I, Morwena Jane V Solivio, hereby submit this original work as part of the requirements for the degree of Doctor of Philosophy in Chemistry.

It is entitled:
Investigation of DNA-Protein Cross-links Generated in the Presence of Biologically Relevant Oxidant Systems

Student’s name: Morwena Jane V Solivio

This work and its defense approved by:

Committee chair: Edward Merino, Ph.D.
Committee member: Joseph Caruso, Ph.D.
Committee member: Pearl Tsang, Ph.D.
Investigation of DNA-Protein Cross-links Generated in the Presence of Biologically Relevant Oxidant Systems

A dissertation submitted to the
Graduate School
of the University of Cincinnati
in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY (Ph.D.)

In the Department of Chemistry
of McMicken College of Arts and Sciences by

Morwena Jane V. Solivio

Bachelor of Science (B.S.), Chemistry
University of the Philippines in the Visayas, Iloilo, Philippines, 2004

Dissertation Advisor: Edward J. Merino, Ph.D.
Abstract

DNA-protein cross-links are detected in cells under oxidative stress, including those exposed to carcinogenic metal ions such as nickel and chromium. These lesions are large, bulky and likely strong blocks to replication and transcription that must be repaired for cell survival. Despite their apparent biological significance, these lesions have received less attention and are less understood than other types of DNA damage. Information on the chemical structure and the mechanism of formation of these lesions is lacking. Its propensity to form in the presence of biologically relevant oxidative conditions also needs to be established.

Complete characterization of DNA-protein cross-links is challenging due to the instability of these adducts and the possible occurrence of multiple oxidative events within the protein. Identification of the DNA bases and amino acids that participate in adduct formation will be valuable in understanding how these cross-links are formed, identifying the protein involved and aiding in the development of a method of detection and characterization of this type of lesion. Determining the propensity of DNA-protein cross-links to form relative to other prevalent and well-understood DNA lesions like 7,8-dihydro-8-oxo-guanine (8OG) will be useful in establishing its consequences as well as its potential as a biomarker for oxidative stress associated diseases, including cancer.

To bypass the complexity of analyzing intact DNA-protein cross-links, the adducts are investigated using model systems, from simple models involving the
nucleoside 2’-deoxyguanosine (dG) and the amino acid lysine, to progressively more complex systems. Investigations involving both small and large molecule models show that guanine oxidation is an important event in cross-link formation. Hydantoin-based adducts with the amino acid lysine, as well as tyrosyl-based cross-links have been identified. It is demonstrated that cross-links occur in significant yields under the oxidative conditions tested and that it is competitive with the other oxidation events occurring on guanine. Proteins having affinity to DNA have been shown to be more pre-disposed to cross-linking compared to non-DNA binding proteins. Product distribution analyses show that in addition to participating in adducts with DNA, highly oxidizable proteins act as sacrificial anodes, protecting DNA from oxidation as introduction of a redox active amino acid, peptide or protein in the oxidation reaction hampers guanine hyper-oxidation, resulting to the build-up of 8OG. This is important in the preservation of genomic integrity as the generation of these hyper-oxidation products increases mutation potential by at least ten-fold.
Acknowledgements

To the Man Above, for this beautiful, wonderful life. The journey has been amazing so far. It is more than I can imagine or wish for. To my family especially to my parents, for giving me so much, and asking for so little. To Dr. Ann Villalobos, for sparking my interest in Chemistry, and for supporting me in every step of my career. To my research advisor, Dr. Edward Merino, for the unwavering guidance through the ups and downs of the last five years. For providing me with all the opportunities to become the best scientist that I can be. To my committee members, Dr. Joseph Caruso and Dr. Pearl Tsang, both of whom I have had the pleasure of working with and learn so much from, for contributing significantly to my scientific and personal growth.

To the Caruso group, especially to my mentor and good friend, Dr. Karolin Kroening, for showing me the ropes when I first entered the department, for teaching me how to do good research and for seeing my potential. To Dr. Larry Sallans, for the patience and for a very productive research collaboration. To the Merino group, Tiffany-Bell Horwath, Dessalegn Nemera, Amy Jones, Tim Joy, Shazna Thowfeik, Anish Vadukoot and Safnaz Farwin, for supporting me in numerous ways, you know what they are.

