is a result of a series of disulfide bond exchange across the protein from the DNA base stack.

We also investigated the dependence of p53 oxidation on DNA sequences. ESMA analysis of biologically derived p53 recognition sequences with varying quantities of guanine doublets and triplets showed efficient p53 oxidation to depend on the presence of low energy GG or GGG sites. Moreover, consistent results were found with biologically derived promoter sequences. Sequence S100A2, with guanine triplets on the same strand, showed the most oxidation of p53, followed by ODC1 and caspase-1. We confirmed these sequence-specific effects by measuring the change in expression level of the genes after induction of DNA CT *in vivo*. S100A2 mRNA levels decreased after photooxidant and light treatment, reflecting the oxidation and dissociation of p53 from the S100A2

site. However, caspase-1 and ODC1 mRNA levels remained the same, indicating less DNA-mediated p53 oxidation.

The results from this study illustrate how protein oxidation at a distance through DNA CT contributes to cellular signaling. This oxidative signaling can control how p53 regulates gene expression under oxidative stress, and this signaling may be disrupted in cancerous cells.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Chapter 1

Chapter 2

Chapter 3

CHAPTER 1

DNA-Mediated Charge Transport in a Biological Context

DNA-Mediated Charge Transport

DNA, which contains the genetic information for all forms of life, is a double-stranded π-stacked heterocyclic-base polymer with unique properties.[1,2] Figure 1.1 shows the structure of B-form DNA and the composition of the bases that compose the genetic code. One of the remarkable abilities of DNA is the ability to conduct charge, termed DNA-mediated charge transport (DNA CT).[3] DNA CT was first hypothesized shortly after the publication of the structure of DNA by Watson and Crick. An electron (or an electron hole) takes advantage of the electronic coupling of the base pairs through their heterocyclic π-stacks to travel through the duplex. Studies of this phenomenon have employed organic and inorganic photooxidants, as well as the biophysical and biochemical properties of DNA.[3-5] DNA CT is extremely sensitive to the structure and base pair stacking of DNA: a mismatch, abasic site, or bend in the DNA will inhibit the CT.[6,7] DNA CT is a fast process (compared to protein CT), and it has been shown to occur in duplexes up to 34 nm (100 base pairs).[8] While the body of *in vitro* work is extensive, less has been done *in vivo*; nor has coupling of DNA CT with protein CT been thoroughly investigated.

DNA CT depends on the redox properties of the bases. Guanine has the lowest potential of the bases (0.9 V), and regions of guanine doublets or triplets have lower oxidation potentials than other combinations of bases.[9,10] When short DNA duplexes are tethered to rhodium-based photooxidants and irradiated, damage accumulates at the GG sites (Figure 1.2).[6,11] Oxidative DNA CT functions best through adenine tracts,[12-14] whereas reductive CT favors runs of thymidine (which has the highest redox potential).[15]

Sequence dependence of DNA CT increases with the distance along the DNA over which the transport occurs.[11,12]

The cell uses DNA CT to locate DNA damage as well as for signaling purposes. Isolated HeLa nuclei incubated with $Rh(phi)_2bpy^{3+}$ (phi = phenanthrenequinone diimine and bpy = 4,4'-dimethyl-2,2'-bipyridine) and irradiated for 30 min suffered damage at protein-bound sections of the genome inaccessible to the rhodium complex, indicating the possibility of long-range DNA CT in organelles.[16] Further work showed that DNA CT was not impeded by the spatial constraints of DNA packaging in nucleosomes.[17] In mitochondria, base oxidation by DNA CT accumulates in mutational hot spots characterized by high guanine content, such as conserved sequence block II.[18,19] These hot spots are associated with both germline and somatic mutations in tumors.[18] Furthermore, repeated rounds of induced DNA CT in mitochondria resulted in another mutational response: guanine to thymine transversion.[20]

DNA CT can also be coupled to protein CT. This coupling serves various functions, including location of damaged DNA and regulation of transcription factors. Base excision repair (BER) proteins, that contain an iron-sulfur cluster, such as MutY and EndoIII, are thought to use reductive DNA CT in the following fashion to "communicate" with each other:[21] going from a 2+ state to a 3+ state, which confers greater DNA affinity. If the helix is without lesions, the electron travels through the DNA until it finds another iron-sulfur containing protein, such as another BER enzyme; it then reduces that iron-sulfur cluster from 3+ to 2+, weakening its ability to bind DNA and causing it to dissociate from the DNA. If there is a lesion, the electron cannot reach the BER enzyme, which remains on the DNA. This process culminates in the redistribution

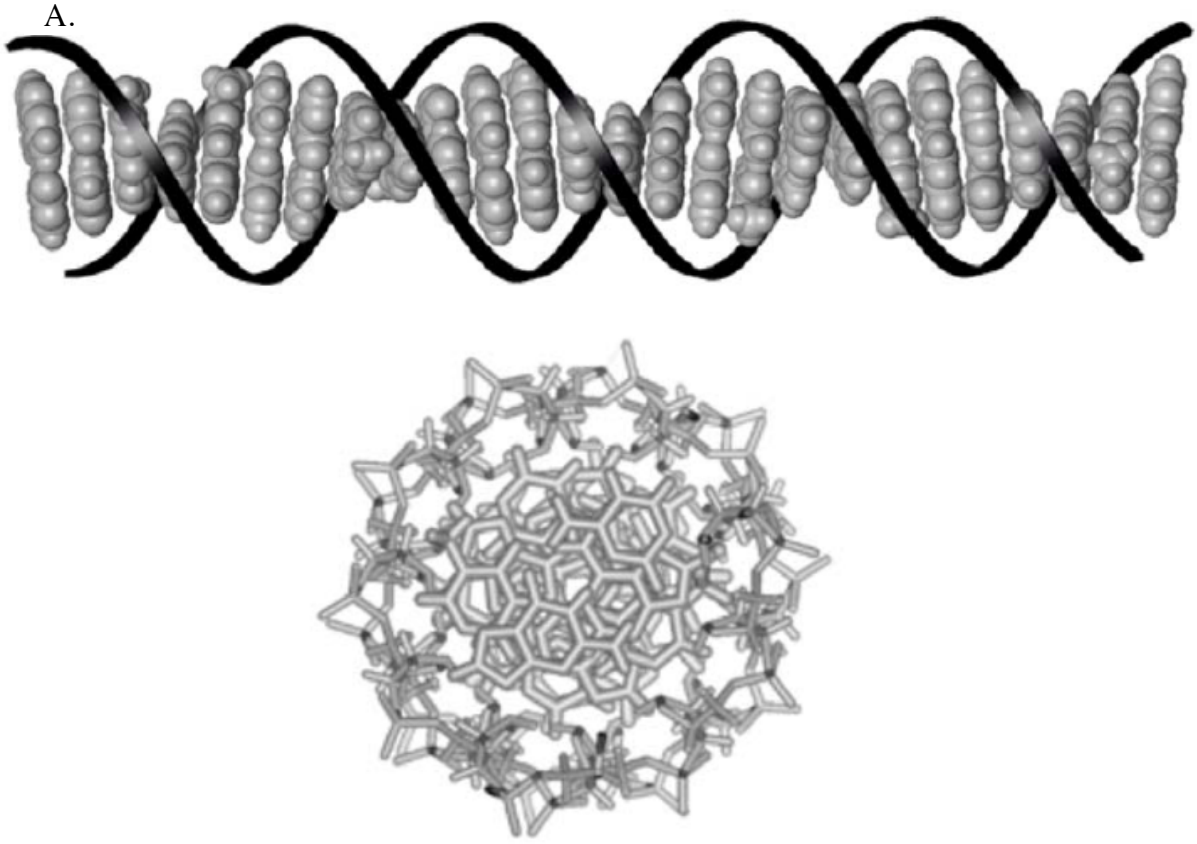
of repair proteins in the vicinity and ultimately the repair of the lesion.[21] Because the iron-sulfur clusters of both EndoIII and MutY are fairly distant from the DNA, the proteins rely on a network of mostly aromatic amino acids to transfer charge. If these amino acids are mutated, as in certain cancer-related human homologues, these mutants retain glycosylase activity but are insensitive to DNA CT.[22,23]

DNA CT has been also found to be a mechanism for the regulation of transcription factors, including the bacterial repressor SoxR. SoxR undergoes a change in redox potential upon DNA binding, from –290 to 200 mV vs. NHE (normal hydrogen electrode).[24] *In vitro,* reduced SoxR inhibits guanine damage induced by ruthenium-based photooxidants in the SoxR binding region of the SoxS promoter. *In vivo,* oxidative DNA CT (an electron hole) oxidizes the 2Fe-2S cluster of one SoxR monomer from 1+ to 2+, leading to a conformational change that upregulates production of SoxS, the superoxide response system.[24,25] This newly discovered regulatory function confers greater sensitivity for the organism in responding to stress; oxidative DNA CT would occur in cells under oxidative stress, in which the up-regulation of SoxS could prevent irreparable oxidative damage to the genome.

Figure 1.1. DNA structure. A. Views of the double helix, emphasizing the base stacking.

B. Chemical structure of the bases.

A.



B.

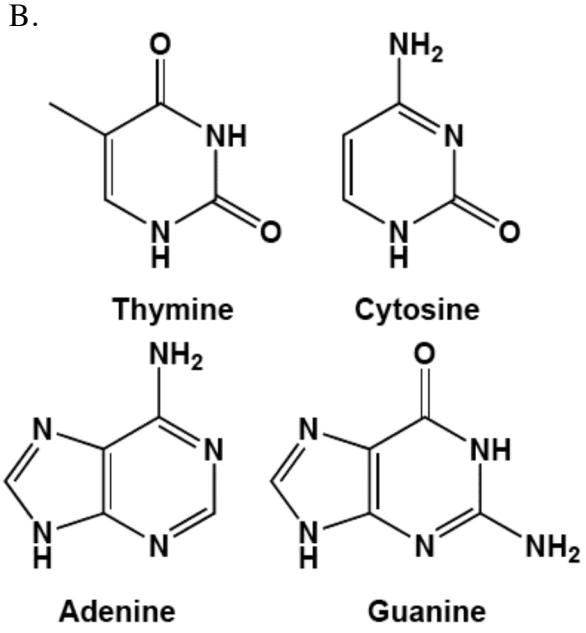
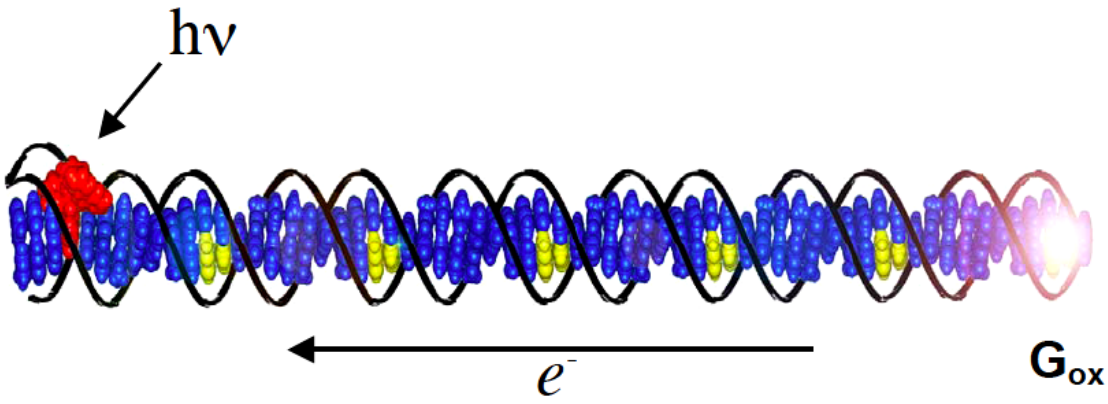


Figure 1.2. Representation of oxidative DNA CT. Adapted from Nunez and coworkers.[11]

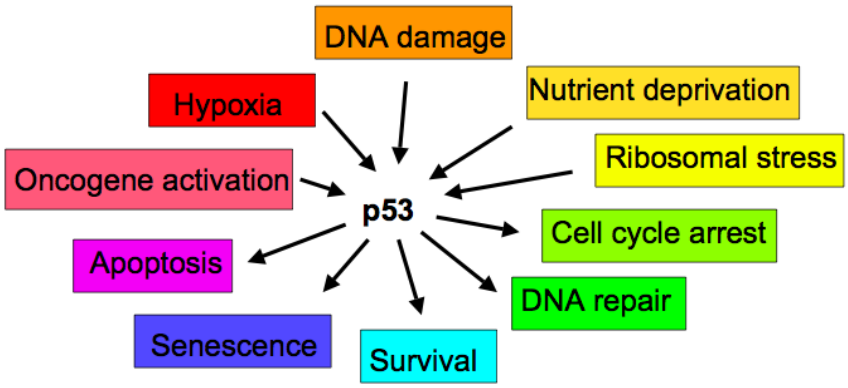


p53

A global transcription factor, p53 is a central hub for regulation of many cellular processes, including growth, reproduction, senescence, DNA repair, and the oxidative stress response (Figure 1.3). When stress levels are low, p53 concentration in the cell is low, and p53 up-regulates the expression of genes that maintain antioxidant function.[26] As oxidative stress increases to critical levels, p53 concentration in the cell increases, and p53 up-regulates pro-oxidant, pro-apoptotic genes to initiate cell death and prevent tumorigenesis.[27] p53 also partners with cathepsin and reactive oxygen species for induced necrosis when mitochondrial apoptotic factors are missing or mutated.[28]

The protein p53 has 393 amino acids, and it binds to its DNA sequence as a tetramer. The domains of human p53 are as follows: the N-terminal transactivation domain (residues 1–62), a proline-rich region (63–94), the central DNA-binding core domain (95–292), a linking domain (293–325), a tetramerization domain (326–355) and a negative regulatory domain (356–393). The DNA binding domain determines the stability of the protein; the N- and C-terminal domains tend to exist as disordered loops.[29,30] Inside the cell, p53 has a half-life of 20 minutes and requires other proteins, lipids, and cellular components to be stable.[31] There are ten cysteines in each monomer of human p53; nine are highly evolutionarily conserved (the exception is cysteine 229),[32] and three (176, 238, and 242) coordinate, together with histidine 179, to a structural zinc ion.[32] Substitution of the zinc ion with other divalent metals results in similar structure to WT p53 but inhibits DNA binding.[33] Cysteine 277 forms a hydrogen bond to a pyrimidine in the recognition sequence, and cysteines 275, 141, 135, and 124 are within 20 Å of the DNA duplex in all p53 crystal structures (as seen in Figure 1.4).[35]

Figure 1.3. Pathways regulated by p53. Adapted from Vousden and Lane.[34]



The half-site recognition sequence usually follows the pattern: 5'-PuPuPuC(A/T)(A/T)GPyPyPy-3', where Pu and Py represent purines and pyrimidines, respectively; however, p53 will bind to sites that are not a 100% match.[35] Conversely, sometimes it will not bind to sites that match the recognition sequence.[32,35,36] This promiscuity in DNA recognition is explained by several key residue contacts. Arginine residues interact with purines in the consensus sequence.[35] Lysine 120 also abuts a purine base in the crystal structure, and acetylation of this residue can lead to stronger DNA binding.[35,37] When the first base is of the sequence is adenine, lysine 120 interacts with a guanine in position 3 and with the opposite strand's thymine in position 9, while cysteine 277 and alanine 276 have favorable van der Waals interactions with the same thymine.[38]

Slightly fewer than half of all human tumors contain p53 mutations, and 80% of these occur in the DNA-binding domain (all cancers have mutations in proteins somewhere in p53-controlled signaling pathways).[39] About 75% of all p53 mutations are missense mutations, which change the sequence of the protein by one amino acid.[40] Some highly prevalent mutations, including substitutions from arginine to other amino acids, prevent DNA binding; however, many mutants retain the ability to bind to DNA, despite other altered properties.[41] In particular, cysteine residues are known to coordinate the response of p53 to oxidative stress, but substitutions of serine for cysteine retain binding in all cases except for the zinc-coordinating cysteines.[42] The redox state of the protein affects its binding capabilities to some DNA sequences but not others; for instance, the binding of p53 to the p21-linked recognition sequence does not depend on the redox state, but the Gadd45-linked sequence shows greater binding to p53 when it is completely reduced.[43] Similarly, the redox state of p53 does not affect its ability to bind to an AT-

rich recognition sequence, but reduction of p53 is required for total binding to a analogous, GC-rich recognition sequence.[43]

The regulation of p53 through DNA-mediated oxidation has been briefly investigated.[44] Oxidation of p53 at a distance (by a covalently attached anthraquinone photooxidant) leads to a conformational change in the protein, which inhibits its association with DNA. However, introduction of a CA mismatch into the bases between the p53 recognition site and the photooxidant prevents DNA CT and thus p53 oxidation. This oxidation occurs for some DNA sequences, but not for others, depending on the cellular role for the gene; for instance, p53 does not show oxidation at a distance when bound to the p21 recognition site, but it does when bound to the Gadd45 recognition site, although both sequences have identical G/C content.[44] This result hinted that dissociation of p53 from the Gadd45 promoter might down-regulate that gene, which would be an appropriate response to oxidative stress. Oxidative DNA CT should correlate with extreme oxidative stress; under those conditions, DNA repair (coordinated by Gadd45)[45] would be insufficient to repair all the cellular genomic damage, and producing cell cycle arrest genes (such as p21)[46] instead would be a cellular priority. Determining any sequence-dependent effects and further identification of other DNA CT regulated, p53-linked genes were intriguing areas of research.

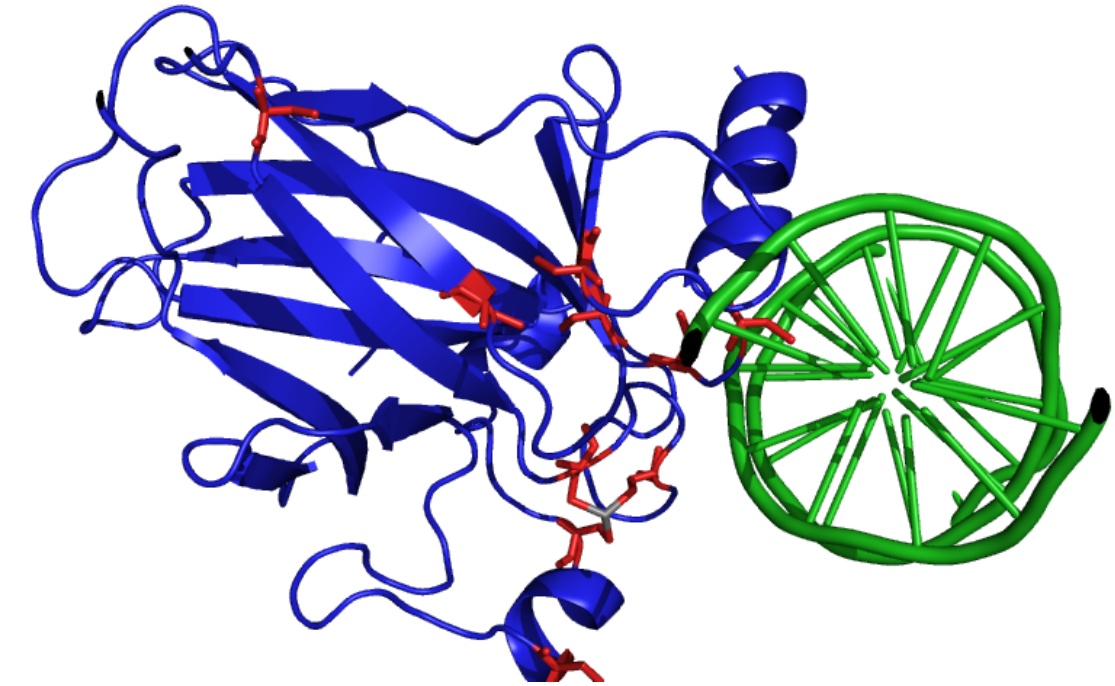
While preliminary mass spectrometry results hinted that cysteine 141 was involved in the oxidation of p53, it was not known if it formed a disulfide bond with any of the nearby cysteines, 124 or 135, or if a noncysteine oxidation product was formed.[44] Other cysteine oxidation products, such as sulfenic acid or sulfonic acid, were also possible. Also, since cysteine 141 is not in direct contact with the DNA, it was

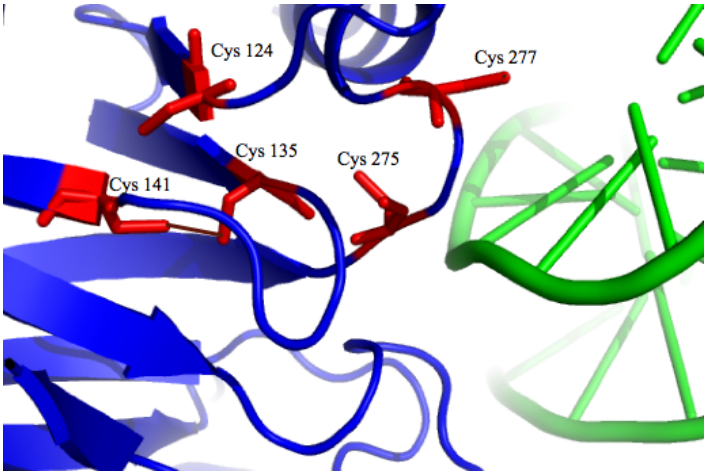
hypothesized that protein charge transport or disulfide bond shuffle (through the network of cysteines) led to cysteine 141 as the final oxidation target.

Augustyn et al. confirmed the *in vivo* oxidation of p53 through DNA CT through photooxidation of stressed HeLa cells.[44] The cells were first exposed to cisplatin in order to upregulate p53. Then the photooxidant $[Rh(phi)_2bpy]^{3+}$, which intercalates into cellular DNA, was added to the cells. Two forms of p53 were observed; these matched the forms detected when cells were treated with hydrogen peroxide. It was concluded that oxidative stress, either from photooxidant-induced DNA-mediated charge transport, or from peroxide, could lead to the oxidation of p53.

Here, some of the previously unsolved questions about the regulation of p53 by DNA CT are addressed. Earlier work suggested that cysteines on the interior of the protein become oxidized; how could oxidation permeate from the DNA to cysteine 141? Through disulfide bond exchange, the disulfide bond could be shuffled to (possibly lower-potential) cysteines farther from the DNA. Clearly identifying which cysteines are oxidized, and which cysteines control this oxidation process, were two key goals of this work. Furthermore, since p53 binds to a wide variety of DNA sequences, this investigation examined sequence-dependent effects in DNA CT-mediated p53 regulation.

Figure 1.4. A) p53 as a monomer. Cysteines are in red. B) Monomer p53 with possibly redox-active cysteines named and in red (PDB 1TSR).





Summary

DNA CT is a well-studied phenomenon, but its applicability to biology is only beginning to be investigated. It is capable of long-range signaling, and its sensitivity to base pair stacking makes it an effective tool for detecting mutations. DNA CT regulates transcription factors, including the transcription factors SoxR in prokaryotes and p53 in vertebrates.[47]

The transcription factor p53, a cellular signaling hub, is key in the response of cells to stresses of all kinds, including oxidative stress.[48] When cellular signals indicating stress are activated, p53 concentration in the cell increases, and when the cell is severely stressed, p53 initiates apoptosis.[40] Furthermore, cancer has been linked to many p53 mutations. The majority of studies on p53 mutants have focused on mutations that interrupt DNA-p53 contacts, or that give p53 novel signaling powers; however, some mutations in the DNA-binding domain do not interrupt binding. These mutations have not been studied extensively. Interestingly, substitution of serine for cysteine at either residue 141 or 277 eliminates sensitivity to oxidative stress in certain cancers.[43,49] Also particularly intriguing is the sequence dependence of protein-DNA binding sensitivity,[43] especially in light of the known sequence sensitivity that DNA CT possesses.[12]

DNA CT regulation of p53 is an exciting new chapter in the study of the "guardian of the genome." Determining the type and extent of the oxidation of p53 at a distance is a key step in understanding the role of DNA CT in biological systems. The influence of the different recognition sequences for p53 on its oxidation may reveal which genes and pathways use DNA CT as a regulation mechanism when the cell is under severe oxidative stress. Determining how mutation of redox-active residues

disrupts this signaling should be a significant step towards understanding and treating cancers.