To Romel Dator, for being my steadfast ally. To Michael Yosief, Papri Bhattacharya, Sanjeewa Rodrigo, Juhlia Kuhlmann, Yun He, Yan Zhang, Siwei Li, Xiaoping Chen, Blake Baird, and Tevhide Ozkaya, for the gift of friendship. To Dr. Svyatoslav Guznov, for always motivating me to do well. To my best friends, Anskie and Katskie, for the loyalty and sustaining friendship. To my sister, Beulah and my brother, Richard, both of whom I know I can always count on.

To Inday Letty, for the unconditional and persistent love and support. To the Missionaries of St Therese for taking me in whenever I need a place to go, and for always making me feel that I belong.

To St. Jerome Academy, for nurturing me both spiritually and academically. For providing me with every opportunity to grow and be prepared for the life outside its walls.

To my professors and friends in the UP in the Visayas, Chemistry Department, especially to my thesis advisor, Ma’am Lani, for doing an excellent job of cultivating my love for Chemistry. And to the University of Cincinnati, Chemistry Department, for giving me the opportunity to be a part of the graduate program and providing me with a good home to grow as a scientist.

Last but not the least, to my Lola Bea, for teaching me what real happiness is all about. This is for you Princess Bea.

Thank you. I hope that one day, I can return even just a fraction of what I was given.
# Table of Contents

## Chapter 1. Introduction

1.1 Reactive oxygen species (ROS), Oxidative Stress and Disease  
1.2 Guanine Oxidation, a Biomarker of Oxidative Stress  
1.3 Generation and Repair of DNA-Protein Cross-links  
1.4 Current Work on Cross-link Structure Determination  
1.5 Research overview  
  1.5 A DNA-Protein Cross-links as Potential Disease Biomarkers  
  1.5 B Goals of this Dissertation  
  1.5 C Research Hypothesis  
  1.5 D Research Strategy  
  1.5 E Overview of Chapter 2  
  1.5 F Overview of Chapter 3  
  1.5 G Overview of Chapter 4  

## Chapter 2. Copper Generated Reactive Oxygen Leads to Formation of Lysine-DNA Adducts

2.1 Introduction  
2.2 Experimental
## Chapter 2

### 2.2.1 Materials

### 2.2.2 Methods

- **2.2.2.1** Copper-Mediated Lysine Addition to dG
- **2.2.2.2** Reaction on DNA Molecules

### 2.3 Results and Discussion

- **2.3.1** Copper-Mediated Lysine Addition to dG
- **2.3.2** Reaction on DNA Molecules

### 2.4 Conclusion

## Chapter 3

### 3.1 Introduction

### 3.2 Experimental

- **3.2.1** Materials
- **3.2.2** Methods
  - **3.2.2.1** DNA-Ribonuclease A Oxidative Cross-linking
  - **3.2.2.2** Primer Extension Experiments
  - **3.2.2.3** LC/MS Product Analysis

---

**Biologically Relevant Oxidants Cause Bound Proteins to Readily Oxidatively Cross-link at Guanine**
### 3.3 Results and Discussion 59

#### 3.3.1 Characterizing DNA-Protein Cross-links Using a Model System 59

#### 3.3.2 Deducing DNA Base Participating in Cross-linking 72

#### 3.3.3 Mass Spectral Analyses of dG-Peptide Cross-links 77

### 3.4 Conclusion 88

### Chapter 4. DNA-Protein Interactions Prevent Hyperoxidized DNA and Allows Buildup of 7,8-Dihydro-8-Oxo-Guanine

#### 4.1 Introduction 91

#### 4.2 Experimental 95

##### 4.2.1 Materials 95

##### 4.2.2 Methods 96

##### 4.2.2.1 Nucleoside Experiments 96

##### 4.2.2.2 Oligonucleotide Experiments 97

#### 4.3 Results and Discussion 98

##### 4.3.1 Nucleoside Studies 98

##### 4.3.2 Oligonucleotide Studies 98
Chapter 5. Future Directions

5.1 Characterization of DNA-Protein Cross-links in More Complex Systems

References

Appendices

A1. Ribonuclease A Binding to DNA
A2. Piperidine Test for DNA Damage
A3. Trypsin Digestion of DNA-Protein Cross-links
A4. Strand Scission as By-Product of DNA-Protein Cross-linking
A5. Sensitivity of DPC Cu(I)-H2O2-Generated DPCs to Loss of Guanine
A8. High Resolution (FT-ICRMS) Mass Spectral Analysis of
N-Ac-AYKTT-dG Adduct 162


A10. Hypercarb Column Chromatograms for dG Hyper-Oxidation Products when Irradiated with Rose Bengal at 560 nm and At 375 nm 170

A11. Peptides Containing Cys and Arg Form Adducts with dsDNA but at a Much Lower Yield Compared to Tyr-Containing Peptides 171

A12. Mass Spectral Analysis of the 3.8-Minute Peak in Figure 4.6 172
# List of Figures

## Chapter 1:

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Endogenous and Exogenous Sources of ROS</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Mitochondria are Important Sources of ROS</td>
<td>6</td>
</tr>
<tr>
<td>1.3</td>
<td>Singlet Oxygen Generation Biological Systems</td>
<td>8</td>
</tr>
<tr>
<td>1.4</td>
<td>Cross-link Structures Currently Identified</td>
<td>18</td>
</tr>
<tr>
<td>1.5</td>
<td>Identification and Characterization of DNA-Protein Cross-links in Cells</td>
<td>20</td>
</tr>
<tr>
<td>1.6</td>
<td>Cross-link Formation can Arise from Guanine Hyper-Oxidation</td>
<td>23</td>
</tr>
</tbody>
</table>

## Chapter 2:

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Reaction Pathway Under Investigation</td>
<td>33</td>
</tr>
<tr>
<td>2.2</td>
<td>Cu-Induced Lysine Addition</td>
<td>39</td>
</tr>
<tr>
<td>2.3</td>
<td>Product Identification Using Mass Spectrometry</td>
<td>41</td>
</tr>
<tr>
<td>2.4</td>
<td>Reaction on a 14-Nucleotide Substrate</td>
<td>44</td>
</tr>
<tr>
<td>2.5</td>
<td>Addition to a 392-Nucleotide DNA</td>
<td>46</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.6</td>
<td>Damage Superimposed on the Sequence of the 392-Nucleotide DNA</td>
<td>43</td>
</tr>
<tr>
<td>3.1</td>
<td>Oxidative DNA-Protein Cross-Link can Form Hydantoin Rings</td>
<td>59</td>
</tr>
<tr>
<td>3.2</td>
<td>Cross-Link Visualization with DNA and Protein Stain</td>
<td>61</td>
</tr>
<tr>
<td>3.3</td>
<td>Dependence of DNA-Protein Cross-linking on Oligonucleotide Structure and Oxidant Type</td>
<td>64</td>
</tr>
<tr>
<td>3.4</td>
<td>Effect of Binding on DNA-Protein Cross-linking</td>
<td>69</td>
</tr>
<tr>
<td>3.5</td>
<td>Guanine Content Correlates with Cross-link Yield</td>
<td>72</td>
</tr>
<tr>
<td>3.6</td>
<td>Guanine is the Major Cross-link Site</td>
<td>74</td>
</tr>
<tr>
<td>3.7</td>
<td>LC-MS Analysis of Rose Bengal Oxidized Adduct Between dG and the Peptide N-Ac-AAKFK</td>
<td>79</td>
</tr>
<tr>
<td>3.8</td>
<td>LC-MS Analysis of Rose Bengal Oxidized Adduct Between dG and the Peptide N-Ac-AYKTT</td>
<td>81</td>
</tr>
<tr>
<td>3.9</td>
<td>LC-MS Analysis of Rose Bengal Oxidized Adduct Between 5’-ACGTC-3’ and the Peptide N-Ac-AYKTT</td>
<td>85</td>
</tr>
</tbody>
</table>
Chapter 4:

Figure 4.1  Proteins Inhibit Singlet Oxygen Hyper-Oxidation of Guanine  91

Figure 4.2  Singlet Oxygen Oxidation of 2’-Deoxyguanosine  99

Figure 4.3  Easily Oxidized Amino Acids and Proteins Enhance 8OdG and Limit Formation of Hyper-Oxidation Products  102

Figure 4.4  Comparison of Oxidation Product Distributions when Peptides or Amino Acids are Added  106

Figure 4.5  Oxidation Product Distribution in the Oligonucleotide Context  111

Figure 4.6  N-Acetyl Cysteine Does Not Inhibit Cross-linking of a Protein Bound to DNA at Physiological Concentrations  114

Chapter 5:

Figure 5.1  DNA-Protein Cross-link Characterization in Complex Systems  121
Chapter 1