REFERENCES

(1) Avery, O. T., MacLeod, C. M., and McCarty, M. *J. Exp. Med.* **1944**, *79*, 137-158.

(2) Hershey, A. D., and Chase, M. *J. Gen. Physiol.* **1952**, *36*, 39-56.

(3) O'Neill, M. A., and Barton, J. K. *Topics Curr. Chem.* **2004**, *236*, 67-115.

(4) Genereux, J. C., and Barton, J. K. *Chem. Rev.* **2010**, *110*, 1642-1662.

(5) Barton, J. K., Olmon, E. D., and Sontz, P. A. *Coord. Chem. Rev.* **2011**, *255*, 619-634.

(6) Hall, D. B., Holmilin, R. E., and Barton, J. K. *Nature* **1996**, *382*, 731-735.

(7) O'Neill, M. A., and Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 13234-13235.

(8) Slinker, J. D., Muren, N. B., Renfrew, S. E., Barton, J. K. *Nat. Chem.* **2011**, *3*, 230-235.

(9) Sugiyama, H., and Saito, I. *J. Am. Chem. Soc.* **1996**, *118*, 7063-7068.

(10) Saito, I., Takayama, M., Sugiyama, H., and Nakatani, K. *J. Am. Chem. Soc.* **1995**, *117*, 6406-6407.

(11) Nunez, M. E., Hall, D. B., and Barton, J. K. *Chem. & Biol.* **1999**, *6*, 85-97.

(12) Williams, T. T., Odom, D. T., and Barton, J. K. *J. Am. Chem. Soc.* **2000**, *117*, 6406-6407.

(13) Liu, C. S., and Schuster, G. B. *J. Am. Chem. Soc.* **2003**, *125*, 6098-6102.

(14) O'Neill, M. A., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16543-16550.

(15) Shao, F., and Barton, J. K. *J. Am. Chem. Soc.* **2007**, *129*, 14733-14738.

(16) Nunez, M. E., Holmquist, G. P., and Barton, J. K. *Biochem.* **2001**, *40*, 12465-12471.

(17) Rajski, S. R., and Barton, J. K. *Biochem.* **2001**, *40*, 5556-5564.

(18) Merino, E. J., and Barton, J. K. *Biochem.* **2007**, *46*, 2805-2811.

(19) Merino, E. J., and Barton, J. K. *Biochem.* **2008**, *47*, 1511-1517.

(20) Merino, E. J., Davis, M. L., and Barton, J. K. *Biochem*. **2009**, *48*, 660-666.

(21) Boal, A. K., Genereux, J. C., Sontz, P. A., Gralnick, J. A., Newman, D. K., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15237-15242.

(22) Boal, A. K., Yavin, E., and Barton, J. K. *J. Inorg. Biochem*. **2007**, *101*, 1913-1921.

(23) Romano, C. R., Sontz, P. A., Barton, J. K. *Biochem*. **2011**, epub June 9.

(24) Gorodetsky, A. A., Dietrich, L. E. P., Lee, P. E., Demple, B., Newman, D. K., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3684-3689.

(25) Lee, P. E., Demple, B., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13164-13168.

(26) Sablina, A. A., Budanov, A. V., Ilyinskaya, G. V., Agapova, L. S., Kravchenko, J. E., Chumakov, P. M. *Nat. Med*. **2005**, *11*, 1306-1313.

(27) Polyak, K., Xia, Y., Zweier, J. L., Kinzler, K. W., Vogelstein, B. *Nature* **1997**, *389*, 300-305.

(28) Tu, H. C., Ren, D., Wang, G. X., Chen, D. Y., Westergard, T. D., Kim, H., Sasagawa, S., Hsieh, J. J., Cheng, E. H. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 1093-1098.

(29) Wells, M., Tidow, H., Rutherford, T. J., Markwick, P., Jensen, M. R., Mylonas, E., Svergun, D. I., Blackledge, M., and Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5762-5767.

(30) Tidow, H., Melero, R., Mylonas, E., Freund, S. M., Grossmann, J. G., Carazo, J. M., Svergun, D. I., Valle, M., and Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12324-12329.

(31) Cañadillas, J. M. P., Tidow, H., Freund, S. M. V., Rutherford, T. J., Ang, H. C., and Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2109-2114.

(32) el-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W., and Vogelstein, B. *Nat. Gen*. **1992**, *1*, 45-49.

(33) Palecek, E., Brázdová, M., Cernocká, H., Vlk, D., Brázda, V., and Vojtesek, B. *Oncogene* **1999**, *18*, 3617-3625.

(34) Vousden, K. H., and Lane, D. P. *Nat. Rev. Mol. Cell Biol*. **2007**, *8*, 275-283.

(35) Cho, Y., Gorina, S., Jeffrey, P. D., and Pavletich, N. P. *Science* **1994**, *265*, 346-355.

(36) Menendez, D., Inga, A., and Resnick, M. A. *Nat. Rev. Cancer* **2009**, *9*, 724-737.

(37) Arbely, E., Natan, E., Brandt, T., Allen, M. D., Veprintsev, D. B., Robinson, C. V., Chin, J. W., Joerger, A. C., Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8251-8256.

(38) Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L, Haran, T. E., and Shakked, Z. *Mol. Cell.* **2006**, *22*, 741-753.

(39) Joerger, A. C., and Fersht, A. R. *Oncogene* **2007**, *26*, 2226-2242.

(40) Brosh, R., and Rotter, V. *Nat. Rev. Cancer* **2009**, *9*, 701-713.

(41) Kato, S., Han, S. Y., Liu, W., Otsuka, K., Shibata, H., Kanamaru, R., Ishioka, C. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 8424-8429.

(42) Rainwater, R., Parks, D., Anderson, M. E., Tegtmeyer, P., and Mann, K. *Mol. Cell. Biol.* **1995**, *15*, 3892-3903.

(43) Buzek, J., Latonen, L., Kurki, S., Peltonen, K., and Laiho, M. *Nucl. Acids Res.* **2002**, *30*, 2340-2348.

(44) Augustyn, K. E., Merino, E., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 18907-18912.

(45) Jung, H. J., Kim, E. H., Mun, J.-Y., Park, S., Smith, M. L., Han, S. S., and Seo, Y. R. *Oncogene* **2007**, *26*, 7517-7525.

(46) Abbas, T., and Dutta, A. *Nat. Rev. Cancer* **2009**, *9*, 400-414.

(47) Merino, E. J., Boal, A. K., and Barton, J. K. *Curr. Opin. Chem. Biol.* **2008**, *2008*, 229-237.

(48) Vogelstein, B., Lane, D., and Levine, A.J. *Nature* **2000**, *408*, 307-310.

(49) Velu, C. S., Niture, S. K., Doneanu, C. E., Pattabiraman, N., and Srivenugopal, K. S. *Biochem.* **2007**, *46*, 7765-7780.

CHAPTER 2

Oxidation of p53 through DNA-Mediated Charge Transport: Dependence on p53 Sequence

INTRODUCTION

p53, a redox-active transcription factor, regulates many major cell functions, including growth cycle arrest, DNA repair, response to oxidative stress, and apoptosis.[1] Its correct function and tight regulation are vital to prevent tumorigenesis; in nearly half of all cancers, p53 is mutated, and 80% of those mutations are found in the DNA binding region.[2] Of those mutations, the majority are missense mutations. While many missense mutations disrupt DNA-p53 binding, around a third of cancer-linked mutations in the DNA binding domain do not disrupt binding and their function in cancer is unknown.[3] The protein p53 binds, as a tetramer, to a DNA consensus sequence consisting of two 10-base pair half-sites: 5'-PuPuPuC(A/T)(A/T)GPyPyPy-3', where Pu and Py represent purines and pyrimidines, respectively; a linker of 0-13 base pairs may also be present.[4] However, many sequences that are not a 100% match will bind to p53, and not all sequences that match the consensus sequence will bind to p53.[4] The consensus sequences of p53 may be found anywhere near the gene, but many of the recognition sequences are found in the promoter or noncoding regions.[5] The protein p53 has greater influence over the control of transcription if the recognition sequence is closer to the start codon.[5] In response to cellular stress, posttranslational modifications of p53 are important for its up-regulation or down-regulation of genes; the modifications include phosphorylation, glycosylation, glutathionylation, acetylation, and others.[1,6] Importantly, some sequences will only bind to p53 in its reduced state.[7]

Located in the DNA-binding domain, the cysteine residues control redox sensitivity of p53.[7] Eight of the ten cysteine residues are highly evolutionarily conserved, including the three involved in coordination to the zinc ion.[8] It is necessary for three of

the reduced cysteine residues (176, 238, and 242, together with histidine 179) to coordinate a structural zinc ion for DNA binding to occur.[9] Different cysteines have been implicated in the response to various oxidative type stresses; for instance, cysteine 141 becomes glutathionylated when HCT116 cells are exposed to the DNA-damaging small molecule camptothecin.[6] In contrast, selenomethionine activates p53 through a redox mechanism involving cysteines 275 and 277,[10] and diamide treatment leads to oxidation of cysteine 182.[11,12] Although DNA binding is retained when cysteines in p53 (that are not coordinated to the zinc ion) are mutated to serine,[13] WS-1 cells lose their redox-differential binding for Gadd45 sequences when cysteine 277 is mutated to the redox silent serine.[7] While the downstream effects of the response of p53 to oxidative stress and DNA damage are well studied, the direct oxidation of p53 and other transcription factors by DNA-mediated charge transport (DNA CT) as a cellular regulatory process has only briefly been studied.

DNA, the near-ubiquitous source of genetic information, has the ability to transport charge along its heterocyclic base polymer π-stack.[14] An oxidant creates the charge (electron hole) by acquiring an electron from the DNA; then the hole migrates to the sequence of bases with low oxidative potential, and there the DNA reacts with water or other species to form an oxidative lesion or a DNA-protein cross-link.[15,16] The clustering of the damage near low oxidation sites such as guanine doublets and triplets,[17,18] has been extensively studied both *in vitro*,[14,16,19] and in biological systems.[20-22] Studies have demonstrated that DNA CT can occur though a 100-mer (34 nm) in its native conformation,[23] but DNA CT is sensitive to base stack perturbations such as mismatches, abasic sites, sharp angles in the DNA, and other structural alterations.[16,24-26]

One-electron DNA CT has been shown to lead to two-electron oxidation products; irradiated thiol- and photooxidant-linked DNA showed formation of a disulfide bond because of oxidative DNA CT.[27] DNA CT has been shown to have a biological role in the response to oxidative stress; in prokaryotes, this response includes DNA repair[28] and the activation of the superoxide response system through the transcription factor SoxR.[29] In eukaryotes, $[Rh(phi)_2bpy]Cl_3$ (phi = phenanthrenequinone diimine and bpy = 4,4'-dimethyl-2,2'-bipyridine) enabled DNA CT to shuttle DNA damage to specific regulatory regions both in irradiated mitochondria[20,22,30] and nuclear extracts.[21,31] The response of proteins to DNA CT often depends on the CT pathway of the protein itself, with cancer-related mutants sometimes interfering with CT deficient pathways.[28,32,33]

Previously, it was shown that DNA CT could be a means of the cellular regulation of p53.[34] Photoinduced dissociation of wild type p53 from photooxidant-conjugated DNA containing various p53 recognition sites was shown through electrophorectic mobility shift (EMSA) assays. When a mismatch was introduced to the synthetically derived consensus sequence, p53 remained bound to the DNA. Photoinduced dissociation was shown to depend on the sequence of the DNA; p53 dissociated from a synthetically derived consensus sequence and the recognition sequences corresponding to regulation of Gadd45 (a DNA repair protein) and MDM2, its negative feedback regulator.[34] p53 did not undergo photoinduced dissociation when bound to recognition sequences of p21, a cell cycle arrest protein.[35] This is beneficial for the cell because under periods of high oxidative stress, the cell will not squander resources to activate genes like Gadd45 that would repair the DNA,[36] or inactivate p53 by increasing levels of MDM2,[37] but instead focus on more extreme methods of controlling damage or starting apoptosis. The

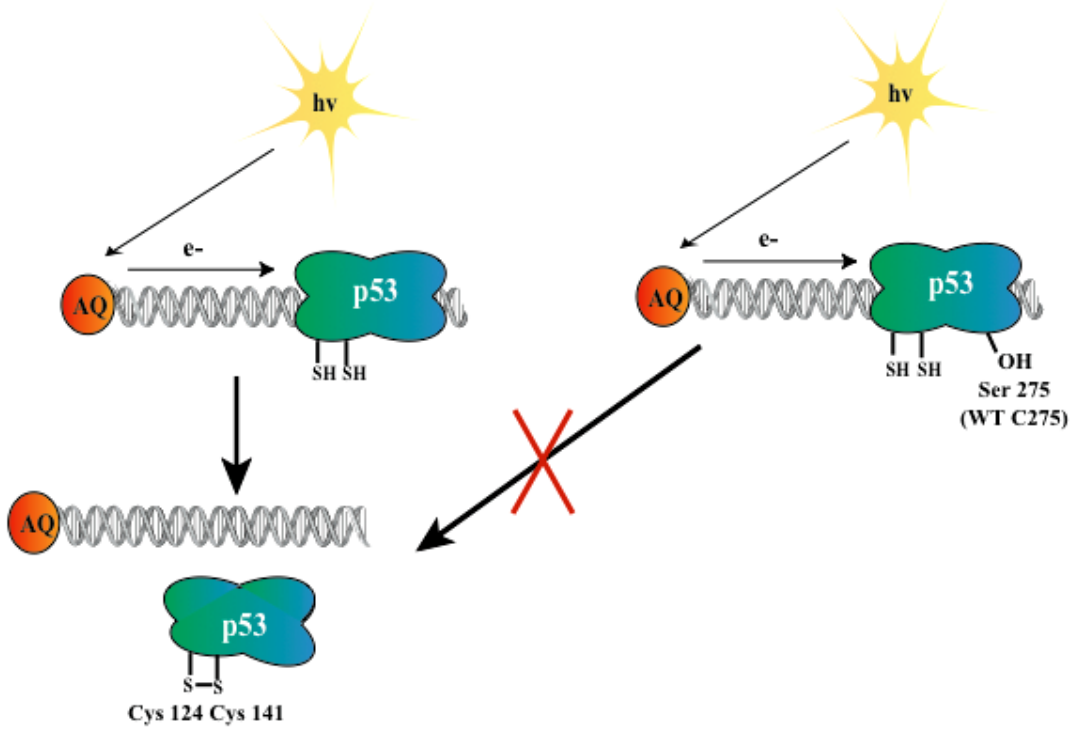
oxidation of cellular p53 was also confirmed through Western blotting of p53 from HCT116 cells after incubation with the photooxidant $[Rh(phi)_2bpy]Cl_3$ and irradiation. Through mass spectrometry experiments of the tryptic peptides from the photooxidation experiments, cysteine 141 was implicated in the photoinduced dissociation of p53. However, it was unknown if other cysteines were involved in the oxidation process, or if other oxidation products (such as sulfenic acid or tyrosine-based adducts) or covalent DNA-protein cross-links were formed.

Furthermore, the charge transport pathway of the p53 mutant N239Y was of interest to us for several reasons. As stated above, the majority of cancer-related p53 mutants are missense mutations,[2] and tyrosine is a redox-active amino acid with a redox potential close to cysteine,[38] the mutant N239Y has been linked to several cancers of tissues known to suffer increased oxidative stress, such as rectal cancer.[39,40] Additionally, the naturally occurring Y236F mutation was of interest because of the reasons previously stated for tyrosine residues and its link to lung cancer.[41] Mutation of various cysteine residues to serine have also been linked to a number of cancers; notably C124S, bronchus;[42] C135S, colon, bladder, skin, among others;[43-45] C141S, bladder, lung, rectal cancer, among others;[46,47] C275S, stomach, lung, bladder;[47-49] and C277S, nasal cavity and middle ear.[50]

The overall goal of this study is to further define the molecular response of p53 to oxidative DNA CT. Through EMSA assays and using site-directed mutagenesis to create redox-defective mutants, this study shows that the protein oxidation response to DNA CT is controlled by cysteine 275. For p53 undergoing DNA-mediated oxidation at a distance, cysteines 124 and 141 are identified through liquid chromatography-tandem

mass spectrometry (LC MS/MS) as the cysteines that form a disulfide bond. This disulfide bond formation inhibits the ability of p53 to bind to DNA, and thus enables the cell to regulate the transcriptional activation of certain genes by p53 through the direct oxidation of p53. Additionally, DNA CT initiates DNA-protein cross-linking, which may serve a biologically interesting role. Overall, this study describes in detail how the cell uses DNA CT to regulate the transcription factor p53.

Figure 2.1. Model of DNA CT oxidation leading to p53 disulfide bond formation and/or protein-DNA cross-linking.



EXPERIMENTAL

Materials. All buffers were freshly made and filtered before use. The NP-40 detergent was purchased as a 10% solution from Pierce (Rockford, IL). Reagents for DNA synthesis were purchased from Glen Research (Sterling, VA). The plasmid containing the quadruple mutant-p53 gene (M133L/V203A/N239Y/N268D) in a pET24a-HLTV vector was a kind gift from Alan Fersht. The TEV protease was purchased from Invitrogen. Unless specified, other enzymes were purchased from Roche (Indianapolis, IN). Other reagents were purchased from Sigma Aldrich (St. Louis, MO) at highest available purity and used without further purification. N-methyliodoacetamide was synthesized according to Ramseier and Chang.[51] The mass plus one proton was found to be 200.0 m/z (not shown).

Synthesis and modification of oligonucleotides. The DNA was synthesized using standard solid phase automated synthesis, modified with anthraquinone prior to cleavage from the beads, and radiolabeled as in Augustyn et al.[34] The DNA containing the Gadd45 recognition site and the competitor DNA strand is shown in Figure 2.2.

Figure 2.2. DNA sequences used in this study.

a) Duplex containing *gadd45* recognition site (in blue). Anthraquinone-linked duplex (AQ): 5′ AQ-AAATCAGCACTACAGCATGCTTAGACATGTTC 3′
3′ TTTAGTCGTGATGTCGTACGAATCTGTACAAG 5′

Duplex without anthraquinone: light control (LC)

5′ AAATCAGCACTACAGCATGCTTAGACATGTTC 3′
3′ TTTAGTCGTGATGTCGTACGAATCTGTACAAG 5′

b) Competitor DNA for gel shift assay

5′ GGAAAAAAAAAAAAAAAAAACC 3′
3′ CCTTTTTTTTTTTTTTTTTTGG 5′

*Protein preparation***.** The triple (M133L/V203A/N268D) mutant (p53SS) was cloned from the quadruple mutant p53 plasmid (N239Y/M133L/V203A/N268D), and subsequent mutants (C124S, C135S, C141S, Y236F, C275S, and C277S) were cloned from the triple mutant with primers shown in Appendix 1. The method used was PCR mutagenesis by overlap extension and gene splicing by overlap extension.[52] *Pfu* (Agilent Technologies) was the preferred polymerase. Briefly, two segments of a gene are PCR amplified independently with primers that contain the mutation of interest, using the protocol from the manufacturer. Then the PCR products are gel-purified (this gel purification step is crucial; other methods of purification are insufficient, and PCR will not work in the next step without purification). The PCR products are combined with 5' primers for the entire insert, and PCR is performed again. The final PCR product and the vector of interest is restriction enzyme digested (*BamHI* and *EcoRI,* using the protocol from the manufacturer) and 5' dephosphorylated with calf intestinal phosphatase. The prepared genes were ligated using ligase from New England Biolabs into the pET24a-HLTV vector with the *BamHI* and *EcoRI* cut sites, once again using the protocol from the manufacturer with the addition of 1-3 μM freshly prepared ATP. Laragen verified the correct sequences. All mutant p53 proteins were overexpressed and purified as in Veprintsev et al.[53] according to the protocol that is in Appendix 1, and then the dithiothreitol was diluted to nanomolar levels with p53 buffer (20 mM Tris, 100 mM KCl, 0.2 mM EDTA, 20% glycerol at pH 8) before the protein was flash-frozen (150-270 μM) and stored at −80 °C under argon. TEV cleavage of the purification tag was successful for only 50-75% of the protein, but the remaining protein with the tag did not

inhibit DNA binding or tetramer formation. The purity of the p53 was 98% or higher for all mutants by SDS-PAGE gel.

Electrophoretic mobility shift assay (EMSA). The EMSA analyses were performed as in Augustyn et al. with some modifications.[34] K_d's of each of the mutants were measured using EMSA with 0.25 nM DNA and increasing concentrations of p53. Then concentrations of p53 for oxidation assays analyzed with EMSA were chosen such that the irradiated sample contained a concentration of p53 that was near or above an order of magnitude greater than the K_d. Irradiation assays were performed at 25 nM p53 tetramer and DNA for most mutants, except for Y236F (50 nM) and C275S (3 μM). Instead of a lamp, an Oriel (Darmstadt, Germany) 1000-W Hg/Xe solar simulator (340-440 nm) was used for irradiation. TBE -10% acrylamide gels (12 well, 20 μL) were purchased from Biorad; samples were run on the gels at 50V for 1-1.5 hr in the cold room. Also, the gel was blotted onto a Hybond-N membrane that was imaged as described previously (this allows for better resolution and longer exposure times; over time, gel bands expand). Results were reported as a percentage of DNA bound normalized to the sample at the zero irradiation time point as determined with ImageQuant software. Data are an average of at least three assays, and the error is reported as the standard deviation.

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Samples were prepared as mass spectrometry samples were prepared previously.[34] After irradiation, Laemmli buffer without reductant (unless noted) was added, and the samples were run on a 7.5% acrylamide Tris-glycine gel in Tris-glycine-SDS running buffer (Biorad). The gels were washed 3 times with ultrapure water and shaking, and stained

with Biosafe Coomassie Blue stain for 1 hr. The gels were destained with several changes of ultrapure water before imaging on a Li-Cor 700 nm filter system. The fraction of the protein that was cross-linked was quantified using Odyssey software. Samples were repeated three times, and error reported is standard deviation.

Mass spectrometry (MS) analysis. The MS samples were prepared as described previously[34] and irradiated for 1 hr with a 1000-W Hg/Xe solar simulator (340-440 nm). The samples were denatured with 6 M guanidine hydrochloride and treated with 50 mM iodoacetamide and 10 mM dimedone for 1 hr in the dark to cap free thiols. Then the iodoacetamide reaction was quenched with DTT, which also reduced the oxidized cysteines. The samples were desalted by purification using a macrotrap protein column (C8) (Matrix Biosciences) on an Agilent HP 1100. The samples were reduced with 1 mM DTT in 8 M urea with sonication for 20 min. Then the samples were treated with 40 mM N-methyliodoacetamide in 8 M urea for one hour in the dark. The reaction was quenched with 75 mM DTT, and the samples were digested with LysC and subsequently trypsin overnight at 37° C. Then the samples were desalted by HPLC purification with a macrotrap peptide column as described above, and they were put on an LTQ-FT-Ultra or an Orbitrap mass spectrometer (Thermo Scientific). The data were analyzed with MASCOT software (Matrix Science). Transient modifications that were included in the analysis are carbamidomethylation (56 amu), N-methyl-amidomethylene (71 amu), dimedone-modification (140 amu) and sulfinic acid for cysteine; and oxidation for both methionine and tyrosine.

RESULTS

Oxidation of p53 at a distance leads to dissociation from DNA. Previous results show that regulation of p53 through DNA-mediated oxidation leads to a conformational change that causes p53 to dissociate from its recognition site on the DNA, likely occurring through formation of a disulfide bond between cysteine 141 and another nearby cysteine. To further determine the chemistry leading to the dissociation, we used an approach combining EMSA analysis to quantify dissociation, site-directed mutagenesis to determine key residues, SDS-PAGE to visualize other interactions including cross-linking, and MS to identify amino acids involved in the oxidation event.

We employed a superstable triple mutant of p53 (M133L/V203A/N268D) (p53SS) amenable to site-directed mutagenesis to investigate the charge transport landscape of the p53-DNA interface. Additionally, we used site-directed mutagenesis to create mutant proteins corresponding to different cancer-related forms of p53. To confirm that the mutants had similar DNA-binding properties to WT p53, we measured K_d's of the mutants for the Gadd45 recognition element, with and without the covalently tethered photooxidant, and we discovered that they were similar to published K_d's for wild-type p53 interacting with the Gadd45 recognition element (Table 2.1).[54] Only C275S had a significantly different K_d of 56 nM, which, while an order of magnitude higher than the other values, is still consistent with the $K_{d's}$ of p53 bound to other biologically important recognition sequences.[54]

DNA-mediated oxidation when bound to the Gadd45 sequence for each of the mutants is shown in Figure 2.3. A representative gel with 25 nM DNA and 25 nM C141S p53 tetramer in the samples is shown in 2.3A, and the aggregate data are shown in

2.3B. We chose concentrations of DNA and p53 for near-complete binding of p53 throughout the scope of the experiment. Because oxidized p53 does not bind to this sequence of DNA,[7] the ratio of the radiolabeled bound DNA to the total signal of DNA gives the total amount of p53 binding in a lane. Oxidation data were normalized to the amount of DNA binding to p53 in the zero irradiation time points (at least 80% of total DNA present). As expected, each mutant shows dissociation from the AQ-DNA when subjected to DNA-mediated oxidation, at a total oxidation amount similar to WT at the same irradiation time point of 30 min except for C275S. The dissociation for the sample without covalent photooxidant is minimal, at most 10% after an hour of irradiation. The extent of oxidation also resembles the result seen for the MDM2 recognition element in the previous study. There is not a statistical difference between the amount of dissociation, and thus oxidation of p53, between the pseudo-WT and other mutants except for C275S. C275S shows minimal oxidation (similar to the sample without anthraquinone) until the 60 min time point; then a slight amount of oxidation is observed. This lack of oxidation is especially significant given that C275S has the weakest binding affinity out of the proteins in this study, but it shows the lowest ability to dissociate upon irradiation. Also, when we calculate apparent K_d's, the K_D for C275S post-irradiation does not change from its normal K_d; however, the K_d's for the other proteins show significant decreases in binding ability after irradiation of the samples. Because C275S is a mutation found in cancers affecting tissues that experience oxidative stress (lung, stomach, and bladder), it is possible that this mutation allows cancers to escaped DNA CT signaling to p53, affording them the opportunity to avoid apoptosis.

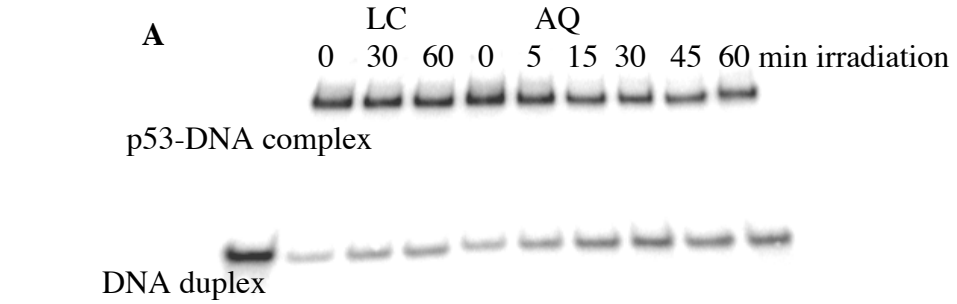
Table 2.1. K_d's for mutants bound to the Gadd45 recognition element. Error is standard deviation and reflects 3 replicates. LC is the duplex without covalent anthraquinone (AQ); AQ is the duplex with the covalent AQ.

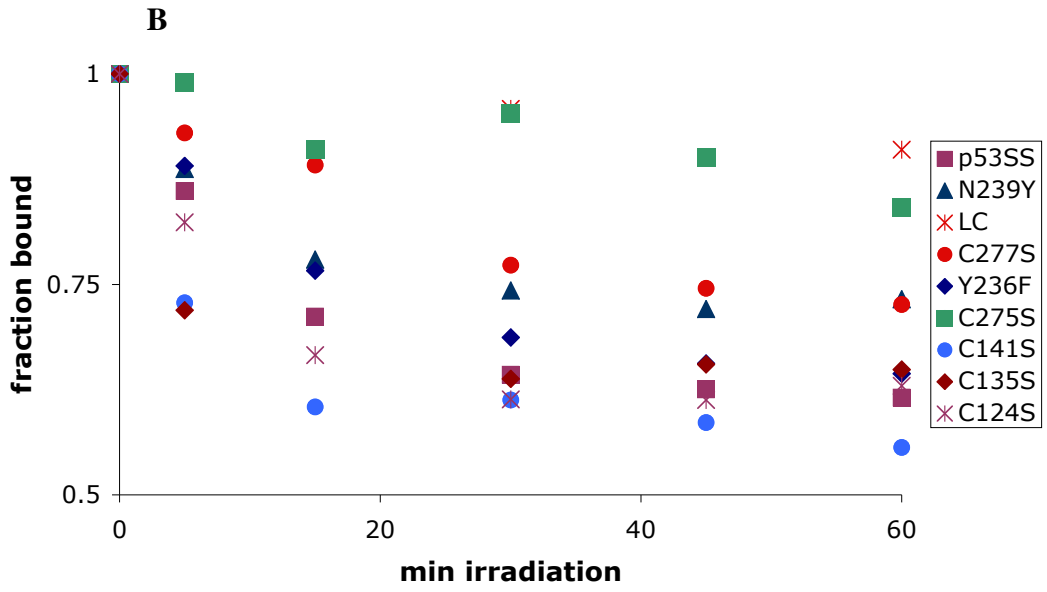
Mutant of p53	**LC K_d (nM)**	**AQ K_d (nM)**
M133L/V203A/N268D	1.6 ± 0.6	2.4 ± 1.1
N239Y/M133L/V203A/N268D	1.2 ± 0.4	1.0 ± 0.1
C124S/M133L/V203A/N268D	3.7 ± 0.5	4.29 ± 0.04
C135S/M133L/V203A/N268D	4.4 ± 2.8	3.1 ± 1.2
C141S/M133L/V203A/N268D	4.6 ± 1.2	3.7 ± 0.3
Y236F/M133L/V203A/N268D	9.7 ± 4.3	8.2 ± 4.7
C275S/M133L/V203A/N268D	56 ± 13	54 ± 8
C277S/ M133L/V203A/N268D	3.0 ± 0.8	2.3 ± 0.5

a) DNA sequence is shown in Figure 2.1.

b) DNA concentration was held constant at 25 nM; binding conditions were 30 min at ambient temperature in 20 mM Tris, 0.2 mM EDTA, 100 mM KCl, 20% glycerol at pH 8.

Figure 2.3. a) Example of data from an EMSA with C141S, 25 nM. b) Data compiled from gel shift experiments with the mutants. DNA:tetramer p53 ratio was held constant at 1:1, and the concentrations used were calculated, based on the K_d's, to be at 75% or more bound p53/DNA. Samples equilibrated at ambient temperature in 1% NP-40, 1% BSA, 5 μM competitor DNA, 20 mM Tris, 0.2 mM EDTA, 100 mM KCl, 20% glycerol at pH 8, before irradiation and EMSA analysis. Error reported is standard deviation, three replicates, and is 5% or less for most time points and most mutants (AQ C135S 5 min and 15 min error is 8.5 and 10%, respectively; AQ C141S at 60 min is 8.6%; and AQ C275S at 60 min is 6%). The error percentage for AQ-3x was 10-12% error.





This result was somewhat unexpected because in C277 is the closest residue to the DNA (3.6 Å) and thus was expected to be a key player in the oxidation response. Also, the mutant N239Y, with its low redox potential for tyrosine and close proximity to the DNA (within 8 Å), is positioned to be oxidized and thus interrupt a disulfide exchange path. However, if DNA CT oxidizes cysteine 275 leading to a disulfide bond that is shuttled through disulfide exchange to cysteine 135 and then 141 and 124, the tyrosine would not have any optimal opportunity to be oxidized.

Covalent cross-linking. Along with oxidation of cysteines, we investigated the possibility of cross-linking during irradiation. Denaturing conditions (Figure 2.4) showed some covalent, nonreversible, nonredox dependent cross-linking also was observed in the AQ-DNA/p53 samples treated with the reductant beta-mercaptoethanol after irradiation (lane 6). Because the irradiated samples treated with reductant were within error of the cross-linked samples without reductant, it was unlikely that p53 is forming intramolecular disulfide bonds (which would be reversible with reductant) (Figure 2.4B). This covalent cross-linking occurred for all tested mutants (Figure 2.5), and the amount of cross-linking was within error for all mutants when quantified. In a related system, In assays of oxidative damage, cross-linked bands were observed for the anthraquinone-containg, irradiated samples; these bands appeared in both the samples with and without p53 (not shown). Kurbanyan et al. have shown that a photooxidant associated with DNA can oxidize a guanine base, leading to a guanine-protein cross-linking.[55] This guanine-protein cross-linking event is the likely explanation for p53-DNA cross-linking and may happen *in vivo*.

Figure 2.4. Example of a SDS-PAGE gel assay, stained for protein with Coomassie blue. Lanes: 1) molecular weight marker 2) DNA + C141S p53, dark control 3) DNA + C141S p53, irradiated 4) AQ-DNA + C141S p53, dark control 5) AQ-DNA + C141S p53, irradiated 6) AQ-DNA + C141S p53, irradiated, plus 15 mM beta-mercaptoethanol added after irradiation. [DNA] = [p53 tetramer] = 0.3 μM.

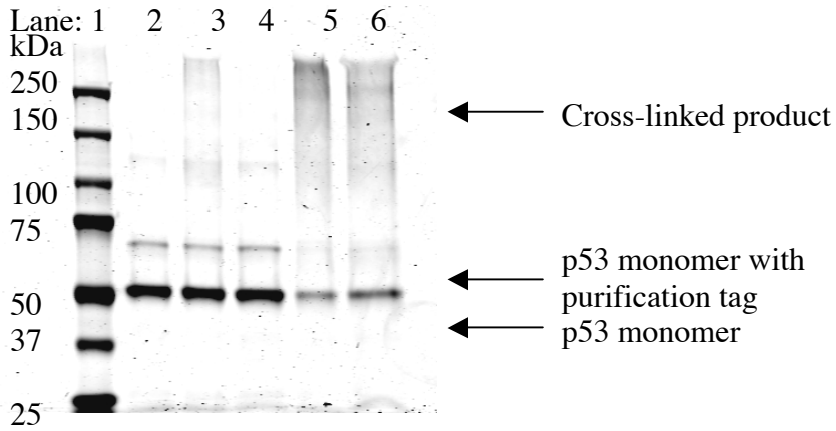
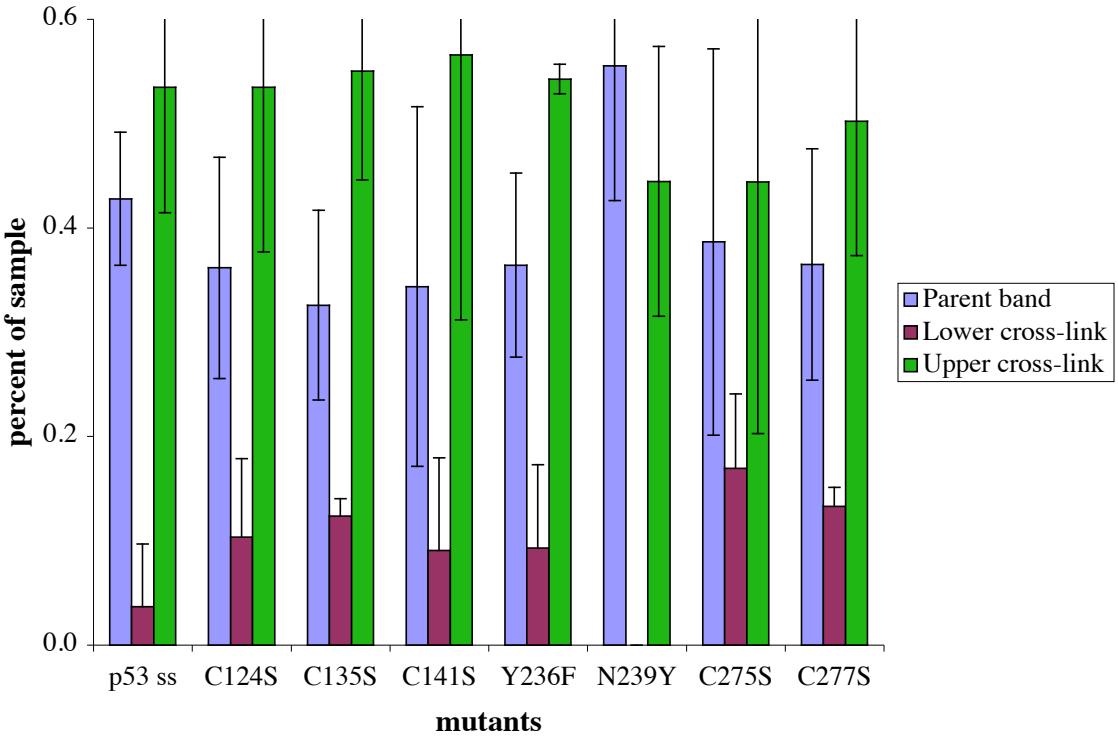


Figure 2.5. Results of the SDS-PAGE analysis that exhibit the cross-linking effects of DNA CT on various p53 mutants (example gel and assay conditions in Figure 2.4). Unless noted, sample concentration was 0.3 μM DNA and p53 tetramer. The amount of cross-linking was analyzed using Odyssey software. Data shown are the fraction of the total signal associated with the identified gel band. Error reported is standard deviation of three samples.



Identification of oxidized residues. To identify the cysteines oxidized through DNA CT, we used liquid chromatography coupled to nanoelectrospray, tandem mass spectrometric analysis of the digested peptides that have undergone differential cysteine modification. The order of modifications is shown in Figure 2.6. The mass of the modification reveals the oxidation state of the cysteine residue. The cysteines that were still in a reduced state after irradiation exhibit carbamidomethylation from reaction with iodoacetamide (an approximate mass gain of 57 amu). Half of the sample was digested with trypsin, and the other half was digested with chymotrypsin to provide a more complete survey of the 10 cysteines within p53. Then the samples were desalted, and the cysteines that had been participating in an oxidative modification are reduced and then reacted with N-methyliodoacetamide, with an approximate mass gain of 71 amu. The peptide samples were desalted again and then analyzed by LC-ESI-MS. The reactivity of iodoacetamide and N-methyliodoacetamide was approximately the same;[51] furthermore, there was little preference for peptides with one modification or the other in control MS samples (data not shown).

Spectral results for the oxidation of cysteine 141 are shown in Figure 2.7. Cysteine 141 was completely carbamidomethylated in the unirradiated sample; however, the irradiated sample showed cysteine 141-containing peptides modified with both alkylating reagents. Likewise, irradiated samples showed that cysteine 124 was also modified with both alkylating reagents. Therefore, p53SS afforded oxidation of cysteines 124 and 141. Control samples without photooxidant showed complete carbamidomethylation of all cysteines. The results for the presence of oxidized cysteines in the triple mutant of p53 under various conditions are shown in Table 2.2. From the

MS experiments, it was unclear if the disulfide bond was an inter- or intramonomer linkage within the p53 tetramer, but as seen in Figure 2.3, little of the cross-linking was redox-reversible, which points to the lack of intramonomer linkage.

After identification of the oxidized cysteine residues, the presence of the disulfide bond was confirmed by ruling out other oxidative events including sulfenic acid (transient cysteine oxidation) or sulfinic acid (permanent cysteine oxidation) formation, tyrosine, and methionine oxidation. The presence of sulfinic acid was not seen in the analysis with MASCOT software. When we treated the samples during or after irradiation with dimedone to trap sulfenic acid production, the peptides did not reflect the additional mass gain (data not shown). Methionine oxidation was seen occasionally in methionines outside of the DNA binding domain and farther than 30 Å away from the DNA, but at that distance, any effect would be negligible. Additionally, we observed some tyrosine oxidation in both the dark and irradiated samples, but this did not seem to affect binding or cysteine oxidation. The only explanation left is disulfide bond formation.

Figure 2.6. Flow chart of MS protocol 1) samples of equimolar DNA and tetramer p53 are irradiated; 2) 50 mM iodoacetamide is added to carbidomethylate free cysteines; samples are desalted by HPLC; 3) DTT quenches the iodoacetamide and breaks the disulfide bond; 4) N-methyliodoacetamide caps the remaining cysteines (shown in red); 5) DTT quenches N-methyliodoacetamide; 6) protein is digested with trypsin and LysC.

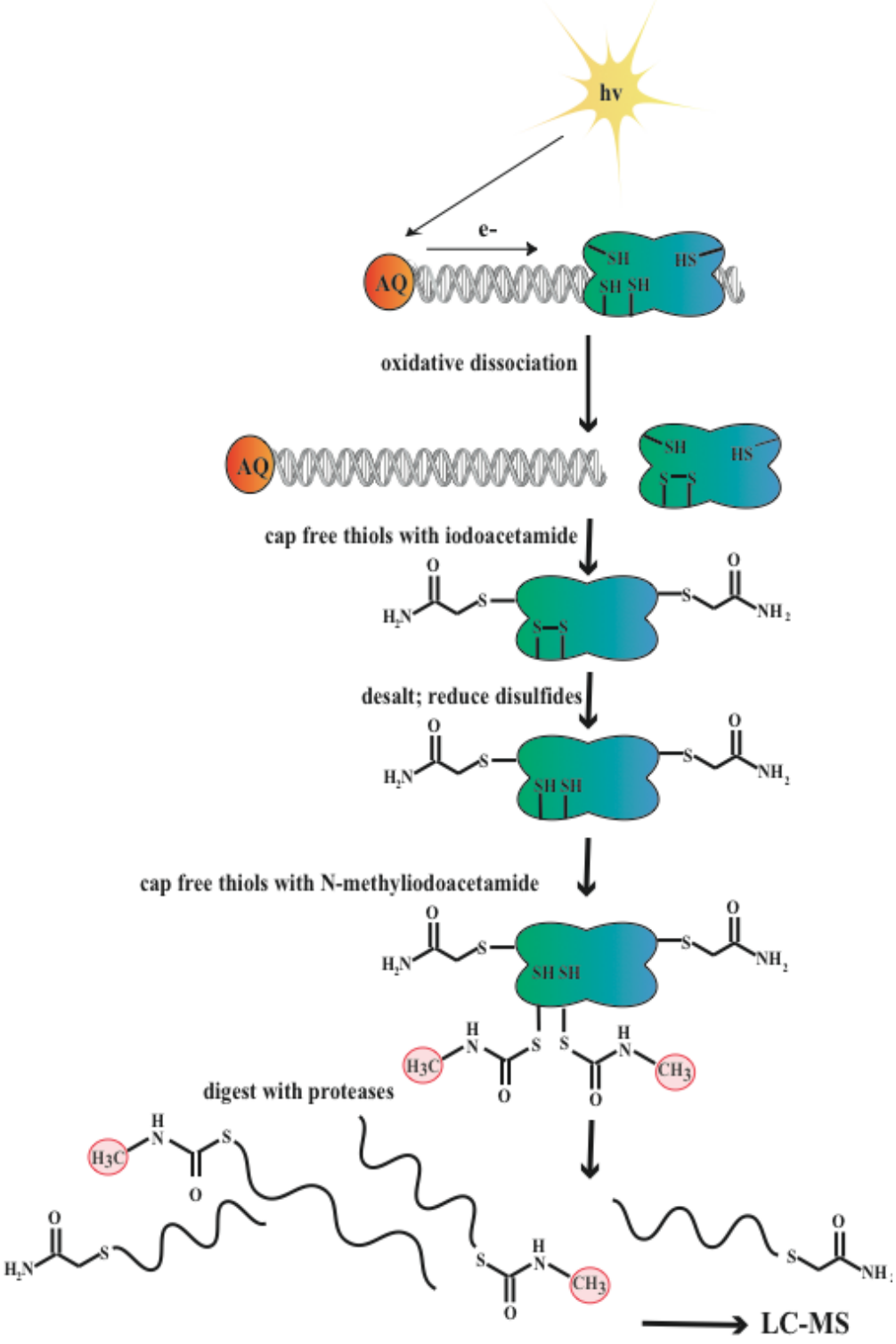


Figure 2.7. Cysteine MS spectra for an irradiated sample of 0.3µM DNA and p53SS treated as in Figure 2.6. Cysteine 141 is partially oxidized; both peptide results from differential thiol labeling are present. The peptide is TCPVQLWVDSTPPPGTR.

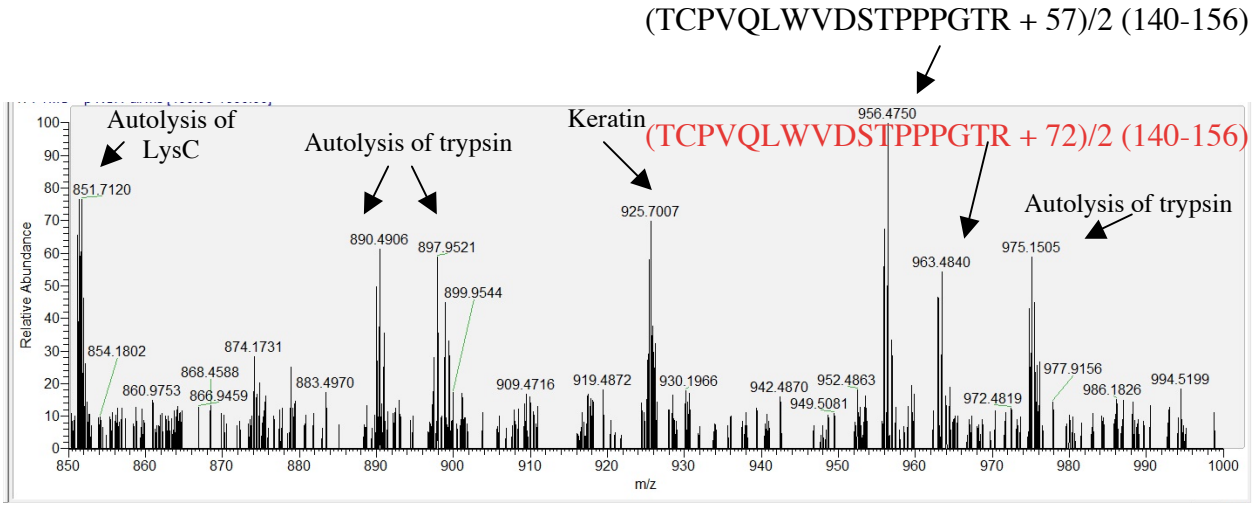


Table 2.2. Oxidation state of cysteine residues in p53

Cysteine	Dark control	Irradiated Sample
124	Reduced	Reduced and oxidized
135	Reduced	Reduced
141	Reduced	Reduced and oxidized
176	Not seen	Not seen
182	Reduced	Reduced
229	Reduced	Reduced
238	Not seen	Not seen
242	Not seen	Not seen
275	Reduced	Reduced
277	Reduced	Reduced

DISCUSSION

Oxidation of p53 at a distance by DNA CT. Here, we show that we can oxidize p53 at a distance and establish by MS that the oxidation product is a disulfide bond. Although disulfide bond formation is a 2-electron process, the anthraquinone photooxidant provides oxidation of only one electron; therefore, an understanding of how this reaction might occur is needed. The DNA-mediated oxidation can cause the thiol to form a radical cation, which will shed a proton and eventually attack a nearby thiol with the release of another proton.[27] The last electron in this process could be shuttled to molecular oxygen as trace amounts of oxygen are required for the formation of the disulfide-linked product.[27] In this experiment, the samples are irradiated while exposed to oxygen, which allows for photooxidant-initiated DNA-mediated disulfide bond formation in p53.