Introduction
1.1 Reactive Oxygen Species (ROS), Oxidative Stress and Disease

Reactive oxygen species (ROS) stem from the metabolism of molecular oxygen\(^1\). ROS are radicals, ions or molecules that have an unpaired electron in their outermost electron shell\(^2\). These species are categorized as radical and non-radical. Radical ROS include superoxide (\(O_2^-\)) and hydroxyl radical (\(^\cdot\)OH) while non-radical ROS comprises hydrogen peroxide (\(H_2O_2\)), singlet oxygen (\(^1O_2\)) and ozone (\(O_3\)), among others. Intracellular sources of ROS include the isoforms of NADPH oxidases (NOXs)\(^3\), the mitochondrial respiratory chain, the flavoenzyme ERO1 in the endoplasmic reticulum during the process of protein folding and disulfide bond formation\(^4\), cytochrome p450s, free copper ions and, iron ions that can transform superoxide and/or hydrogen peroxide to the potent oxidant, hydroxyl radical via Fenton reaction\(^5\). Exogenous sources of ROS include environmental pollutants, ionizing radiation, inflammatory cytokines, pathogens and xenobiotics (Figure 1.1)\(^5b\).
Figure 1.1 Endogenous and Exogenous Sources of ROS. Reactive oxygen species are formed by a variety of endogenous and exogenous sources, as by-product of normal cellular functions, metabolism and response to inflammation, or from exposure of organisms to environmental pollutants, ionizing radiation, pathogens and xenobiotics.

Oxidants are generated as a result of normal intracellular metabolism in the mitochondria, peroxisomes and a number of cytosolic enzyme systems. ROS play important roles in host defense, oxidative biosynthetic reactions, as well as function as signaling agents, particularly in higher organisms. Lowering ROS levels below the homeostatic set point may interrupt the physiological role of oxidants in cellular proliferation and host defense. In a similar manner, elevated levels of ROS, due either to increased ROS production or reduced antioxidant defense, may also be injurious and result to ageing and age-related
diseases. Under normal physiological conditions, the cellular level of ROS is constantly maintained to prevent cellular damage. Neutralization of ROS can be facilitated both by enzymatic and non-enzymatic molecules. Enzymatic ROS detoxification can be brought about by antioxidant enzymes, which specifically scavenge different kinds of ROS. For example, superoxide dismutases (SODs) are ubiquitous metalloenzymes that catalyze the dismutation of superoxide anion to oxygen and hydrogen peroxide. SODs use metal ions such as copper (Cu$^{2+}$), zinc (Zn$^{2+}$), manganese (Mn$^{2+}$) or iron (Fe$^{2+}$) as cofactors. These enzymes are located in various cellular compartments and are highly specific in regulating the biological processes linked to each compartment. For instance, superoxide can be dismutated into hydrogen peroxide either in the mitochondrial matrix by manganese superoxide dismutase (MnSOD) or in the cytosol by copper/zinc superoxide dismutase (Cu/ZnSOD) (Figure 1.2). CuZnSOD is widely distributed throughout the cell cytoplasm, nucleus and intermembrane space of the mitochondrial matrix, while MnSOD is found only in the mitochondrial matrix. The third identified isoform of SOD, extracellular superoxide dismutase (EC-SOD) is found in the extracellular matrix of tissues. CuZnSOD deficiency is associated with amyloid protein oligomerization characteristic of Alzheimer's diseases. Deficiency in MnSOD results in early neonatal death in gene knockout mice. Additionally, a variety of cancer cells have reduced levels of MnSOD compared to their normal counterpart. EC-SOD is the major isoform found in vascular vessel walls and its deficiency is linked with cardiovascular diseases. On the other hand, catalases for most eukaryotes are generally
localized in the cytosol and peroxisomes\textsuperscript{15} and facilitate the decomposition of hydrogen peroxide to water and oxygen. Overexpression of catalase in ApoE-/- mice results in the retardation of atherosclerosis\textsuperscript{16}. Catalase does not play a central role in scavenging ROS in the mitochondria since it is mainly localized in the peroxisomes, still its overexpression in the mitochondria has been shown to decrease oxidative damage, inhibit cardiac pathology and extend the lifespan in mice\textsuperscript{17}, suggesting the importance of this enzyme in suppressing cardiovascular inflammation, damage and atherosclerosis. Additional ROS-regulating enzymes include thioredoxin reductases\textsuperscript{18}, peroxiredoxins\textsuperscript{19}, methionine sulfoxide reductases\textsuperscript{20}, and the glutathione system, which include glutathione reductase, glutathione peroxidases (GPX) and glutathione S-transferases (GST)\textsuperscript{21}. In addition, many small molecules that react with ROS non-enzymatically can be recycled and replenished affording them a ROS-buffering capacity. These non-enzymatic antioxidants include flavonoids, vitamins A, C and E\textsuperscript{2}, as well as pyruvate, $\alpha$-ketoglutarate, oxaloacetate\textsuperscript{5b} and more importantly, glutathione which is present in millimolar concentrations within cells\textsuperscript{22}.