DNA CT leading to protein disulfide bond exchange. The disulfide bond exchange pathway in p53 (illustrated in Figure 2.8 and Table 2.3) would start with either cysteine 277 (which participates in a hydrogen bond to a base in the DNA, 3.6 Å away) or cysteine 275 (4.9 Å from the DNA, and 7.5 Å away from cysteine 277), and then progress to cysteines 135, 141, and 124 that would eventually form the observed and thus more stable disulfide bond. From the EMSA results in Figure 2.2, cysteine 275 is the key point in this disulfide exchange pathway. If residue 275 is a redox-silent serine residue, the disulfide bond exchange does not occur. All of these cysteines in the proposed disulfide exchange pathway are highly conserved.[8] Recently, Fersht and coworkers described the reactivity of the cysteine residues in a pseudo WT p53 containing the N239Y mutation to certain ligands with cysteines 141 and 124 as the most reactive

cysteines, which agrees with our data from the peptide analysis.[56] Additionally, Landridge-Smith and coworkers have shown that N-ethylmalemide reacts most readily with cysteines 182 and 277, but this result reflects how the difference ligand choice affects steric considerations and reaction mechanism that lead to different measures of reactivity for the different cysteines.[57]

This is a reasonable result given that cysteines 124 and 141 are within 20 Å of the DNA, and are thus close enough to be oxidized at a distance through a disulfide exchange pathway; recently, more of these exchange pathways are being discovered, many with regulatory functions.[58,59] While it has been shown that the maximum absolute distance allowed between the α-carbons of two cysteine residues that may form a disulfide bond is 4.6 Å, thermally induced backbone motion allows for disulfide bond formation over distances of up to 20 Å.[60] The p53 DNA binding region has, through NMR studies, been shown to have a flexible and dynamic hydrophobic core, so the initiation and continuation of a disulfide bond exchange could occur.[61]

Method of identification for disulfide bonds. There are few methods for direct detection of disulfide bonds in proteins in the literature,[62] and none of them currently are applicable to a protein as large and complex as p53. The literature supports the differential thiol method as reported here as long as the sensitivity is high enough to detect the various peptides, and for larger proteins, if the enzyme digestions cover most of the protein.[11,63] While many methods of differential thiol labeling employ tags with the same chemical formulae but different isotope content to avoid any differential reactivity, we instead chose a set of modifications, from iodoacetamide and N-methyliodoacetamide, that fortunately did not exhibit any difference in reactivity.

Furthermore, we bypassed steric issues by denaturing p53 in guanidinium chloride so that the internal cysteines would not be blocked by other residues. By using a differential thiol labeling, plus scanning for other oxidative modifications, we are able to rule out other possibilities of oxidative modifications. Methionine oxidation and tyrosine oxidation have been linked to the oxidative response in the cell,[64,65] but in this case, the oxidations seen were far away from the DNA and were not linked to the binding capabilities of the protein.

Figure 2.8. Structural environment of cysteine residues that could be involved in disulfide bond formation. A. Monomeric p53. Cysteines are in red. B. A close-up of the DNA-p53 interface (PDB: 1TSR).

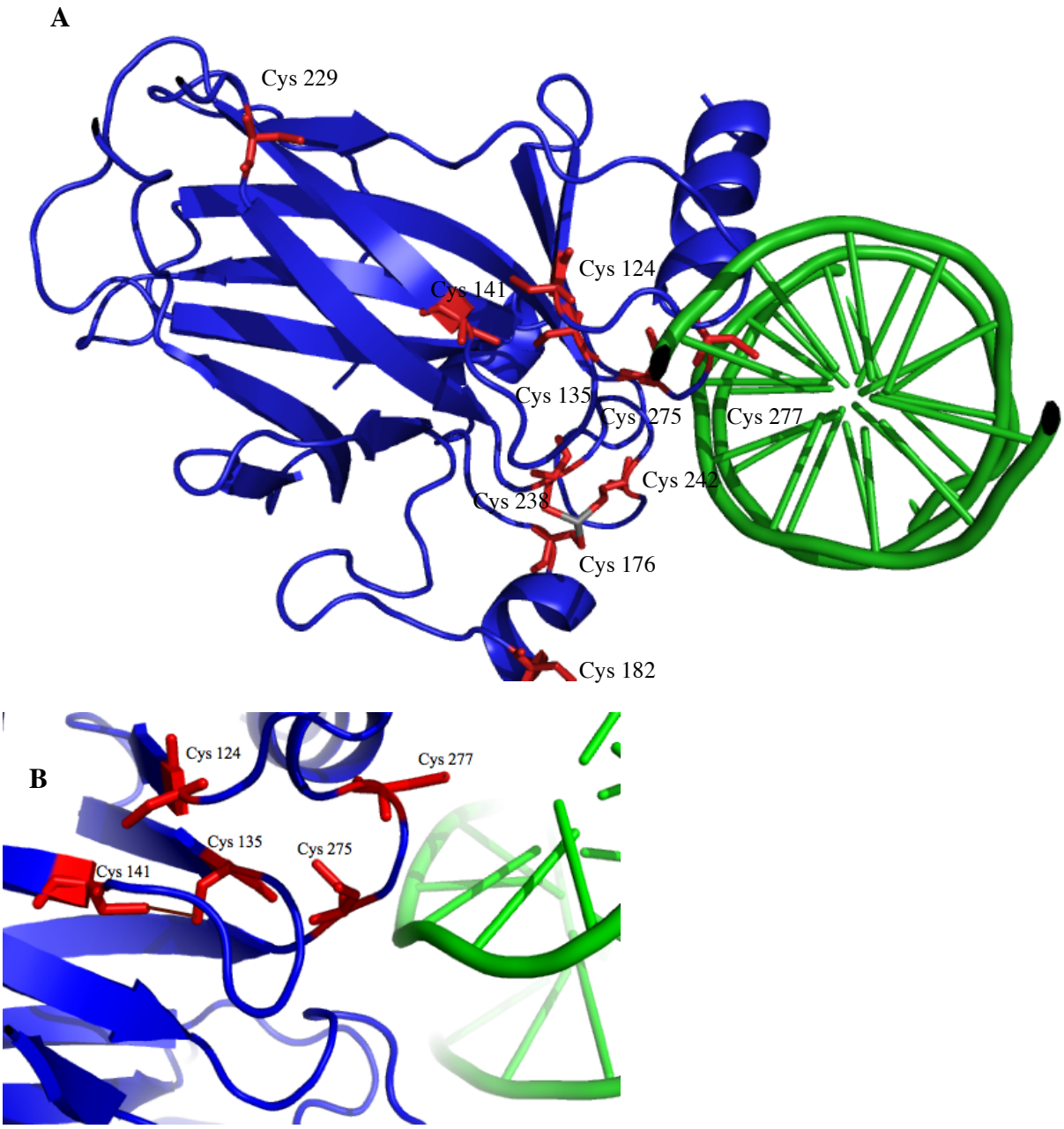


Table 2.3. Distance between cysteine residues (measured sulfur to sulfur atom) (Å)

DNA-Cys 277[1]	3.6
DNA-Cys 275[1]	4.9
Cys 275-Cys 277[2]	7.5
Cys 135-Cys 275[2]	6.9
Cys 135-Cys 141[2]	3.4
Cys 124-Cys 135[2]	5.5
Cys 124-Cys 141[2]	3.9

[1] DNA to sulfur atom (Å)
[2] Sulfur to sulfur atom (Å)

The oxidation of cysteines versus the oxidation of tyrosines. While the p53 oxidation results from the EMSA assays are within error of each other, the other biochemical analyses in the literature point to a different chemical reactivity between the proteines. N239Y, while stabilizing p53,[66] is linked to rectal cancer.[39,40] Conversely, it also shows greater activation of proapoptotic cellular signaling pathways, specifically through the signaling protein Bax.[67] N239Y replaces asparagine, a CT-neutral residue with the larger amino acid tyrosine, which has the approximate reduction potential of cysteine (both 0.9 V at ph 7).[38] This shift in redox potential confers the possibility to change the charge transport pathway of p53. Tyrosine is as easily oxidized as cysteine,[38] but the reactivity of the residues is highly dependent on their immediate environment. Tyrosine oxidation products were not seen in the MS peptide results. Also, the EMSA results for the oxidation of the quadruple mutant are statistically identical with those of the triple mutant (Figure 2.3), so the extent of oxidation is most likely the same. Furthermore, the extent of DNA-p53 cross-linking as shown in Figure 2.5 is equivalent for the N239Y mutant as the pseudo-wild type p53 that does not have any tyrosines within 15 Å of the DNA.

Implications of protein-DNA cross-linking. As reported elsewhere, oxidative conditions favoring the oxidation of guanine often lead to DNA-protein cross-linking. In Figure 2.5, cross-linking of the various mutants to p53 showed that the cross-linking was not a specific cysteine-dependent, or N239Y, or Y236F, process. We do not see evidence for intra-molecular disulfide bonds in this experiment, or other reversible redox bonds. Furthermore, future experiments should switch to Western blots of the protein for two reasons. The assurance that the cross-linking protein is p53 will be guaranteed, and

Western blots require loading controls, which should make quantification easier and more reliable.

Lysine and tyrosine are the two most likely residues to participate in cross-linking.[68] In the crystal structure of the DNA core domain,[9] lysine 120 associates with the DNA near a guanine doublet; moreover, it is 4.86 Å away from the guanine base itself. Because DNA-protein cross-linking blocks transcription and replication, and are lethal if not repaired,[69,70] p53-DNA cross-links induced by oxidative CT may have severe effects on the ability of the cell to function, perhaps signaling to the cell that it should initiate apoptosis. This signaling is predicted by our model of p53 regulation by DNA CT: oxidation at a distance leads to cell death.

These intriguing cross-linking studies could be further bolstered by several control experiments. With irradiation, anthraquinone photooxidants may form superoxide radicals that could initiate random cross-linking separate from DNA CT effects.[59] To show that DNA CT initiates this cross-linking, the introduction of a mismatch in the bases separating the photooxidant from the p53 binding site should inhibit cross-linking. Also, the inclusion of a superoxide "sink," such as superoxide dismutase, should not have an effect on the cross-linking. Further chemical characterization of the cross-linked products should also be demonstrated.

*Cysteines involved in the oxidation of p53 under various conditions***.** Thioredoxin-human p53 interactions in engineered yeast are dependent on cysteine 275, as well as the three cysteines that coordinate the structural zinc ion.[71] Cysteine 182 was a candidate for oxidation as demonstrated by two other studies.[11,12] Intriguingly, the redox-dependent differential binding of p53 from UVC-irradiated cells to recognition sites was

determined to be contingent on cysteine 277,[7] which is not positioned to be the protein CT "gateway" in our model. Here, the key residue is cysteine 275, although the redox state of the cysteines coordinated to the zinc ion was not seen in the MS analysis. Studies completed by Srivenugopal and coworkers indicated that cysteine 141 in cellular p53 can be modified by glutathionylation, which requires the formation of a mixed disulfide bond, after oxidant and DNA-damaging treatments.[6] In these studies, either p53 without additional treatment was used, or it was collected from cells undergoing some general form of oxidizing or DNA-damaging stress.

The multiple cysteine residues on p53 allow it exquisite control over its response to different types of oxidative conditions. If the oxidation is DNA-mediated, the cysteine based network will channel the oxidation away from the DNA and to cysteines that will then form a disulfide bond, undergo disulfide bond exchange to funnel it deeper into the protein, and dissociate from genes not needed for the cellular response. If the cellular conditions in general are oxidizing, glutathionylation will occur, prohibiting p53 from binding to DNA and activating gene transcription.[6] Cysteine 182 is also well positioned on the surface of p53 to intercept diffusing oxidizing molecules such as hydrogen peroxide and protect the integrity of nuclear DNA,[12] and cysteine 277, coordinated to the DNA, may serve as a last defense for DNA against non-CT oxidative stress.[7]

C275S has been correlated with lung cancer,[49] a decidedly deadly disease. Lung cancer kills more people than any other type of cancer.[72] With its constant exposure to the environment, lung tissue is often exposed to chemicals that cause oxidative stress, such as cigarette smoke,[73] which is highly linked to lung cancer.[73] If signaling to p53 through DNA CT is inhibited by C275S in lung tissue, the lung cells would be at greater risk for

tumorigenesis because the cell would not, through this pathway, initiate the p53-dependent correct response to severe stress: apoptosis. While some cellular pathways are redundant, p53 occupies a unique, central position for many signaling pathways,[37] and therefore, disruptions in p53 function signaling often have a disparate effect on cellular fate.[1]

CONCLUSIONS

p53 regulation by DNA CT is an exciting new chapter in the study of the "guardian of the genome." Determining the type and extent of this kind of chemistry at a distance is a key step in comprehending the role of DNA CT in biological systems. The results so far indicate that DNA-mediated oxidation of p53 is a chemically distinct way for the cell to enable a response to extensive oxidative damage to the DNA, as opposed to general oxidative insults. Cysteine 275 is the key amino acid governing the oxidation of p53. Formation of a disulfide bond between cysteines 124 and 141 must happen for this signaling event to occur. Covalent cross-linking, which may prohibit transcription inside the cell, is also observed between DNA and p53 after induction of oxidative DNA CT *in vitro*. Future studies will continue to elucidate the details of this fascinating process and enhance understanding of the complex cellular stress response.

Acknowledgements: We are grateful to the Ellison Foundation for financial support. We also thank the Protein Exploration Laboratory at the Beckman Institute for equipment support.

REFERENCES

(1) Vogelstein, B., Lane, D., and Levine, A.J. *Nature* **2000**, *408*, 307-310.

(2) Brosh, R., and Rotter, V. *Nat. Rev. Cancer* **2009**, *9*, 701-713.

(3) Kato, S., Han, S. Y., Liu, W., Otsuka, K., Shibata, H., Kanamaru, R., Ishioka, C. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 8424-8429.

(4) el-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W., and Vogelstein, B. *Nat. Gen.* **1992**, *1*, 45-49.

(5) Beckerman, R., and Prives, C. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a000935.

(6) Velu, C. S., Niture, S. K., Doneanu, C. E., Pattabiraman, N., and Srivenugopal, K. S. *Biochem.* **2007**, *46*, 7765-7780.

(7) Buzek, J., Latonen, L., Kurki, S., Peltonen, K., and Laiho, M. *Nucl. Acids Res.* **2002**, *30*, 2340-2348.

(8) Paveletich, N. P., Chambers, K. A., and Pabo, C. O. *Genes Dev.* **1993**, *7*, 2256-2564.

(9) Cho, Y., Gorina, S., Jeffrey, P. D., and Pavletich, N. P. *Science* **1994**, *265*, 346-355.

(10) Seo, Y. R., Kelley, M. R., and Smith, M. L. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14548-14553.

(11) Held, J. M., Danielson, S. R., Behring, J. B., Atsriku, C., Britton, D. J., Puckett, R. L., Schilling, B., Campisi, J., Benz, C. C., Gibson, B. W. *Mol. Cell. Prot.* **2010**, *9*, 1400-1410.

(12) Sun, X. Z., Vinci, C., Makmura, L., Han, S., Tran, D., Nguyen, J., Hamann, M., Grazziani, S., Sheppard, S., Gutova, M., Zhou, F., Thomas, J., Momand, J. *Antioxid. Redox Signal.* **2003**, *5*, 655-666.

(13) Rainwater, R., Parks, D., Anderson, M. E., Tegtmeyer, P., and Mann, K. *Mol. Cell. Biol.* **1995**, *15*, 3892-3903.

(14) O'Neill, M. A., and Barton, J. K. *Topics Curr. Chem.* **2004**, *236*, 67-115.

(15) Hall, D. B., Kelley, S.O., and Barton, J. K. *Biochem.* **1998**, *37*, 15933-15940.

(16) Nunez, M. E., Hall, D. B., and Barton, J. K. *Chem. & Biol.* **1999**, *6*, 85-97.

(17) Sugiyama, H., and Saito, I. *J. Am. Chem. Soc.* **1996**, *118*, 7063-7068.

(18) Saito, I., Takayama, M., Sugiyama, H., and Nakatani, K. *J. Am. Chem. Soc.* **1995**, *117*, 6406-6407.

(19) Hall, D. B., Holmilin, R. E., and Barton, J. K. *Nature* **1996**, *382*, 731-735.

(20) Merino, E. J., and Barton, J. K. *Biochem.* **2007**, *46*, 2805-2811.

(21) Nunez, M. E., Holmquist, G. P., and Barton, J. K. *Biochem.* **2001**, *40*, 12465-12471.

(22) Merino, E. J., and Barton, J. K. *Biochem.* **2008**, *47*, 1511-1517.

(23) Slinker, J. D., Muren, N. B., Renfrew, S. E., Barton, J. K. *Nat. Chem.* **2011**, *3*, 230-235.

(24) Rajski, S. R., Kumar, S., Roberts, R. J., and Barton, J. K. *J. Am. Chem. Soc.* **1999**, *121*, 5615-5616.

(25) Rajski, S. R., and Barton, J. K. *Biochem.* **2001**, *40*, 5556-5564.

(26) Hall, D. B., and Barton, J. K. *J. Am. Chem. Soc.* **1997**, *119*, 5045-5046.

(27) Takada, T., and Barton, J. K. *J. Am. Chem. Soc.* **2005**, *127*, 12204-12205.

(28) Boal, A. K., Genereux, J. C., Sontz, P. A., Gralnick, J. A., Newman, D. K., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15237-15242.

(29) Lee, P. E., Demple, B., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13164-13168.

(30) Merino, E. J., Davis, M. L., and Barton, J. K. *Biochem.* **2009**, *48*, 660-666.

(31) Nunez, M. E., Noyes, K. T., and Barton, J. K. *Chem. Biol.* **2002**, *9*, 403-415.

(32) Romano, C. R., Sontz, P. A., Barton, J. K. *Biochem.* **2011**, epub June 9.

(33) Lin, J. C., Singh, R. R., Cox, D. L. *Biophys. J.* **2008**, *95*, 3259-3268.

(34) Augustyn, K. E., Merino, E., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 18907-18912.

(35) Abbas, T., and Dutta, A. *Nat. Rev. Cancer* **2009**, *9*, 400-414.

(36) Jung, H. J., Kim, E. H., Mun, J.-Y., Park, S., Smith, M. L., Han, S. S., and Seo, Y. R. *Oncogene* **2007**, *26*, 7517-7525.

(37) Vousden, K. H., and Lane, D. P. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 275-283.

(38) Milligan, J. R., Tran, N. Q., Ly, A., Ward, J. F. *Biochem.* **2004**, *43*, 5102-5108.

(39) Rebischung, C., Gerard, J. P., Gayet J., Thomas, G., Hamelin, R., and Laurent-Puig, P. *Int. J. Cancer* **2002**, *100*, 131-135.

(40) Webley, K. M., Shorthouse, A. J., and Royds, J. A. *J. Pathol.* **2000**, *191*, 361-367.

(41) Hiroshima, K., Toyozaki, T., Kohno, H., Ohwada, H., Fujisawa, T. *Pathol. Int.* **1998**, *48*, 869-876.

(42) Brattström, D., Bergqvist, M., Lamberg, K., Kraaz, W., Scheibenflug, L., Gustafsson, G., Inganas, M., Wagenius, G., and Brodin, O. *Med. Onco.* **1998**, *15*, 255-261.

(43) Dumaz, N., Drougard, C., Sarasin, A., and Daya-Grosjean, L. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10529-10533.

(44) Calistri, D., Rengucci, C., Seymour, I., Lattuneddu, A., Polifemo, A. M., Monti, F., Saragoni, L., and Amadori, D . *J. Cell Physiol.* **2005**, *204*, 484-488.

(45) Hernández, S., López-Knowles, E., Lloreta, J., Kogevinas, M., Jaramillo, R., Amorós, A., Tardón, A., García-Closas, R., Serra, C., Carrato, A., Malats, N., and Real, F. X. *Clin. Cancer Res.* **2005**, *11*, 5444-5450.

(46) Russo, A., Migliavacca, M., Zanna, I ., Valerio, M. R., Latteri, M. A., Grassi, N., Pantuso, G., Salerno, S., Dardanoni, G., Albanese, I., La Farina, M., Tomasino, R. M., Gebbia, N., and Bazan, V. *Cancer Epidemiol. Biomarkers Prev .* **2002**, *11*, 1322-1331.

(47) Lu, M. L., Wikman, F., Orntoft, T. F., Charytonowicz, E., Rabbani, F., Zhang, Z., Dalbagni, G., Pohar, K. S., Yu, G., and Cordon-Cardo C.. *Clin. Cancer Res.* **2002**, *8*, 171-179.

(48) Uchino, S., Noguchi, M., Ochiai, A., Saito, T., Kobayashi, M., and Hirohashi, S. *Int. J. Cancer* **1993**, *54*, 759-764.

(49) Gao, H. G., Chen, J. K., Stewart, J., Song, B., Rayappa, C., Whong, W. Z., and Ong, T. *Carcinogenesis* **1997**, *18*, 473-478.

(50) Hongyo, T., Hoshida, Y., Nakatsuka, S., Syaifudin, M., Kojya, S., Yang, W. I., Min, Y. H., Chan, H., Kim, C. H., Harabuchi, Y., Himi, T., Inuyama, M., Aozasa, K., and Nomura, T. *Oncol. Rep.* **2005**, *13*, 265-271.

(51) Ramseier, U., and Chang, J.-Y. *Anal. Biochem.* **1994**, *221*, 231-233.

(52) Vallejo, A. N., Pogulls, R. J., and Pease, L. R. In *Cold Spring Harbor Protocols: 2008*; Cold Spring Harbor Laboratory Press: 2008; Vol. 2008.

(53) Veprintsev, D. B., Freund, S. M., Andreeva, A., Rutledge S. E., Tidow, H., Canadillas, J. M., Blair, C. M., and Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2005**, *103*, 2115-2119.

(54) Weinberg, R. L., Veprintsev, D. B., Bycroft, M., and Fersht, A. R. *J. Mol. Biol.* **2005**, *348*, 589-596.

(55) Kurbanyan, K., Nguyen, K. L., To, P., Rivas, E. V., Lueras, A. M., Kosinski, C., Steryo, M., González, A., Mah, D. A., and Stemp, E. D *Biochem.* **2003**, *42*, 10269-10281.

(56) Kaar, J. L., Basse, N., Joerger, A. C., Stephens, E., Rutherford, T. J., and Fersht, A.R. *Prot. Sci.* **2010**, *19*, 2267-2278.

(57) Scotcher, J., Clarke, D. J., Weidt, S. K., Mackay, C. L., Hupp, T. R., Sadler, P. J., and Langridge-Smith, P. R. R. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 888-897.

(58) Messens, J., Van Molle, I., Vanhaesebrouck, P., Limbourg, M., Van Belle, K., Wahni, K., Martins, J. C., Loris, R., Wyns, L. *J. Mol. Biol.* **2004**, *339*, 527-537.

(59) Kadokura, H., Katzen, F., and Beckwith J. *Annu. Rev. Biochem.* **2003**, *72*, 111-135.

(60) Bass, R. B., Butler, S. L., Chervitz, S. A., Gloor, S. L., and Falke, J. J. *Methods Enzymol.* **2007**, *423*, 25-51.

(61) Cañadillas, J. M. P., Tidow, H., Freund, S. M. V., Rutherford, T. J., Ang, H. C., and Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2109-2114.

(62) Kim, H. I., Beauchamp, J. L. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 157-168.