A balance between ROS production and removal by antioxidant defense system is important to maintain cellular redox homeostasis. A disturbance of this balance favoring pro-oxidative conditions can result to a condition known as oxidative stress. Elevated levels of ROS and the resulting oxidative stress are capable of imposing biological damage and are implicated in aging and the pathology of various conditions, including cancer, cardiovascular, inflammatory and neurodegenerative diseases\textsuperscript{6, 23}. 
Figure 1.2 Mitochondria are Important Sources of ROS. Electron leakage in mitochondrial electron transport chain is responsible for converting up to 2% of the total oxygen consumed into superoxide anion. The presence of MnSOD in the mitochondrial matrix or ZnCuSOD in the cytosol allows for the conversion of the generated superoxide anion into hydrogen peroxide, which can then be transformed by the enzymes catalase (CAT) or glutathione peroxidase (GPx) into water and oxygen. Hydrogen peroxide is highly diffusible and those that escape enzymatic degradation can travel long distances, reaching other cellular components including the nucleus and reacting with nuclear DNA. Additionally, it can react with cellular transition metals like copper or iron and undergo Fenton reaction to form the highly reactive ROS, hydroxyl radical.

The most important endogenous source of ROS is the mitochondrial electron transport chain (Figure 1.2), which is responsible for converting 1-2% of the total oxygen consumed into superoxide anions. A 130 lb person can produce 160-320 mmol superoxide each day from mitochondrial respiration alone. In some cases, ROS overproduction have been associated to mutations in a mitochondrial gene that encodes a component of the electron transport chain.
Damaged mitochondria in turn produce excess ROS, which attack the mitochondrial DNA. Mitochondrial DNA is considered to be more vulnerable to oxidative damage compared to nuclear DNA due to its proximity to the main source of oxidant formation, as well as due to its limited DNA repair system and the absence of protective histones. Increasing damage to mitochondrial DNA inevitably leads to compromised mitochondrial function and integrity, which results in increased ROS generation and set in motion a vicious cycle of increased DNA damage leading to increased ROS production and vice versa. Oxidative damage to mitochondrial DNA in the heart and brain have been shown to be inversely related to the maximum lifespan of mammals, which indicates that accumulation of mitochondrial DNA damage can result to age-related disorders like cancer and neurodegeneration. The mammalian brain is considered to be particularly susceptible to the damaging effects due to its high oxygen consumption and that persistent exposure to oxidizing conditions can result to progressive neuronal loss associated with aging and neurodegeneration. Some neurological disorders associated with mitochondrial dysfunction include, Alzheimer’s, Parkinson’s and Huntington’s diseases, and amyotrophic lateral sclerosis. Nuclear DNA is not necessarily secure from oxidative damage. Non-radical ROS can penetrate membranes and can impart oxidative damage at locations remote from the oxidant source. For example, hydrogen peroxide generated from the mitochondria, can diffuse long distances, crossing membranes and can travel to the nucleus. In the process, it may undergo Fenton reaction, forming the strong oxidant hydroxyl radical, and impose
damage to nuclear DNA. Copper has been found in isolated genomic DNA and can likely catalyze these reactions in the cell\(^{29}\).

**Figure 1.3 Singlet Oxygen Generation in Biological Systems**  
(A) Light mediated formation of singlet oxygen occurs via initial absorption of solar energy to ground state triplet (molecular) oxygen by either protein-bound or other chromophores, including aromatic compounds and conjugated alkenes.  
(B) Non-
light mediated formation of the oxidant can occur from the activity of several peroxidases and other enzymatic reactions.