(63) Hurd, T. R., Prime, T. A., Harbour, M. E., Lilley, K. S., Murphy, M. P. *J. Biol. Chem.* **2007**, *282*, 22040-22051.

(64) Cao, G., Lee, K. P., van der Wijst, J., de Graaf, M., van der Kemp, A., Bindels, R. J. M., Hoenderop, J. G. J. *J. Biol. Chem.* **2010**, *285*, 26081-26087.

(65) Candeias, L. P., Turconi, S., Nugent, J. H. A. *Biochim. Biophys. Acta* **1998**, *1363*, 1-5.

(66) Nikolova, P. V., Henckel, J., Lane, D. P. Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14675-14680.

(67) Khoo, K. H., Mayer, S., and Fersht, A. R. *J. Biol. Chem.* **2009**, *284*, 30974-30980.

(68) Xu, S., Muller, J. G., Ye, Y., and Burrows, C. J. *J. Am. Chem. Soc.* **2008**, *130*, 703-709.

(69) Reardon, J. T., Cheng, Y., and Sancar, A. *Cell Cycle* **2006**, *5*, 1366-1370.

(70) Minko, I. G., Kurtz, A., Croteau, D. L., Van Houten, B., Harris, T. M., and Lloyd, R. S. *Biochem*. **2005**, *44*, 3000-3009.

(71) Stoner, C. S., Pearson, G. D., Koc, A., Merwin, J. R., Lopez, N. I., and Merrill, G. F. *Biochem*. **2009**, *48*, 9156-9169.

(72) Department of Health and Human Services: Centers for Disease Control and Prevention; Registries, National Program of Cancer, Ed.; USA.gov: Atlanta, 2007; Vol. 2011, p http://apps.nccd.cdc.gov/uscs/.

(73) Carnevali, S., Petruzzelli, S., Longoni, B., Vanacore, R., Barale, R., Cipollini, M., Scatena, F., Paggiaro, P., Celi, A., and Giuntini, C. *Am. J. Physiol. Lung Cell. Mol. Physiol*. **2003**, *284*, L955-L963.

CHAPTER 3

Biological Effects of DNA Sequence on DNA-Mediated p53 Regulation

*Acknowledgments: Katie Schaefer did all of the K_d assays and EMSA experiments in this chapter. Paul Lee synthesized and purified $[Rh(phi)_2bpy]Cl_3$.

INTRODUCTION

DNA-mediated charge transport (DNA CT) occurs when charge is passed through the π-stack formed by the heterocyclic bases;[1] it has been extensively studied both *in vitro* and in biological systems.[2-4] This sensitive, fast process is influenced greatly by the identity of the base pairs in the sequence. First, because guanine has the lowest oxidation potential of the bases, guanine doublets and triplets are the most likely oxidized locations.[5,6] Second, the sequence dependence of DNA CT increases with the distance traveled.[2,7,8] Additionally, sequence dictates structure; DNA CT best occurs in B-form DNA,[9-11] and sequence effects that lead to A- or Z-form may decrease CT processes.[12] Mismatched base pairs and acute bending in the DNA induced by protein binding all attenuate DNA CT;[1,9,13-15] however, gaps in the sugar phosphate backbone of DNA do not attenuate DNA CT.[16] In a more subtle demonstration of the sensitivity of DNA CT, the stronger base stacking exhibited by purine tracts, especially adenine runs, is favored by oxidative DNA CT,[17-20] while reductive DNA CT favors the more easily reduced pyrimidine stretches.[21] Sequence also affects dynamical processes that are required for DNA CT.[20,22]

In organisms, DNA CT may lead to wide-ranging genetic outcomes. In mitochondria, DNA CT-induced damage accumulates in regulatory sequences with high G/C content, such as regulatory element 2.[23,24] After several rounds of induced oxidative CT, G/C regions in guanine-rich regulatory regions mutate to A/T regions in mitochondria, showing exquisite regulation and adaptation in the face of oxidative stress.[25] In nuclei undergoing photooxidant-induced DNA CT, damage was localized to the p53 gene and a transcriptionally active promoter site for phosphoglycerate kinase

(PGK1), which was protected from direct photooxidant binding by the association of transcription factors.[26] BER proteins, such as Endo III and MutY, use DNA CT to locate damaged bases.[27] DNA CT has been shown to regulate the superoxide response system in *E. coli* through DNA-mediated oxidation of one of the iron-sulfur clusters of the transcription factor SoxR.[28,29] However, the effect of the DNA sequence, namely G/C base pair placement, and the subtle effects that guanine doublets and triplets may have on the structure and function of DNA CT on proteins inside the cell have not been addressed to the full extent.

The global transcription factor p53 is a key player in the regulation of the oxidative stress response and other vital cellular pathways.[30] Mutant p53 has been a target of much scientific research; just under 50% of all tumors contain mutant p53, and 80% of its mutations are found in the DNA-binding region of the protein. The protein p53 binds as a tetramer to a variety of different sequences, which usually exhibits two repetitions of the pattern 5'-PuPuPuC(A/T)(A/T)GPyPyPy-3' separated by a spacer of 0-13 base pairs (Pu and Py represent purines and pyrimidines, respectively).[31] The protein p53 may bind to sites that are not a 100% match, and sometimes p53 may not bind to sites that parallel the recognition sequence.[31] The recognition sequences of p53 may be present anywhere in the locus of the regulated gene, but usually the recognition sequences are found in the promoter or noncoding regions; moreover, the closer the sequence is to the gene, the greater the transcriptional control p53 may exert.[32]

The protein p53 requires a structural zinc ion bound to histidine 179 and the reduced cysteines 176, 238, and 242 in order to bind to DNA. The C-terminus of p53 assists the core DNA-binding region in the search for DNA binding regions, and eventual

regulation of function through acetylation.[32] Additionally, p53 exhibits redox sensitivity when bound to certain sequences; while p53 that is completely reduced will bind to a Gadd45 recognition sequence, its overall binding to the Gadd45 sequence will decrease by half if it is removed to nonreducing conditions.[33] However, p53 will remain bound to the p21 binding sequence in both reducing and nonreducing conditions.[33] Likewise, p53 shows greater binding to GC rich sequences that match its recognition element under reducing conditions, but it shows relative insensitivity to its redox state when bound to AT rich sequences that also match the general recognition element.[33]

From extensive analyses of p53 interaction with DNA, a few general trends have emerged. Many apoptotic genes exhibit higher dissociation constants, K_D's, for p53 than other cell growth regulating genes;[34] however, there are enough counterexamples that this can only be described as a pattern, not an absolute necessity. Also, p53 tetramers have been observed forming different binding modes and organization based on the sequence of the bound DNA, post-translational modifications, and associations with other proteins; this leads to a wider range of transcriptional outcomes and exquisite control over different promoter regions.[35]

Previous work in the Barton group showed that p53 was oxidized at a distance through oxidative DNA CT. As reported in Augustyn et al.,[36] photooxidation of conjugated anthraquinone-DNA, containing p53 recognition sequences, leads to DNA-mediated oxidation of the protein. However, this effect was sequence dependent. For instance, when p53 was bound to DNA (modified as above) that contained either the MDM2 (the principle antagonist of p53)[37] recognition site or the Gadd45 (involved in DNA repair)[38] recognition site, p53 was oxidized. If the DNA contained the p21 (cell

cycle arrest)[39] recognition site, p53 did not show any appreciable oxidation, although it had the same G/C content as the Gadd45 site. In the cellular context, it is hypothesized that this sequence dependence is driven by cell fate decisions of p53. Oxidative DNA CT happens under severely stressful conditions, and the severely stressed cell will down-regulate DNA repair genes (Gadd45) and genes (MDM2) that limit key apoptotic players, like p53, in favor of genes that will reflect the threat to genomic integrity (p21 as a cell cycle arrest protein). Oxidized p53 was also observed through a Western blot of lysate from HCT116N cells that had undergone induced CT. In the previous paper,[36] it was suggested that cysteine 141 forms a disulfide bond, and in the previous chapter of this thesis, the molecular oxidation of p53 is discussed more thoroughly.

Furthermore, sequence-specific studies have been continued.[40] Using the synthetically derived consensus sequence that we had used for initial studies in EMSA assays, and increasing or decreasing the amount of guanine doublets and triplets in the purinic region of the consensus sequence, we have found that DNA-mediated p53 oxidation follows the pattern AAA<GGA<GGG/GGG (both strands of the duplex) <GGG (same strand of the duplex). When assays of oxidative damage to the DNA were performed, the damage to the DNA was localized in areas of guanine doublets and triplets, with damage to those guanine sites following the same trend as the DNA-mediated p53 oxidation. However, adding p53 to those samples limited the amount of damage inflicted on the DNA, with near-complete rescue of the DNA occurring when the ratio of DNA to p53 was 1:1.

Here, we investigate the effect of sequence content of the p53 regulatory elements on DNA-mediated p53 oxidation at a distance. Biologically important p53 recognition

sequences matching the p53 consensus sequence, and covalently attached to anthraquinone photooxidants, were constructed with varying guanine doublet or triplet content and subjected to electrophoretic mobility shift assays (EMSA's) after irradiation. The genes were chosen based on their p53 binding sequence content (Figure 3.1) and the function of the genes: S100A2 (S100 calcium binding protein A2) is a protein thought to be involved with tumorgenic suppression,[41,42] and it has two sets of guanine triplets in its p53 binding site; ODC1 (ornithine decarboxylase 1) is an enzyme that catalyzes putresicine, which is toxic to cells and thus important for cell death,[43] and it has multiple guanines on both sides of the duplex. Caspase-1, in addition to having mostly adenines in its p53 binding sequences, is important for apoptotic function.[44] These assays correlated well with the conclusions of the EMSA assays on synthetic sequences. Moreover, these data are biologically significant. RT-qPCR of the mRNA products shows possible regulation of p53 through DNA CT within the organism.

MATERIALS AND METHODS

Electrophoretic mobility shift assay (EMSA). All oligomers were constructed as Augustyn et al.[36] and are shown in Figure 3.1. Protein production was accomplished using the V203A/N268D/M133L (3x) mutant of p53, as described in Chapter 2 of this thesis, and the purification of p53 was carried out with minor modifications according to the procedure established by Fersht and coworkers and detailed in Appendix 1.[45] K_d's were measured as in Chapter 2, and are shown in Table 1. EMSA's after sample irraditation proceeded as described in Chapter 2, and the images were quantified in ImageQuant. The ratio of protein to DNA was always 1:1, although the exact

concentrations were calculated to reach 80% binding, based on the determined K_d's. Error reported is standard deviation and reflects at least 3 replicates.

RT-qPCR. Reverse transcriptase-quantitative PCR was performed to monitor the changes in transcription after induced DNA CT in cancer cells displaying wild-type p53. HCT116N cells were maintained at 37 °C with 5% carbon dioxide in RPMI medium 1640 supplemented with 10% FBS, 2 mM L-glutamine, 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, 100 units/mL penicillin, 100 μg/mL streptomycin, and 400 μg/mL Geneticin (G418) in tissue culture flasks and dishes (Corning Costar, Acton, MA). The rhodium photooxidant $[Rh(phi)_2bpy]Cl_3$ was synthesized, as described previously,[46] by Paul Lee. After 2 passages, cells were grown to 50% confluence in clear six-well plates (4 plates for the four conditions—untreated, irradiated, rhodium treated, and rhodium treated and irradiated), and 10 μM $[Rh(phi)_2bpy]Cl_3$ in phosphate buffered saline (PBS) or the same volume of only PBS was added to fresh media (the solution never more than 10% total of the media; otherwise, the cells die overnight) and grown overnight. The next morning, the cells were washed 2x with PBS and grown 30 min in media containing 10 μM cisplatin (dissolve in PBS first). After 2x washes in PBS, the cells were grown in media for 2 hr, and subsequently washed and irradiated with the solar simulator from Oriel (Darmstadt, Germany,1000-W Hg/Xe (340-440 nm)) in PBS (1 mL per well) for 1 hr. The cells were washed with PBS and grown 4 hours; then the cells were trypsinized for removal from the dishes, centrifuged, and washed with PBS before being frozen in liquid nitrogen and stored at –80ºC.

RNA isolation was accomplished with the RNeasy kit from Qiagen; cells were re-suspended in 350 μL lysis buffer, and they were lysed by passage 15-20 times through a

21-gauge needle, then the samples were treated according to the manufacturer's instructions and storage at –80ºC. cDNA libraries were constructed through one round of PCR using the Transcriptor First Strand cDNA Synthesis Kit from Roche. A PCR reaction using water instead of mRNA was also set up as a blank sample. The PCR program that was used had the following parameters: 10 min at 25ºC, followed by 30 min at 50ºC, and 5 min at 85ºC. The libraries were diluted 20x with ultrapure, sterile water, and stored at -80ºC.

Quantitative PCR was accomplished with the SybreGreen master mix (2x) available from Roche, primers from and validated by realtimeprimers.com, and carried out on an ABI 7300 qPCR instrument (Applied Biosystems) according to the protocol from the manufacturer. All primers were originally 50 μM in PBS, and diluted to 0.1 μM for quantitative PCR in Sybre Green master mix solution (Roche). The primers were as follows: human S100 calcium binding protein A2, forward: 5' GATCAGGTTGAGGCAGGTTT 3', reverse: 5' TCATTTCCCCCTTACTCAGC 3'; human ornithine decarboxylase 1 (ODC1), forward: 5' CACATGTAAAGCCCCTTCTG 3', reverse: 5' TAGTAGATCGTCGGCCTCTG 3'; caspase-1, forward: 5' TACAGAGCTGGAGGCATTTG 3', reverse: 5' GATCACCTTCGGTTTGTCCT 3'; Gadd45, forward: 5' AGAGCAGAAGACCGAAAGGA 3', reverse: 5' CAGAGCCACATCTCTGTCGT 3'; p21/Waf-1: forward: 5' AAACTTTGGAGTCCCCTCAC 3', reverse: 5' AAAGGCTCAACACTGAGACG 3'; human glyceraldehydes phosphate dehydrogenase (GAPD), forward: 5' GAGTCAACGGATTTGGTCGT 3', reverse: 5' TTGATTTTGGAGGGATCTCG 3'. Melting temperatures of the PCR products were assayed after completion of the cycle to

rule out primer dimerization and incomplete PCR product formation. mRNA levels of the selected genes S100A2, ODC1, and caspase-1 were normalized to the mRNA of GAPDH, a housekeeping gene, and error reported is standard error of the mean (SEM). The blank sample from the cDNA library construction step was also subjected to qPCR using the primers as listed above, in order to rule out any background signal or contamination. Data reflect 5 biological replicates unless otherwise indicated, and assays were repeated on three different cell passages to confirm the trends.

For RT-qPCR of Gadd45 and p21, slight differences in procedure were followed because these genes were assayed before the development of the above, superior technique. Cells were treated with 50 μM cisplatin for 30 min at 50% confluency at 37°C, 5% CO_2. Then the cells were washed 2x with PBS and grown overnight in the presence of 7.5 μM $[Rh(phi)_2bpy]Cl_3$ or vehicle. The next day, after washing 3x with PBS, the cells were irradiated in PBS for 30 min as above, and then either harvested immediately or allowed to grow at 37°C, 5% CO_2, for 4 hours before harvest. RNA extraction, cDNA library construction, and qPCR techniques were repeated as above, and the error reported is from the combination of 3 separate experiments.

Figure 3.1. DNA sequences used in this experiment. p53 binding areas are in the outlined box, and guanines are in red.

RESULTS

Correlation of the oxidation of p53 with guanine doublets and triplets in the recognition sequence. A master transcription factor, p53 regulates a variety of proteins in multiple cellular pathways by binding to sequences of DNA containing the following pattern: 5'-PuPuPuC(A/T)(A/T)GPyPyPy-3' separated by a spacer of 0-13 base pairs (Pu and Py represent purines and pyrimidines, respectively). This relative promiscuity potentially indicates regulation by base pair identity within the binding site correlating to biological function.

To standardize binding conditions, this study used promoters with sequences with had similar K_d values, and the K_d's were experimentally validated with the superstable mutant of p53 (Table 3.1). As seen in Figures 3.2 and 3.3, the oxidation results, as seen in the gel shift assays, follow the same trend as the synthetically derived sequences.[40] S100A2, with its two sets of guanine triplets on the same strand of the duplex, shows the most total dissociation (reflecting p53 oxidation) of p53 and its DNA. Next is the ODC1 sequence, with its guanine doublets and triplets on either strand of the duplex; the availability of regions of low oxidative potential on both strands of the duplex may have prevented the electron hole from staying in one region long enough to oxidize the bound p53. Finally, caspase-1 shows the least amount of dissociation, close in total amount of oxidation to the anthraquinone-lacking light control.

The natural and synthetic sequence gel shift results are shown side-by-side in Figure 3.4. Except for the 30 min irradiations of Con-2 and S100A2, all of the analogous sequences show total p53-DNA dissociation that is within error of each other. Con-2 and S100A2 both show the maximum amount of p53 oxidation. Con-3 and ODC, both with

sets of guanine doublets or triplets on either side of the DNA duplex, show less oxidation, possibly because of the distribution and equilibrium of electron holes. Finally, caspase-1 and Con-0 both show little oxidation, which is to be expected with their lack of guanine doublets or triplets.

Effects of DNA-mediated p53 oxidation in vivo. *In vivo,* we would expect that for cases where p53 is oxidized and does not bind to the DNA, transcription would be altered. Genes that are activated by p53 binding, such as S100A2 and caspase-1 would show down-regulation, while genes that are repressed by p53 binding, such as ODC1, would be up-regulated. The method to induce DNA CT *in vivo* is seen in Figure 3.5. When HCT116N cells (which contain wild type, active p53) are treated with a rhodium photooxidant, then cisplatin to up-regulate p53 and finally irradiated, the mRNA transcript levels (after 4 hours of recovery) are significant (Figure 3.6). The amount of control mRNA from the housekeeping gene GAPDH (Glyceraldehyde 3-phosphate dehydrogenase, a key enzyme in glycolysis) remains the same, within normal variation, for all samples (data not shown). S100A2 mRNA levels decrease in the rhodium + light samples as compared to the dark control + rhodium samples. However, treatment with rhodium photooxidant, with or without irradiation, leads to a major increase in the amount of S100A2 mRNA present; this may be caused by the influence of other signaling pathways, or it may be a way to protect the cell from yet unknown effects of the photooxidant.

Caspase 1 and ODC1 do not exhibit any change in expression upon the addition of the rhodium-based photooxidant. ODC1 levels are more or less unchanged; a slight increase due to p53 oxidation was expected, but the equilibrium process of electron hole

distribution between the guanine doublets and triplets within the cell, plus the error due to normal fluctuations in transcript levels, could erase any small potential increase. Caspase-1 levels are slightly higher but within error of the untreated cells, which agrees with the hypothesis and *in vitro* data that p53 bound to caspase-1 promoter sites does not undergo DNA-mediated oxidation.

To enhance the impact of the investigation performed by Augustyn et al.,[36] Gadd45 and p21 mRNA levels were also measured under conditions of induced oxidative stress. Like S100A2, Gadd45 levels change upon introduction to the rhodium photooxidant. However, Gadd45 levels decrease by almost an order of magnitude (Figure 3.7A). Also, this effect is seen immediately after irradiation; both the 0 and 4 hr time points after irradiation exhibit this effect. Despite multiple assays, the difference between the Gadd45 rhodium-treated, and rhodium plus irradiated samples is not appreciable with the error for the experiment (usually 10% for most applications of RT-PCR),[47-49] although the low expression levels increase the amount of error in the measurements.[47] Likely the rhodium photooxidant treatment itself, after the cisplatin treatment, for the conditions used has turned off Gadd45 expression so that the effects of additional oxidation cannot be seen. This change in expression could be due to an unknown signaling pathway diverting cellular resources directly away from DNA damage repair, or it could be the side effect of a different response to rhodium treatment. The p21 levels, however, show a slight decrease due to light exposure immediately (0 hr) after irradiation (Figure 3.7B). After 4 hours, however, the p21 expression levels for each of the conditions (untreated, irradiated, rhodium treated, rhodium treated and irradiated) are the same within error.

Table 3.1. K_d's of sequences shown in Figure 3.1. DNA concentration was held constant at 100 nM, and p53 was titrated to reach full binding capacity. After EMSA analysis, the data were fit to the one-site binding equation and analyzed with Origin software.

Sequence[a]	K_d (nM)[b]
Caspase-1	612 ± 79
S100A2	331 ± 22
ODC-1	315 ± 53

a) Oligonucleotides used had the sequences as shown in Figure 3.1.
b) Measurements were made in buffer containing 20 mM Tris, 0.2 mM EDTA, 100 mM KCl, 20% glycerol at pH 8.

Figure 3.2. EMSA example. The sequence Con-0 is 5' AAATCAGCACTAAAACATGTCTAAACATGTCT 3'. The sequence Con-2 is 5' AAATCAGCACTAGGGCATGTCTGGGCATGTCT 3'. DNA and tetramer p53 concentrations are 100 μM each. "LC " refers to strands without anthraquinone, and "AQ" refers to strands with the tethered anthraquinone photooxidatnt. The samples were allowed to bind at 4°C for 30 minutes, and then separated on a 10% acrylamide-TBE gel at 50V for 1 hr before transfer to a Hybond-N membrane at 175 mA for 55 min. Then the membrane was phosphorimaged, scanned, and analyzed with ImageQuant software.

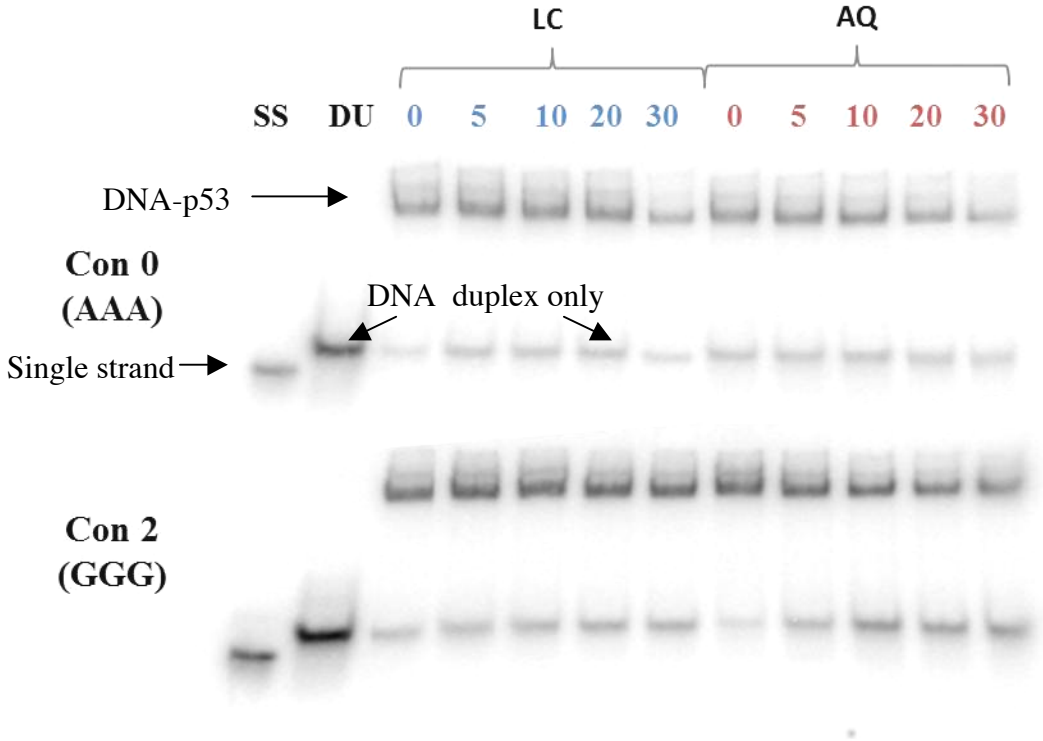


Figure 3.3. Gel shift assay results after ImageQuant analysis for naturally derived sequences. Samples of DNA containing the recognition sites and tetramer p53 in a 1:1 ratio (concentration depended on the K_d) were allowed to bind at 4° for 30 min in 20 mM Tris, 100 mM KCl, 0.2 mM EDTA, 20% glycerol at pH 8. Then they analyzed through EMSA, phosphorimaged, and the ratio of unbound to total DNA, normalized to the 0 time point, allowed for the determination of the p53 dissociation. 3x p53 refers to the triply-stabilized mutant p53. LC refers to samples without the tethered anthraquinone photooxidant.

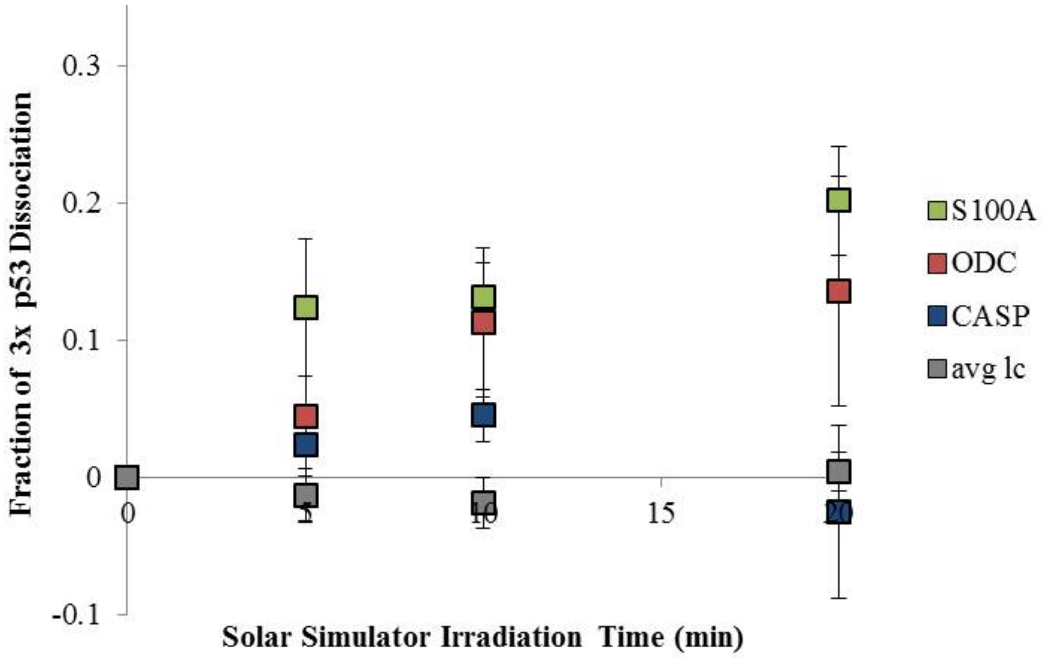


Figure 3.4. Synthetic vs. natural sequence results for the oxidation of p53. Analysis of the different sequences is as described in Figure 3.3. The synthetically-derived sequences are named Con-0 (AAA), Con-1 (AGG), Con-2 (GGG), and Con-3 (GGG/GGG), and were analyzed using DNA and tetramer p53 in a 1:1 ratio (concentration depended on the K_d). LC refers to samples without the covalently linked anthraquinone photooxidant.

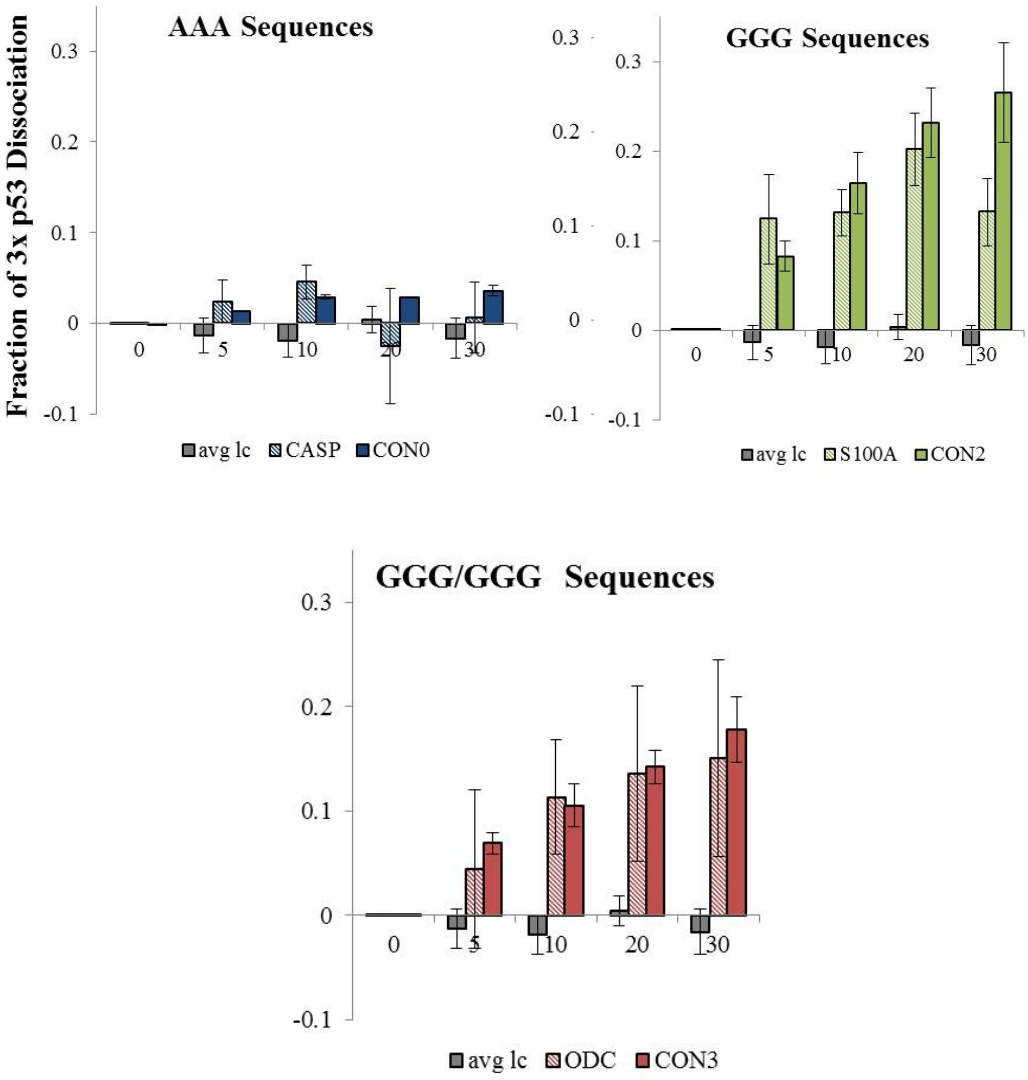


Figure 3.5. Procedure for induction of DNA CT inside of cells for RT-qPCR analysis. The cells used are HCT116N, which contain wild type p53, and irradiation was carried out with a solar simulator. Lysis was accomplished with multiple passages through a needle while suspending in buffer containing guanidine thiocyanate.

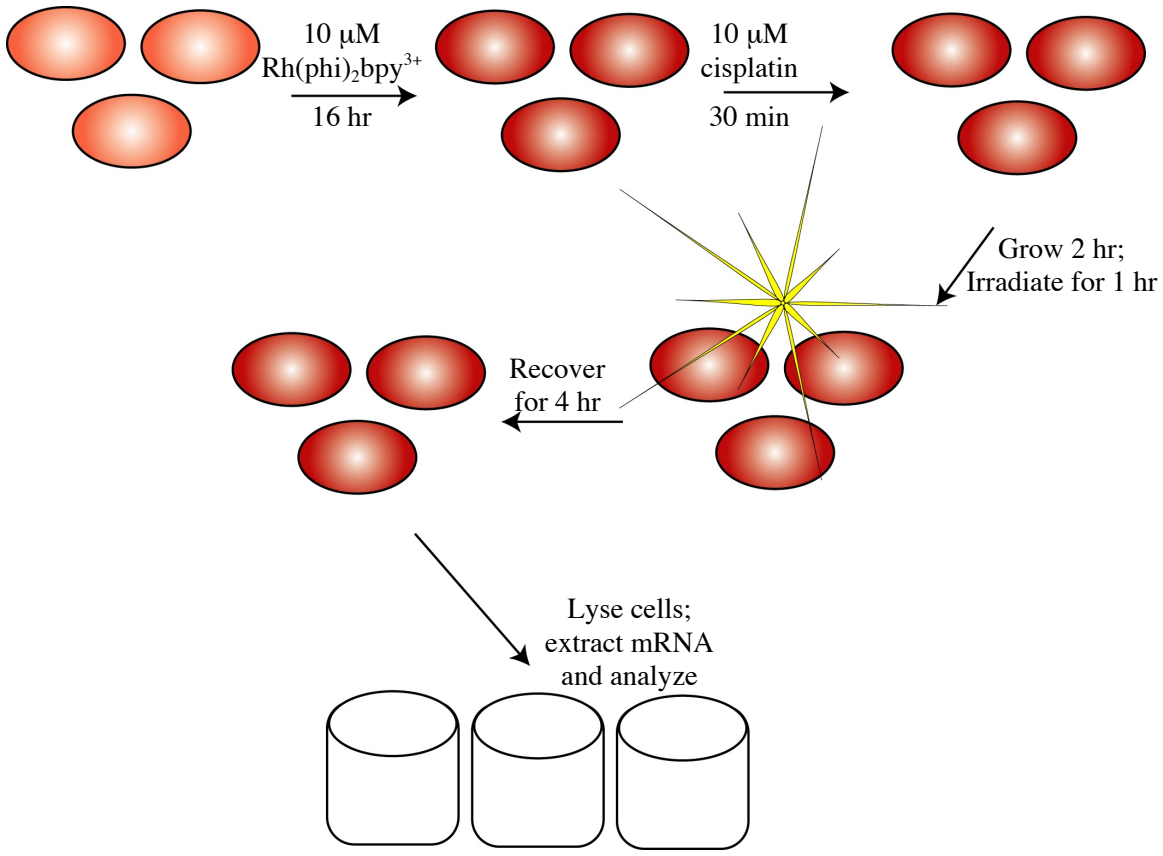


Figure 3.6. RT-qPCR results from conditions leading to p53 oxidation *in vivo*. HCT116N cells were treated with 10 μM $[Rh(phi)_2bpy]Cl_3$ in media for 16 hr, then treated with 10 μM cisplatin before irradiation for 1 hr. mRNA levels were standardized to GAPDH (a metabolic housekeeping gene) using the $\Delta\Delta C_t$ method, and are expressed relative to the untreated control. A) S100A2, B) ODC1, C) caspase-1.

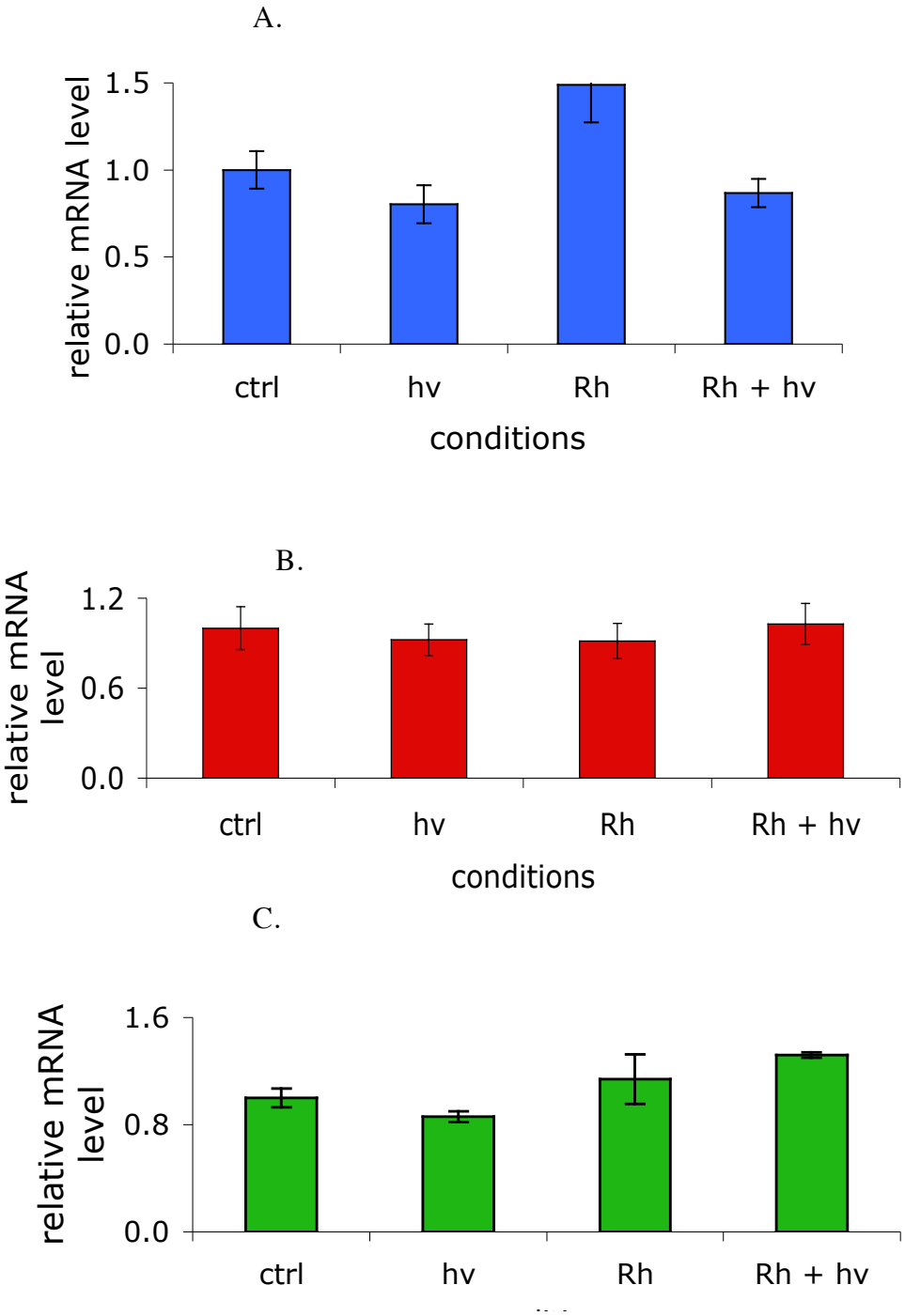
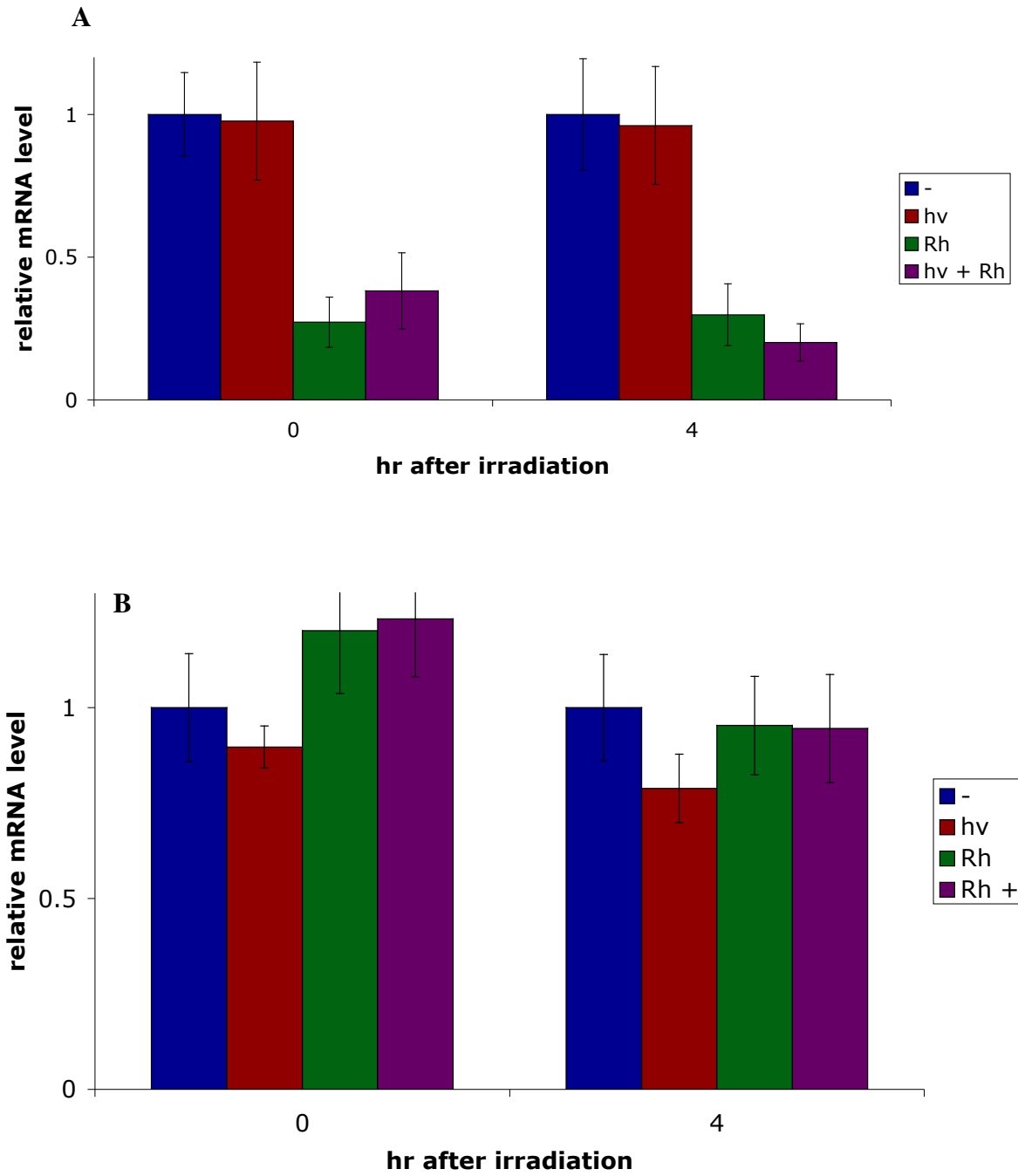


Figure 3.7. RT-qPCR results from other conditions leading to p53 oxidation *in vivo*. Procedure is described in materials and methods. mRNA levels were standardized to GAPDH (a metabolic housekeeping gene) using the $\Delta\Delta C_t$ method, and are expressed relative to the untreated control. A) Gadd45, B) p21/Waf-1.



DISCUSSION

Oxidation of p53 at a distance is dependent on the DNA sequence. The results as described above show a finely regulated, sequence-dependent protein response to DNA CT regulation. As seen in Figure 3.3, p53 oxidation is correlated to increasing sections of low oxidation potential; that is, guanine doublets and triplets on the same strand. In Figure 3.4, the results of the EMSA studies for the naturally occurring p53 recognition sites are directly compared with the results for the synthetically derived p53 recognition sites. This is confirmed in biologically important p53 recognition elements for S100A and ODC1, with caspase-1 as the adenine-rich control, for both *in vitro* and *in vivo* conditions (Figure 3.6). An additional level of control is seen in ODC, where guanine doublets and triplets on both sides of the DNA duplex lead to wide electron hole distribution and less p53 oxidation compared to S100A2, but more p53 oxidation than caspase-1.

This augments the work of Augustyn et al.,[36] who showed a contrasting oxidation response in p53 between recognition elements corresponding to Gadd45 (DNA repair) and p21 (cell cycle arrest), which had identical G/C percentage but differing sequences. The protein p53, bound to Gadd45 sequences, was oxidized at a distance to about the same degree as the Con-2 or S100A2 sequences, while little oxidation was seen for the p21 sequences after the same amount of irradiation time and using the same covalently attached anthraquinone photooxidant. While the results of this study show that the highest extent of DNA-mediated p53 oxidation occurs when guanine triplets occupy each of the purinic regions on one strand of the DNA duplex, the earlier p21 results suggests that one guanine triplet is not enough, or perhaps the p21 sequence has slight structural

differences in either the base stacking. Balagurumoorthy and coworkers report a kink, which would attenuate DNA CT, in the DNA during the p53-p21 recognition sequence binding.[50] Alternatively, the p53 tetramer-DNA contacts are positioned so that the protein is not able to be oxidized at a distance.

The question of how much this effect depends on the structure of the DNA, compared to sequence or oxidation potentials alone, remains to be addressed. Because the structure (B-, A-, or Z-form) of the DNA depends on the sequence,[51] it will be difficult to uncouple sequence and structural effects from each other. To further complicate the issue, Fersht and coworkers have shown that tetramer p53 has as many as four different binding modes, depending on the sequence of the recognition site;[35] these different binding modes may affect the structure of the DNA to either allow or attenuate DNA CT,[13,14] based on the cellular decision making progress. Also, posttranslational modifications are known to affect p53 binding affinity;[52] whether or not acetylation, phosphorylation, and other modifications affect the sensitivity of p53 to DNA remains to be seen. More work that varies the guanine pattern (GAG, GGA, guanine triplet in one of the purinic regions of a half-site but not the other, guanines in the linker region), in addition to studies categorizing p53 structural forms when bound to many different DNA sequences, will be required to understand completely the sequence and structural factors that lead to p53 oxidation at a distance.

Biological consequences. The implication for biological control of these recognition elements is novel as well as extensive in scope. Under conditions of high oxidative stress, when the cell is in jeopardy and oxidative DNA CT is likely to occur, the cell will shift its priorities away from DNA repair (Gadd45) and early tumorgenic

suppression (S100A2) by the DNA-mediated oxidation and thus dissociation of p53 from those promoters. This is well demonstrated by the RT-qPCR results (Figure 3.6), where a down-regulation of S100A2 is seen under conditions of induced DNA CT, but caspase-1 expression is unchanged. Under these stressful oxidative conditions, the genomic integrity will be threatened beyond the point where DNA repair is useful, and the proper cell response will be apoptosis or necrosis so as to avoid transmission of a faulty genome. The cell will not, however, stop the production of cell cycle arrest proteins (p21)[36] and apoptotic proteins (caspase-1) by DNA CT, because these are vital to preventing the cell from progressing to cancer. Future experiments will address more of these p53-controlled genes and illuminate pathways of DNA CT regulation using results from induced cellular DNA CT studies.

While the RT-qPCR experiments were optimized for rhodium photooxidant identity, concentration, irradiation time, and p53 induction, (detailed in my lab notebooks) the interfering pathways within the cell do not always allow for the unambiguous determination of the effect of some stimulus on a particular gene. For example, the effects of the addition of $[Rh(phi)_2bpy]Cl_3$ to the cells are extreme in the case of S100A2, where the rhodium-treated sample that was not irradiated shows a large increase in S100A2 transcript (Figure 3.6). While we speculate that this effect could be traced to the tumor suppression effects of S100A2 in general, the specific reason for this up-regulation is unknown. Another example is the suppression of Gadd45 with the introduction of rhodium photooxidant (Figure 3.7A); this is constant from the 0 hr samples to the 4 hr samples, and could reflect the response of the cell to metallointercalator interaction with the DNA. In Figure 3.7B, p21 levels slightly

decrease after irradiation, but this decrease is not present four hours later. Further studies of the effects of DNA CT will require the complete characterization of the cellular response and interfering pathways to $Rh(phi)_2bpy]Cl_3$ or other photooxidants without the use of light, as well as samples exposed to light but not photooxidants, to ensure the success of those experiments.

This study also presumes that GAPDH is a suitable housekeeping gene for standardization of mRNA reference levels. The most well-known role for GAPDH is in the metabolic pathway, and it has been used as a standard for up to a third of published RT-qPCR studies.[53] However, GAPDH has been linked to other cellular functions, most notably, a death cascade.[54] While GAPDH levels did not vary appreciably between cellular replicates in this experiment (accounting for variation in cell numbers from seeding procedures, and within pipetting error for the cDNA and qPCR steps), the assumption that it will stay constant for every situation is not valid.[55,56] GAPDH was chosen as a reference for this assay due to its stability in the face of these particular conditions for the cells, as shown experimentally (in my lab book); previous experiments (not shown in this chapter; in lab book Mercer IV) have shown reasons why ß-actin cannot be used in these experiments (too much variation during oxidative conditions), as well as a link to variations during hypoxic conditions.[56] Future studies may standardize results to three housekeeping genes, from three different cellular functions, for the most accurate report of experimental findings.[55,57]

In the context of other work on gene expression, the experimental results as shown above are useful in constructing a model of how the cell responds to various types oxidative stress. p53-regulated expression has been widely studied using RT-qPCR and

other techniques that monitor mRNA levels,[58-60] but little work has been done to characterize complete expression levels of p53-regulated genes as oxidative stress increases, most likely due to fluctuations in gene expression from differing cell type and experimental conditions. Therefore, comparing the global change in gene expression levels after induction of DNA CT with the change in gene expression levels after treatment with increasing concentrations of a general oxidant, such as hydrogen peroxide or diamide, would also greatly expand the understanding of how cells react and specific pathways used in the response to oxidative stress.

In cancer cells, S100A2 has been reported to be overexpressed[61] or underexpressed,[42] depending on the type of cancer. S100A2 also is up-regulated in normal human skin that is exposed to UV radiation.[62] ODC-1 is repressed by p53 until apoptosis[58] and shows very stable increased expression patterns upon induction of apoptosis and through cell death. Caspase-1 levels tend to rise after p53 levels in the cell increase, and the accumulation of both proteins is strongly correlated with apoptosis.[63] The data as presented in this chapter show that the conditions under which this experiment was performed did not immediately induce apoptosis; otherwise, ODC1 and caspase-1 levels would be increased 4-fold as compared to the untreated control.[58,63] A goal of future experiments should be the determination of the end result of extreme oxidative DNA CT *in vivo*; the idea is that extreme oxidative CT should eventually lead to apoptosis, but necrosis could be an alternate cellular strategy.

Other possibilities of DNA-mediated CT regulation. As described in this chapter, DNA CT regulation of p53 is sequence specific, and has been demonstrated within cells through the change in expression levels of DNA CT-regulated genes. The validation of

this experiment allows for the exploration of how other redox-active transcription factors are regulated by DNA CT *in vivo*. This study also shows the advantage in investigating transcription factors that directly interact with guanine doublets and triplets in their recognition sites; many important cancer-related transcription factors have multiple guanines, especially in doublets and triplets, in their preferred DNA binding sequences.[64] Further research will show that guanine content in areas of important cellular control may imply regulation by DNA-mediated CT.

CONCLUSIONS

DNA CT is a sequence-dependent process. The protein p53 recognizes dozens of recognition elements within the genome, whose sequence context could provide a way for DNA CT regulation to distinguish between different genes for certain cellular outcomes. This study shows that natural consensus sequences with different base patterns within the purinic region show the following trend in increasing p53 oxidation: AAA< GGA < GGG/GGG< GGG. Furthermore, the oxidation trend is substantiated in naturally derived sequences containing the same sequence motifs, and the regulation of these genes is spotlighted through the changes in mRNA levels as a function of irradiation of the photooxidant-treated cells. The implications of this study involve how the cell can further use DNA CT in controlling p53 function exquisitely and open a new field of *in vivo* protein-DNA CT studies, which will eventually lead to a better understanding of cellular regulation as a whole, both in normal and cancerous cells.

REFERENCES

(1) O'Neill, M. A., and Barton, J. K. *Topics Curr. Chem.* **2004**, *236*, 67-115.

(2) Genereux, J. C., and Barton, J. K. *Chem. Rev.* **2010**, *110*, 1642-1662.

(3) Merino, E. J., Boal, A. K., and Barton, J. K. *Curr. Opin. Chem. Biol.* **2008**, *2008*, 229-237.

(4) Barton, J. K., Olmon, E. D., and Sontz, P. A. *Coord. Chem. Rev.* **2011**, *255*, 619-634.

(5) Sugiyama, H., and Saito, I. *J. Am. Chem. Soc.* **1996**, *118*, 7063-7068.

(6) Saito, I., Takayama, M., Sugiyama, H., and Nakatani, K. *J. Am. Chem. Soc.* **1995**, *117*, 6406-6407.

(7) Meggers, E. Michel-Beyerle, M. E., and Giese, B. *J. Am. Chem. Soc.* **1998**, *120*, 12950-12955.

(8) Nunez, M. E., Hall, D. B., and Barton, J. K. *Chem. & Biol.* **1999**, *6*, 85-97.

(9) Kelley, S. O., and Barton, J. K. *Science* **1999**, *283*, 375-381.

(10) O'Neill, M. A., Becker, H. C., Wan, C., Barton, J. K., Zewail, A. H. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 5896-5900.

(11) O'Neill, M. A., and Barton, J. K. *J. Am. Chem. Soc.* **2002**, *124*, 13053-13066.

(12) Boon, E. M., and Barton, J. K. *Bioconj. Chem.* **2003**, *14*, 1140-1147.

(13) Rajski, S. R., Kumar, S., Roberts, R. J., and Barton, J. K. *J. Am. Chem. Soc.* **1999**, *121*, 5615-5616.

(14) Rajski, S. R., and Barton, J. K. *Biochem.* **2001**, *40*, 5556-5564.

(15) Kelley, S. O., Boon, E. M., and Barton, J. K. *Nucl. Acids Res.* **1999**, *27*, 4830-4837.

(16) Liu, T., and Barton, J. K. *J. Am. Chem. Soc.* **2005**, *127*, 10160-10161.

(17) Williams, T. T., Odom, D. T., and Barton, J. K. *J. Am. Chem. Soc.* **2000**, *117*, 6406-6407.

(18) O'Neill, M. A., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16543-16550.

(19) Liu, C. S., and Schuster, G. B. *J. Am. Chem. Soc.* **2003**, *125*, 6098-6102.

(20) O'Neill, M. A., and Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 13234-13235.

(21) Shao, F., and Barton, J. K. *J. Am. Chem. Soc.* **2007**, *129*, 14733-14738.

(22) Shao, F., Augustyn, K. E., and Barton, J. K. *J. Am. Chem. Soc.* **2005**, *127*, 17445-12452.

(23) Merino, E. J., and Barton, J. K. *Biochem.* **2007**, *46*, 2805-2811.

(24) Merino, E. J., and Barton, J. K. *Biochem.* **2008**, *47*, 1511-1517.

(25) Merino, E. J., Davis, M. L., and Barton, J. K. *Biochem.* **2009**, *48*, 660-666.

(26) Nunez, M. E., Holmquist, G. P., and Barton, J. K. *Biochem.* **2001**, *40*, 12465-12471.

(27) Boal, A. K., Genereux, J. C., Sontz, P. A., Gralnick, J. A., Newman, D. K., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15237-15242.

(28) Gorodetsky, A. A., Dietrich, L. E. P., Lee, P. E., Demple, B., Newman, D. K., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3684-3689.

(29) Lee, P. E., Demple, B., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13164-13168.

(30) Vogelstein, B., Lane, D., and Levine, A.J. *Nature* **2000**, *408*, 307-310.

(31) el-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W., and Vogelstein, B. *Nat. Gen.* **1992**, *1*, 45-49.

(32) Beckerman, R., and Prives, C. *Cold Spring Harb Perspect Biol.* **2010**, *2*, a000935.

(33) Buzek, J., Latonen, L., Kurki, S., Peltonen, K., and Laiho, M. *Nucl. Acids Res.* **2002**, *30*, 2340-2348.

(34) Menendez, D., Inga, A., and Resnick, M. A. *Nat. Rev. Cancer* **2009**, *9*, 724-737.

(35) Melero, R., Rajagopalan, S., Lázaro, M., Joerger, A. C., Brandt, T., Veprintsev, D. B., Lasso, G., Gil, D., Scheres, S. H., Carazo, J. M., Fersht, A. R., Valle, M. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 557-562.

(36) Augustyn, K. E., Merino, E., and Barton, J. K. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 18907-18912.

(37) Perry, M. E. *Cold Spring Harb Perspect Biol.* **2010**, *2*, a000968.

(38) Jung, H. J., Kim, E. H., Mun, J.-Y., Park, S., Smith, M. L., Han, S. S., and Seo, Y. R. *Oncogene* **2007**, *26*, 7517-7525.

(39) Abbas, T., and Dutta, A. *Nat. Rev. Cancer* **2009**, *9*, 400-414.

(40) Schaefer, K. N., Geil, W. M., and Barton, J. K. *In preparation* **2011**.

(41) Wicki, R., Franz, C., Scholl, F. A., Heizmann, C. W., Schäfer, B. W. *Cell Calcium* **1997**, *22*, 243-254.

(42) Cao LY, Y. Y., Li H, Jiang Y, Zhang HF. *World J. Gastroenterol.* **2009**, *15*, 4183-4188.

(43) Pegg, A. E. *J. Biol. Chem.* **2006**, *281*, 14529-14532.

(44) Grutter, M. G. *Curr. Opin. Struct. Biol.* **2000**, *10*, 649-655.

(45) Tidow, H., Melero, R., Mylonas, E., Freund, S. M., Grossmann, J. G., Carazo, J. M., Svergun, D. I., Valle, M., and Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12324-12329.

(46) Sitlani, A., and Barton, J. K. *Biochem.* **1994**, *33*, 12100-12108.

(47) Nolan, T., Hands, R. E., Bustin, S. A. *Nat. Protoc.* **2006**, *1*, 1559-1582.

(48) Pritchard, C. C., Hsu, L., Delrow, J., and Nelson, P. S. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13266-13271.

(49) Huggett, J., Dheda, K., Bustin, S., and Zumla, A. *Genes Immun.* **2005**, *6*, 279-284.

(50) Balagurumoorthy, P., Lindsay, S. M., Harrington, R. E. *Biophys. Chem.* **2002**, *101-102*, 611-623.

(51) Basham, B., Schroth, G. P., and Ho, P. S. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 6464-6468.

(52) Arbely, E., Natan, E., Brandt, T., Allen, M. D., Veprintsev, D. B., Robinson, C. V., Chin, J. W., Joerger, A. C., Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8251-8256.

(53) Suzuki, T., Higgins, P. J., and Crawford, D. R. *Biotechniques* **2000**, *29*, 332-337.

(54) Hara, M. R., Thomas, B., Cascio, M. B., Bae, B. I., Hester, L. D., Dawson, V. L., Dawson, T. M., Sawa, A., Snyder, S. H. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3887-3889.

(55) Radonic, A., Thulke, S., Mackay, I. M., Landt, O., Siegert, W., Nitsche, A. *Biochem. Biophys. Res. Commun.* **2004**, *313*, 856-862.

(56) Zhong, H., and Simons, J. W.. *Biophys. Res. Commun.* **1999**, *259*, 523-526.

(57) Neuvians, T. P., Gashaw, I., Sauer, C. G., von Ostau, C., Kliesch, S., Bergmann, M., Häcker, A., Grobholz, R. *J. Biotechnol.* **2005**, *117*, 163-171.

(58) Mirza, A., Wu, Q., Wang, L., McClanahan, T., Bishop, W. R., Gheyas, F., Ding, W., Hutchins, B., Hockenberry, T., Kirschmeier, P., Greene, J. R., Liu, S. *Oncogene* **2003**, *22*, 3645-3654.

(59) Liu, S., Mirza, A., and Wang, L. *Methods Mol. Biol.* **2001**, *281*, 33-54.

(60) Wei, C. L., Wu, Q., Vega, V. B., Chiu, K. P., Ng, P., Zhang, T., Shahab, A., Yong, H. C., Fu, Y., Weng, Z., Liu, J., Zhao, X. D., Chew, J. L., Lee, Y. L., Kuznetsov, V. A., Sung, W. K., Miller, L. D., Lim, B., Liu, E. T., Yu, Q., Ng, H. H., Ruan, Y. *Cell* **2006**, *124*, 207-219.

(61) El-Rifai, W., Moskaluk, C. A., Abdrabbo, M. K., Harper, J., Yoshida, C., Riggins, G. J., Frierson, H. F. Jr, Powell, S. M. *Cancer Res.* **2002**, *62*, 6823-6826.

(62) Li, Y., Gudjonsson, J. E., Woods, T. L., Zhang, T., Johnston, A., Stoll, S. W., Elder, J. T. *Arch. Dermatol. Res.* **2009**, *301*, 205-217.

(63) Gupta, S., Radha, V., Furukawa, Y., and Swarup, G. *J. Biol. Chem.* **2001**, *276*, 10585-10588.

(64) Wilson, D., Charoensawan, V., Kummerfeld, S. K., and Teichmann, S. A. *Nucl. Acids Res.* **2008**, *36*, D88-D92.

CHAPTER 4

Conclusions

DNA-mediated charge transport is a well-studied phenomenon whose biological impact is only beginning to be investigated. Besides a role in DNA repair, the potential regulation of transcription factors controlling redox-related (and cancer misregulated pathways) through DNA CT are a unique opportunity for the cell to take advantage of redox-active motifs in the protein. Previous investigations of the redox-active transcription factors SoxR and p53 showed the promise of such studies. However, specific explorations into the chemistry of p53 oxidation, oxidation limitation of CT-deficient p53 mutants (and link to cancers), and base-specific regulation *in vivo* had not been completed.

The studies described in this thesis contribute the next step in the exploration of p53 regulation in cells through oxidation at a distance. Specifically, the control of the chemical oxidation of p53 was investigated using site-directed mutagenesis to disrupt the disulfide bond exchange pathway and determine the key residues. Additionally, it was shown that chemical cross-linking occurs as a byproduct of this oxidation. Preliminary mass spectrometry data suggests the formation of a disulfide bond between cysteine 141 and cysteines 124 is the end product of the oxidation. The sequence dependence of the oxidation was demonstrated. Furthermore, the regulation of p53 through DNA-mediated oxidation *in vivo* was shown through the change in relative mRNA levels after induced DNA-mediated oxidation.

Future work on p53 will include further studies regarding the oxidation of p53 *in vivo*. Immunoprecipitation assays from cells that have undergone induced DNA CT, followed by mass spectrometry analysis, will be examined to show if disulfide bonds, and

perhaps other posttranslational modifications, result from regulation by DNA CT; certain forms of mutant p53 (C275S) but not others (C135S and others) might show insensitivity to DNA-mediated oxidation at a distance. Furthermore, plasmid-based luminescent transcriptional assays in cells that have undergone photooxidant-induced DNA CT will confirm patterns of the extent of oxidation based on guanine doublet and triplet distribution, and this data will be useful for the verification of oxidation at a distance through other sequences. More work on wild-type and mutant p53 acting as a repressor, and how that repression could be reversed through oxidation at a distance, will be a priority. Microarray assays investigating the full extent of the change in gene expression through DNA-mediated oxidation of p53 will provide important information, especially in comparison to CT-deficient forms of mutant p53.

While this work encompasses effort from many years of experiments, it is only the first step in the investigation of how DNA CT regulates proteins, and partners with protein CT inside eukaryotic cells. Many transcription factors have redox-active motifs that currently appear to have little use; DNA CT could regulate such proteins, such as NF-κB, c-jun, and c-fos. Future work will use cell permeable photooxidants, irradiation, and common cell assays, including transcriptional assays, RT-PCR, Western blots, and ELISA assays to quantify the response of these transcription factors, and thereby, the cell, to oxidative DNA CT.

The understanding of the cellular regulation through DNA-mediated oxidation is major advance in the understanding of cellular regulation and undoing of this regulation in the case of cancer. To treat cancers effectively, complete knowledge of pathways involving p53 and other transcription factors must be achieved. The information

described in this thesis advances the body of scientific knowledge about the regulation of proteins through DNA-mediated oxidation and the chemical means through which this fascinating process occurs.

APPENDIX A

Protein purification procedure (based on instructions from the Fersht lab):

This construct has a 6 x His tag, Lipoyl domain (aa1-85 of the dihydrolipoyl acetyltransferase domain of Bacillus stearothermophilus) and TEV (Tobacco Etch virus) protease cleavage site cloned between the NdeI and BamHI sites of pET24a. The p53 gene is cloned BamHI/EcoRI. Cleavage of the expressed product with TEV protease results in a protein with an extra Gly/Gly/Ser at the amino terminus.

Transform competent expression cells e.g. E.coli BL21(DE3) (50 μL) with plasmid (1 μL), plate onto TYE plate containing kanamycin (30 μg/mL) and incubate at 37 ^{0}C overnight.

The following morning, make a starter culture by introducing a few colonies from the plate into 2xTY(1.5% tryptone, 1% yeast extract, 0.5% NaCl) medium (10 mL for every liter to be inoculated) containing kanamycin (30μg/ml). Grow at 37^{0}C in a shaking incubator for several hours till the culture reaches the log phase of its growth (OD_{600} = 0.2 – 0.8). Introduce 10 mL of this inoculant into each 2L Erlenmeyer flask containing 1L of 2xTY medium plus kanamycin (30μg/ml) and incubate at 37^{0}C with shaking at ~250 rpm, until an OD_{600} of 0.8 is reached (this is crucial—0.75 is the minimum, and 0.80 is the absolute maximum).

Induce cells by addition of 1 mM IPTG (filter-sterilized). Also add zinc sulphate to a final concentration of 0.1 mM, and express for 16 hours at 22 ^{0}C (or RT), shaking at ~220 rpm. Harvest cells by centrifugation for 20 min at 4500 rpm in Sorvall RC 3B Plus rotor cooled to 4^{0}C. Cell pellets can be snap frozen and stored at -80^{0}C if it is not possible for purification to proceed immediately.

Resuspend pelleted cells in ice cold Cell Cracking Buffer (100 ml per liter (or 10-15 grams of pelleted cells) of culture) and from now on keep samples on ice. Homogenise cell suspension using a homogenizer to remove any remaining lumps, and lyse cells by passing three to five times through a cell cracker, such as the microfluidizer in the Clemons lab. Clear the lysate of cell debris by centrifugation in Oakridge tubes in a SS34 rotor at 12,000 rpm for 20 min at 4^{0}C, and by subsequent filtration of the supernatant through a 0.22 μm filter at 4^{0}C.

Pre-equilibrate a nickel column (e.g. HiTrap Chelating HP column from GE Healthcare) with Nickel Buffer A (5CV). Load crude lysate onto column using peristaltic pump over one hour in cold room. Wash the column with Buffer A until OD_{280} returns to is baseline (5 CV). Run a gradient of 0-100% Buffer B over 5 CV, while collecting fractions (use my Histrap method on the FPLC). Elution is typically at around 100 mM imidazole. Pool fractions of interest, dilute ten-fold with dilution buffer, concentrate to 10-20 mL, and digest overnight at 4^{0}C with TEV protease (1mg of TEV to every 100mg of p53—usually entire vial from Invitrogen).

The following morning, preequilibrate a heparin column (e.g. HiTrap Heparin HP column from GE Healthcare) with Heparin Buffer A (5CV). Load the diluted protein onto the column. Wash the column with Buffer A (5CV). Run a gradient of 0-100% Buffer B over 20 CV (use my heparin method on the FPLC). Collect fractions. Elution is typically around 300 mM salt.

According to the Fersht lab, if protein is still insufficiently pure, concentrate to below 10 mL and load onto a preequilibrated HiLoad 26/60 Superdex 200 prep grade size exclusion column (GE Healthcare). However, in my hands, this only led to visualization

of tetramer and dimer formation, not any purification. Dilute in normal p53 buffer. p53 should be concentrated using Centricon concentrators (YM 30). Aliquot, flush with argon, freeze in liquid nitrogen, and store at -80°C.

Buffers

All buffers should be made ahead of time with ultrapure water, filtered (0.22 μm filter), and cooled to 4°C.

Cell Cracking Buffer: 50 mM KPi, pH 8; 300 mM NaCl; 10mM imidazole,; 15 mM β-mercapto-ethanol (add just before use) and 'Complete' protease inhibitor (1 tablet /50 mL)

Nickel Buffer A: 50 mM KPi, pH8; 300 mM NaCl; 10 mM imidazole; 15 mM β-mercapto-ethanol (add just before use)

Nickel Buffer B: 50 mM KPi, pH8; 300 mM NaCl; 250 mM imidazole; 15 mM β-mercapto-ethanol (add just before use)

Dilution Buffer: 25 mM phosphate, pH 7.5, 10 % glycerol, 2 mM DTT(add DTT immediately before use).

Heparin Buffer A: 25 mM phosphate, pH 7.5 and 10% glycerol.

Heparin Buffer B: 25 mM phosphate, pH 7.5, 1 M NaCl, and 10% glycerol.

Gel Filtration Buffer: 25 mM phosphate, pH 7.2, 300 mM NaCl, 5 mM DTT and 10% glycerol. Should be degassed.

p53 storage buffer: 20 mM Tris, 100 mM KCl, 0.2 mM EDTA, and 20% glycerol, pH 8.

APPENDIX B

Sequencing results for plasmids

1) 3x p53 (M133L/V203A/N268D)

DNA sequence forward:

ATGGAGGAGCCGCAGTCAGATCCTAGCGTCGAGCCCCCTCTGAGTCAGGAAACATTTTCAGAC
CTATGGAAACTACTTCCTGAAAACAACGTTCTGTCCCCCTTGCCGTCCCAAGCAATGGATGAT
TTGATGCTGTCCCCGGACGATATTGAACAATGGTTCACTGAAGACCCAGGTCCAGATGAAGCT
CCCAGAATGCCAGAGGCTGCTCCCCCCGTGGCCCCTGCACCAGCAGCTCCTACACCGGCGGCC
CCTGCACCAGCCCCCTCCTGGCCCCTGTCATCTTCTGTCCCTTCCCAGAAAACCTACCAGGGCA
GCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTGACTTGCACGTACTC
CCCTGCCCTCAACAAGCTGTTTTGCCAACTGGCCAAGACCTGCCCTGTGCAGCTGTGGGTTGA
TTCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTACAAGCAGTCACAGCACAT
GACGGAGGTT

DNA sequence backwards:

CTCCCCTGCCCTCACAAGCTGTTTGCCAACTGGCCAAGACTGCCCTGTGCAGCTGTGGGTGAT
TCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTACAAGCAGTCACAGCACATG
ACGGAGGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGATAGCGATGGTCTGGCCCCT
CCTCAGCATCTTATCCGAGTGGAAGGAAATTTGCGTGCTGAGTATTTGGATGACAGAAACACT
TTTCGACATAGTGTGGTGGTGCCCTATGAGCCGCCTGAGGTTGGCTCTGACTGTACCACCATC
CACTACAACTACATGTGTAACAGTTCCTGCATGGGCGGCATGAACCGGAGGCCCATCCTCACC
ATCATCACACTGGAAGACTCCAGTGGTAATCTACTGGGACGGGACAGCTTTGAGGTGCGTGTT
TGTGCCTGTCCTGGGAGAGACCGGCGCACAGAGGAAGAGAATCTCCGCAAGAAAGGGGAGCC
TCACCACGAGCTGCCCCCAGGGAGCACTAAGCGAGCACTGCCCAACAACACCAGCTCCTCTCC
CCAGCCAAAGAAGAAACCACTGGATGGAGAATATTTCACCCTTCAGATCCGTGGGCGTGAGC
GCTTCGAGATGTTCCGAGAGCTGAATGAGGCCTTGGAACTCAAGGATGCCCAGGCTGGGAAG
GAGCCAGGGGGGAGCAGGGCTCACTCCAGCCACCTGAAGTCCAAAAAGGGTCAGTCTACCTC
CCGCCATAAAAAACTCATGTTCAAGACAGAAGGGCCTGACTCAGACTGA

DNA sequence middle:

CCTTTTGCGCTGTGGGTTGTTCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTAC
AAGCAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGA
TAGCGATGGTCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAGGAAATTTGCGTGCTGAGTA
TTTGGATGACAGAAACACTTTTCGACATAGTGTGGTGGTGCCCTATGAGCCGCCTGAGGTTGG
CTCTGACTGTACCACCATCCACTACAACTACATGTGTAACAGTTCCTGCATGGGCGGCATGAA
CCGGAGGCCCATCCTCACCATCATCACACTGGAAGACTCCAGTGGTAATCTACTGGGACGGGA
CAGCTTTGAGGTGCGTGTTTGTGCCTGTCCTGTGAGAGACCGGCGCACAGAGGAAGAGAATCT
CCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCCCAGGGAGCACTAAGCGAGCACTGCCCA
ACAACACCAGCTCCTCTCCCCAGCCAAAGAAGAAACCACTGGATGGAGAATATTTCACCCTTC
AGATCCGTGGGCGTGAGCGCTTCTAGATGTTCCGAGAGCTGAATGAGGCCTTGGAACTCATGG
AGCCCCCCT

a) Sequencing data for C141S

ATGGAGGAGCCGCAGTCAGATCCTAGCGTCGAGCCCCCTCTGAGTCAGGAAACATTTTCAGAC
CTATGGAAACTACTTCCTGAAAACAACGTTCTGTCCCCCTTGCCGTCCCAAGCAATGGATGAT
TTGATGCTGTCCCCGGACGATATTGAACAATGGTTCACTGAAGACCCAGGTCCAGATGAAGCT
CCCAGAATGCCAGAGGCTGCTCCCCCCGTGGCCCCTGCACCAGCAGCTCCTACACCGGCGGCC
CCTGCACCAGCCCCCTCCTGGCCCCTGTCATCTTCTGTCCCTTCCCAGAAAACCTACCAGGGCA
GCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTGACTTGCACGTACTC
CCCTGCCCTCAACAAGCTGTTTTGCCAACTGGCCAAGACCTCCCCTGTGCAGCTGTGGGTTGA
TTCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATG

Reverse:
CGCGCCATGGCCATCTACAAGCAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCCA
CCATGAGCGCTGCTCAGATAGCGATGGTCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAGG
AAATTTGCGTGCTGAGTATTTGGATGACAGAAACACTTTTCGACATAGTGTGGTGGTGCCCTA
TGAGCCGCCTGAGGTTGGCTCTGACTGTACCACCATCCACTACAACTACATGTGTAACAGTTC
CTGCATGGGCGGCATGAACCGGAGGCCCATCCTCACCATCATCACACTGGAAGACTCCAGTG
GTAATCTACTGGGACGGGACAGCTTTGAGGTGCGTGTTTGTGCCTGTCCTGGGAGAGACCGGC
GCACAGAGGAAGAGAATCTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCCCAGGGAG
CACTAAGCGAGCACTGCCCAACAACACCAGCTCCTCTCCCCAGCCAAAGAAGAAACCACTGG
ATGGAGAATATTTCACCCTTCAGATCCGTGGGCGTGAGCGCTTCGAGATGTTCCGAGAGCTGA
ATGAGGCCTTGGAACTCAAGGATGCCCAGGCTGGGAAGGAGCCAGGGGGGAGCAGGGCTCA
CTCCAGCCACCTGAAGTCCAAAAAGGGTCAGTCTACCTCCCGCCATAAAAAACTCATGTTCAA
GACAGAAGGGCCTGACTCAGACTGA

b) Sequencing data for C275S
Sequence forward:
ATGCACCACCACCACCACCACAGCGGCGCTTTTGAATTTAAGCTGCCGGACATTGGCGAAGGC
ATCCACGAAGGTGAAATTGTCAAATGGTTTGTGAAACCGGGCGATGAAGTGAACGAAGACGA
TGTATTGTGCGAAGTGCAAAATGACAAGGCGGTTGTCGAAATTCCCTCCCCGGTCAAAGGGA
AAGTGCTTGAAATCCTCGTCCCGGAGGGAACAGTGGCAACGGTCGGGCAAACGCTCATCACG
CTCGATGCGCCGGGTTATGAAAACATGACGACCGGCAGCGACACCGGCGAAAACCTGTACTT
CCAGGGTGGATCCATGGAGGAGCCGCAGTCAGATCCTAGCGTCGAGCCCCCTCTGAGTCAGG
AAACATTTTCAGACCTATGGAAACTACTTCCTGAAAACAACGTTCTGTCCCCCTTGCCGTCCCA
AGCAATGGATGATTTGATGCTGTCCCCGGACGATATTGAACAATGGTTCACTGAAGACCCAGG
TCCAGATGAAGCTCCCAGAATGCCAGAGGCTGCTCCCCCCGTGGCCCCTGCACCAGCAGCTCC
TACACCGGCGGCCCCTGCACCAGCCCCCTCCTGGCCCCTGTCATCTTCTGTCCCTTCCCAGAAA
ACCTACCAGGGCAGCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTG
ACTTGCACGTACTCCCCTGCCCTCAACAAGCTGTTTTGCCAACTGGCCAAGACCTGCCCTGTGC
AGCTGTGGGTTGATTCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTACAAGC
AGTCACAGCACATGACGGAGGTTGTG

Sequence reverse + flipped
TACAAGCAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTC
AGATAGCGATGGTCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAGGAAATTTGCGTGCTGA
GTATTTGGATGACAGAAACACTTTTCGACATAGTGTGGTGGTGCCCTATGAGCCGCCTGAGGT
TGGCTCTGACTGTACCACCATCCACTACAACTACATGTGTAACAGTTCCTGCATGGGCGGCAT
GAACCGGAGGCCCATCCTCACCATCATCACACTGGAAGACTCCAGTGGTAATCTACTGGGACG
GGACAGCTTTGAGGTGCGTGTTAGTGCCTGTCCTGGGAGAGACCGGCGCACAGAGGAAGAGA
ATCTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCCCAGGGAGCACTAAGCGAGCACTG
CCCAACAACACCAGCTCCTCTCCCCAGCCAAAGAAGAAACCACTGGATGGAGAATATTTCAC
CCTTCAGATCCGTGGGCGTGAGCGCTTCGAGATGTTCCGAGAGCTGAATGAGGCCTTGGAACT
CAAGGATGCCCAGGCTGGGAAGGAGCCAGGGGGGAGCAGGGCTCACTCCAGCCACCTGAAGT
CCAAAAAGGGTCAGTCTACCTCCCGCCATAAAAAACTCATGTTCAAGACAGAAGGGCCTGAC
TCAGACTGA

c) C277S_4 results from Laragen.
DNA sequence. Forward:
ATGCACCACCACCACCACCACAGCGGCGCTTTTGAATTTAAGCTGCCGGACATTGGCGAAGGC
ATCCACGAAGGTGAAATTGTCAAATGGTTTGTGAAACCGGGCGATGAAGTGAACGAAGACGA
TGTATTGTGCGAAGTGCAAAATGACAAGGCGGTTGTCGAAATTCCCTCCCCGGTCAAAGGGA
AAGTGCTTGAAATCCTCGTCCCGGAGGGAACAGTGGCAACGGTCGGGCAAACGCTCATCACG
CTCGATGCGCCGGGTTATGAAAACATGACGACCGGCAGCGACACCGGCGAAAACCTGTACTT
CCAGGGTGGATCCATGGAGGAGCCGCAGTCAGATCCTAGCGTCGAGCCCCCTCTGAGTCAGG
AAACATTTTCAGACCTATGGAAACTACTTCCTGAAAACAACGTTCTGTCCCCCTTGCCGTCCCA
AGCAATGGATGATTTGATGCTGTCCCCGGACGATATTGAACAATGGTTCACTGAAGACCCAGG

TCCAGATGAAGCTCCCAGAATGCCAGAGGCTGCTCCCCCCGTGGCCCCTGCACCAGCAGCTCC
TACACCGGCGGCCCCTGCACCAGCCCCCTCCTGGCCCCTGTCATCTTCTGTCCCTTCCCAGAAA
ACCTACCAGGGCAGCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTG
ACTTGCACGTACTCCCCTGCCCTCAACAAGCTGTTTTGCCAACTGGCCAAGACCTGCCCTGTGC
AGCTGTGGGTTGATTCCACACCCCCGCCCGGCACCCGCGTCCGCG

Middle:
ACCCCCGCCGGCACCCGCGTCCGCGCCATGGCCATCTACAAGCAGTCACAGCACATGACGGA
GGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGATAGCGATGGTCTGGCCCCTCCTCA
GCATCTTATCCGAGTGGAAGGAAATTTGCGTGCTGAGTATTTGGATGACAGAAACACTTTTCG
ACATAGTGTGGTGGTGCCCTATGAGCCGCCTGAGGTTGGCTCTGACTGTACCACCATCCACTA
CAACTACATGTGTAACAGTTCCTGCATGGGCGGCATGAACCGGAGGCCCATCCTCACCATCAT
CACACTGGAAGACTCCAGTGGTAATCTACTGGGACGGGACAGCTTTGAGGTGCGTGTTTGTGC
CAGTCCTGGGAGAGACCGGCGCACAGAGGAAGAGAATCTCCGCAAGAAAGGGGAGCCTCAC
CACGAGCTGCCCCCAGGGAGCACTAAGCGAGCACTGCCCAACAACACCAGCTCCTCTCCCCA
GCCAAAGAAGAAACCACTGGATGGAGAATATTTCACCCTTCAGATCCGTGGGCGTGAGCGCT
TCGAGATGTTCCGAGAGCTGAATGAGGCCTTGGAACTCAAGGATGCCCAGGCTGGGAAGGAG
CCAGGGGGGAGCAGGGCTCACTCCAGCCACCTGAAGTCCAAAAAGGGTCAGTCTACCTCCCG
CCATAAAAAACTCATGTTCAAGACAGAAGGGCCTGACTCAGAC

d) Y236F_A
ATGCACCACCACCACCACCACAGCGGCGCTTTTGAATTTAAGCTGCCGGACATTGGCGAAGGC
ATCCACGAAGGTGAAATTGTCAAATGGTTTGTGAAACCGGGCGATGAAGTGAACGAAGACGA
TGTATTGTGCGAAGTGCAAAATGACAAGGCGGTTGTCGAAATTCCCTCCCCGGTCAAAGGGA
AAGTGCTTGAAATCCTCGTCCCGGAGGGAACAGTGGCAACGGTCGGGCAAACGCTCATCACG
CTCGATGCGCCGGGTTATGAAAACATGACGACCGGCAGCGACACCGGCGAAAACCTGTACTT
CCAGGGTGGATCCATGGAGGAGCCGCAGTCAGATCCTAGCGTCGAGCCCCCTCTGAGTCAGG
AAACATTTTCAGACCTATGGAAACTACTTCCTGAAAACAACGTTCTGTCCCCCTTGCCGTCCCA
AGCAATGGATGATTTGATGCTGTCCCCGGACGATATTGAACAATGGTTCACTGAAGACCCAGG
TCCAGATGAAGCTCCCAGAATGCCAGAGGCTGCTCCCCCCGTGGCCCCTGCACCAGCAGCTCC
TACACCGGCGGCCCCTGCACCAGCCCCCTCCTGGCCCCTGTCATCTTCTGTCCCTTCCCAGAAA
ACCTACCAGGGCAGCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTG
ACTTGCACGTACTCCCCTGCCCTCAACAAGCTGTTTTGCCAACTGGCCAAGACCTGCCCTGTGC
AGCTGTGGGTTGATTCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTACAAGC
AGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGATAGC
GATGGTCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAGGAAATTTGCGTGCTGAGTATTTG

NCATCTTCTGNNCNTCCCAGAAAACCTNCCAGGGCAGCTACGGTTTCCGTCTGGGCTTCTTGC
ATTCTGGGACAGCCAAGTCTGTGACTTGCACGTACTCCCCTGCCCTCAACAAGCTGTTTTGCCA
ACTGNCCAAGACCTGCCCTGTGCAGCTGTGGGTTGATTCCACACCCCCGCCCGGCACCCGCGT
CCGCGCCATGGCCATCTACAAGCAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCC
ACCATGAGCGCTGCTCAGATAGCGATGGTCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAG
GAAATTTGCGTGCTGAGTATTTGGATGACAGAAACACTTTTCGACATAGTGTGGTGGTGCCCT
ATGAGCCGCCTGAGGTTGGCTCTGACTGTACCACCATCCACTACAACTTCATGTGTAACAGTT
CCTGCATGGGCGGCATGAACCGGAGGCCCATCCTCACCATCATCACACTGGAAGACTCCAGTG
GTAATCTACTGGGACGGGACAGCTTTGAGGTGCGTGTTTGTGCCTGTCCTGGGAGAGACCGGC
GCACAGAGGAAGAGAATCTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCCCAGGGAG
CACTAAGCGAGCACTGCCCAACAACACCAGCTCCTCTCCCCAGCCAAAGAAGAAACCACTGG
ATGGAGAATATTTCACCCTTCAGATCCGTGGGCGTGAGCGCTTCGAGATGTTCCGAGAGCTGA
ATGAGGCCTTGGAACTCAAGGATGCCCAGGCTGGGAAGGAGCCAGGGGGGAGCAGGGCTCA
CTCCAGCCACCTGAAGTCCAAAAAGGGTCAGTCTACCTCCCGCCATAAAAAACTCATGTTCAA
GACAGAAGGGCCTGACTCAGACTGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGA
GCACCACCACCACCACCACTG

e) C135S_A sequencing

ATGCACCACCACCACCACCACAGCGGCGCTTTTGAATTTAAGCTGCCGGACATTGGCGAAGGC
ATCCACGAAGGTGAAATTGTCAAATGGTTTGTGAAACCGGGCGATGAAGTGAACGAAGACGA
TGTATTGTGCGAAGTGCAAAATGACAAGGCGGTTGTCGAAATTCCCTCCCCGGTCAAAGGGA
AAGTGCTTGAAATCCTCGTCCCGGAGGGAACAGTGGCAACGGTCGGGCAAACGCTCATCACG
CTCGATGCGCCGGGTTATGAAAACATGACGACCGGCAGCGACACCGGCGAAAACCTGTACTT
CCAGGGTGGATCCATGGAGGAGCCGCAGTCAGATCCTAGCGTCGAGCCCCCTCTGAGTCAGG
AAACATTTTCAGACCTATGGAAACTACTTCCTGAAAACAACGTTCTGTCCCCCTTGCCGTCCCA
AGCAATGGATGATTTGATGCTGTCCCCGGACGATATTGAACAATGGTTCACTGAAGACCCAGG
TCCAGATGAAGCTCCCAGAATGCCAGAGGCTGCTCCCCCCGTGGCCCCTGCACCAGCAGCTCC
TACACCGGCGGCCCCTGCACCAGCCCCCTCCTGGCCCCTGTCATCTTCTGTCCCTTCCCAGAAA
ACCTACCAGGGCAGCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTG
ACTTGCACGTACTCCCCTGCCCTCAACAAGCTGTTTAGCCAACTGGCCAAGACCTGCCCTGTG
CAGCTGTGGGTTGATTCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGNCATCTACAAG
CAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGATAG
CGATGGTCTGGCCCCTCCTCAGCATCTTATCCGAGTGG

CCAGGGCAGCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTGACTTG
CACGTACTCCCCTGCCCTCAACAAGCTGTTTAGCCAACTGGCCAAGACCTGCCCTGTGCAGCT
GTGGGTTGATTCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTACAAGCAGTC
ACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGATAGCGATG
GTCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAGGAAATTTGCGTGCTGAGTATTTGGATG
ACAGAAACACTTTTCGACATAGTGTGGTGGTGCCCTATGAGCCGCCTGAGGTTGGCTCTGACT
GTACCACCATCCACTACAACTACATGTGTAACAGTTCCTGCATGGGCGGCATGAACCGGAGGC
CCATCCTCACCATCATCACACTGGAAGACTCCAGTGGTAATCTACTGGGACGGGACAGCTTTG
AGGTGCGTGTTTGTGCCTGTCCTGGGAGAGACCGGCGCACAGAGGAAGAGAATCTCCGCAAG
AAAGGGGAGCCTCACCACGAGCTGCCCCCAGGGAGCACTAAGCGAGCACTGCCCAACAACAC
CAGCTCCTCTCCCCAGCCAAAGAAGAAACCACTGGATGGAGAATATTTCACCCTTCAGATCCG
TGGGCGTGAGCGCTTCGAGATGTTCCGAGAGCTGAATGAGGCCTTGGAACTCAAGGATGCCC
AGGCTGGGAAGGAGCCAGGGGGGAGCAGGGCTCACTCCAGCCACCTGAAGTCCAAAAAGGG
TCAGTCTACCTCCCGCCATAAAAAACTCATGTTCAAGACAGAAGGGCCTGACTCAGACTGAAT
TCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACC

f) C124S sequencing

TAGAATAATTTTNNTTAACTTTAAGAAGGAGANATACATATGCACCACCACCACCACCACAGC
GGCGCTTTTGAATTTAAGCTGCCGGACATTGGCGAAGGCATCCACGAAGGTGAAATTGTCAAA
TGGTTTGTGAAACCGGGCGATGAAGTGAACGAAGACGATGTATTGTGCGAAGTGCAAAATGA
CAAGGCGGTTGTCGAAATTCCCTCCCCGGTCAAAGGGAAAGTGCTTGAAATCCTCGTCCCGGA
GGGAACAGTGGCAACGGTCGGGCAAACGCTCATCACGCTCGATGCGCCGGGTTATGAAAACA
TGACGACCGGCAGCGACACCGGCGAAAACCTGTACTTCCAGGGTGGATCCATGGAGGAGCCG
CAGTCAGATCCTAGCGTCGAGCCCCCTCTGAGTCAGGAAACATTTTCAGACCTATGGAAACTA
CTTCCTGAAAACAACGTTCTGTCCCCCTTGCCGTCCCAAGCAATGGATGATTTGATGCTGTCCC
CGGACGATATTGAACAATGGTTCACTGAAGACCCAGGTCCAGATGAAGCTCCCAGAATGCCA
GAGGCTGCTCCCCCCGTGGCCCCTGCACCAGCAGCTCCTACACCGGCGGCCCCTGCACCAGCC
CCCTCCTGGCCCCTGTCATCTTCTGTCCCTTCCCAGAAAACCTACCAGGGCAGCTACGGTTTCC
GTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTGACTTCCACGTACTCCCCTGCCCTCAA
CAAGCTGTTTTGCCAACTGGCCAAGACCTGCCCTGTGCAGCTGTGGGTTGATTCCACACCCCC
GCCCGGCACCCGCGTCCGCGCCATGGNCATCTACAAGCAGTCACAGCACATGACGGAGGTTG
TGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGATAGCGATGGTCTGGCCCCTCCTCAGCATC
TTATCCGA

CAGCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTGACTTCCACGTAC
TCCCCTGCCCTCAACAAGCTGTTTTGCCAACTGNCCAAGACCTGCCCTGTGCAGCTGTGGGTT
GATTCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTACAAGCAGTCACAGCAC
ATGACGGAGGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGATAGCGATGGTCTGGCC

CCTCCTCAGCATCTTATCCGAGTGGAAGGAAATTTGCGTGCTGAGTATTTGGATGACAGAAAC
ACTTTTCGACATAGTGTGGTGGTGCCCTATGAGCCGCCTGAGGTTGGCTCTGACTGTACCACC
ATCCACTACAACTACATGTGTAACAGTTCCTGCATGGGCGGCATGAACCGGAGGCCCATCCTC
ACCATCATCACACTGGAAGACTCCAGTGGTAATCTACTGGGACGGGACAGCTTTGAGGTGCGT
GTTTGTGCCTGTCCTGGGAGAGACCGGCGCACAGAGGAAGAGAATCTCCGCAAGAAAGGGGA
GCCTCACCACGAGCTGCCCCCAGGGAGCACTAAGCGAGCACTGCCCAACAACACCAGCTCCT
CTCCCCAGCCAAAGAAGAAACCACTGGATGGAGAATATTTCACCCTTCAGATCCGTGGGCGTG
AGCGCTTCGAGATGTTCCGAGAGCTGAATGAGGCCTTGGAACTCAAGGATGCCCAGGCTGGG
AAGGAGCCAGGGGGGAGCAGGGCTCACTCCAGCCACCTGAAGTCCAAAAAGGGTCAGTCTAC
CTCCCGCCATAAAAAACTCATGTTCAAGACAGAAGGGCCTGACTCAGACTGAATTCGAGCTCC
GTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACT

g) F134V sequencing
forward:
CTAGAATAATTTTNNTTAACTTTAAGAAGGAGANATACATATGCACCACCACCACCACCACAG
CGGCGCTTTTGAATTTAAGCTGCCGGACATTGGCGAAGGCATCCACGAAGGTGAAATTGTCAA
ATGGTTTGTGAAACCGGGCGATGAAGTGAACGAAGACGATGTATTGTGCGAAGTGCAAAATG
ACAAGGCGGTTGTCGAAATTCCCTCCCCGGTCAAAGGGAAAGTGCTTGAAATCCTCGTCCCGG
AGGGAACAGTGGCAACGGTCGGGCAAACGCTCATCACGCTCGATGCGCCGGGTTATGAAAAC
ATGACGACCGGCAGCGACACCGGCGAAAACCTGTACTTCCAGGGTGGATCCATGGAGGAGCC
GCAGTCAGATCCTAGCGTCGAGCCCCCTCTGAGTCAGGAAACATTTTCAGACCTATGGAAACT
ACTTCCTGAAAACAACGTTCTGTCCCCCTTGCCGTCCCAAGCAATGGATGATTTGATGCTGTCC
CCGGACGATATTGAACAATGGTTCACTGAAGACCCAGGTCCAGATGAAGCTCCCAGAATGCC
AGAGGCTGCTCCCCCCGTGGCCCCTGCACCAGCAGCTCCTACACCGGCGGCCCCTGCACCAGC
CCCCTCCTGGCCCCTGTCATCTTCTGTCCCTTCCCAGAAAACCTACCAGGGCAGCTACGGTTTC
CGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTGACTTGCACGTACTCCCCTGCCCTCA
ACAAGCTGGTTTGCCAACTGGCCAAGACCTGCCCTGTGCAGCTGTGGGTTGATTCCACACCCC
CGCCCGGCACCCGCGTCCGCGCCATGNNCATCTACAAGCAGTCACAGCACATGACGGAGGTT
GTGAGGCGCTGCCCCCACCATGAGCGCTGCTCANATAGCGATGGTCTGGNCCCTCCTCAGCAT
CTTA

Reverse:
AAGACCTGCCCTGTGCAGCTGTGGGTTGATTCCACACCCCCGCCCGGCACCCGCGTCCGCGCC
ATGGCCATCTACAAGCAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCCACCATGA
GCGCTGCTCAGATAGCGATGGTCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAGGAAATTT
GCGTGCTGAGTATTTGGATGACAGAAACACTTTTCGACATAGTGTGGTGGTGCCCTATGAGCC
GCCTGAGGTTGGCTCTGACTGTACCACCATCCACTACAACTACATGTGTAACAGTTCCTGCAT
GGGCGGCATGAACCGGAGGCCCATCCTCACCATCATCACACTGGAAGACTCCAGTGGTAATCT
ACTGGGACGGGACAGCTTTGAGGTGCGTGTTTGTGCCTGTCCTGGGAGAGACCGGCGCACAG
AGGAAGAGAATCTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCCCAGGGAGCACTAA
GCGAGCACTGCCCAACAACACCAGCTCCTCTCCCCAGCCAAAGAAGAAACCACTGGATGGAG
AATATTTCACCCTTCAGATCCGTGGGCGTGAGCGCTTCGAGATGTTCCGAGAGCTGAATGAGG
CCTTGGAACTCAAGGATGCCCAGGCTGGGAAGGAGCCAGGGGGGAGCAGGGCTCACTCCAGC
CACCTGAAGTCCAAAAAGGGTCAGTCTACCTCCCGCCATAAAAAACTCATGTTCAAGACAGA
AGGGCCTGACTCAGACTGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCAC
CACCACCACCACTGAG