Not all ROS is derived from the mitochondrial electron transport chain. Singlet oxygen is another important ROS that does not stem from the mitochondria, but has been shown to occur naturally in several biological systems (Figure 1.3). It is formed from exposure of biological systems to visible or ultraviolet light in the presence of an appropriate sensitizer (Figure 1.3A). Cells contain a range of chromophores that can act as sensitizing agents for singlet oxygen formation. Examples of these are flavins, phenols, quinones, vitamins, DNA\textsuperscript{30} and proteins containing chromophoric amino acids like tryptophan, tyrosine and phenylalanine\textsuperscript{31}. A range of peroxidase (i.e. myeloperoxidase\textsuperscript{32}, lactoperoxidase\textsuperscript{33}, horseradish peroxidase\textsuperscript{34} and chloroperoxidase\textsuperscript{35}) lipoxygenase,\textsuperscript{36} NADH oxidase\textsuperscript{37}-mediated reactions also form this oxidant (Figure 1.3B). Eosonophils\textsuperscript{38} and macrophages\textsuperscript{39} have been shown to produce this oxidant, likely due to the activity of the above enzymes. Singlet oxygen is therefore a likely intermediate in a range of biological systems exposed to oxidative stress.

Oxidative damage to DNA can result to genetic instability, which is a hallmark for cancer\textsuperscript{40}. An increased rate of oxidative DNA damage or in some cases, a deficient DNA repair system is observed in patients with diseases associated with increased cancer risk, like Fanconi Anemia, chronic hepatitis, cystic fibrosis and several autoimmune diseases\textsuperscript{41}. ROS overproduction by tumour cells has been extensively investigated\textsuperscript{2}. Elevated levels of ROS have
been associated with aggressive forms of cancer including breast, lung, liver colon and prostate cancer\textsuperscript{42}. An increased steady state level of DNA base damage, characteristic of hydroxyl radical attack has been reported in inflammatory breast disease, where malignant progression can occur\textsuperscript{42b,43}. Many cancers stem from chronic irritation, infection or inflammation. Recent work have illustrated that inflammation is a critical component of tumor progression\textsuperscript{44}. For example, carcinoma of the hepatic cells is often associated with chronic infection by hepatitis B or C virus\textsuperscript{45}. It has been suggested that the measurement of DNA damage and mutation in human liver associated with persistent chronic hepatitis may be predictive for the onset of liver cancer\textsuperscript{43}.

1.2 Guanine Oxidation, a Biomarker of Oxidative Stress

Exposure of DNA to oxidative conditions can result to lesions including abasic sites, DNA strand breaks, DNA-DNA cross-links, DNA-protein cross-links and a multitude of modifications to the heterocyclic DNA bases\textsuperscript{46}. 2’-deoxyguanosine (dG) is preferentially attacked during oxidative DNA damage due to its low redox potential (1.3 V vs NHE) among the four DNA bases\textsuperscript{47}. The remaining DNA bases are not as susceptible to oxidation, with potentials of 1.4, 1.6 and 1.7 V vs NHE for 2’-deoxyadenosine (dA), 2’-deoxycytidine (dC) and thymidine (dT), respectively\textsuperscript{48}.

Genome wide investigations on the formation of oxidative lesions show that the guanine content of DNA is a strong indicator of overall oxidative DNA damage\textsuperscript{49}. Electron loss created in duplex DNA ultimately ends up at a guanine residue by means
of hole migration. Saito et al demonstrated both experimentally and by ab initio calculations that the 5'-G residues of 5'-GG-3' steps in B-form DNA are the most easily oxidized due to the GG stacks and can act as thermodynamic sinks in hole migration across the DNA π stack \(^5\). The same group documented that the sequence context of guanine affects its propensity to be oxidized. According to their findings, the easiest to oxidize is 5'-GGG with an ionization potential of 6.34 eV \(^5\). The order of the remaining sequences from the lowest to the highest ionization potential is, 5'-CGG (6.44 eV) < 5'-AGG ≈ 5'-GGA (6.50 eV) < 5'-TGG (6.52 eV) < 5'-GGT (6.59 eV) < 5'-GGC ≈ 5'-CGA (6.63 eV) < 5'-AGA (6.73 eV) < 5'-TGA (6.76 eV) < 5'-CGT (6.91 eV) < 5'-AGT (6.93 eV) < 5'-CGC ≈ 5'-TGT (6.96 eV) < 5'-AGC (7.01 eV) < 5'-TGC (7.12 eV) \(^5\). The damage at guanine repeats is caused by a phenomenon called DNA-mediated charge transport where holes or guanine radicals form via electron abstraction on DNA \(^5\). The holes migrate with greater occupancy at the most stable sites (regions with the lowest ionization potential), resulting to more oxidative lesions at these DNA sequences, making them prone to damage; a phenomenon termed funneling \(^5\). It has been suggested that this is a means with which the cell preserves DNA. It has been demonstrated via computational methods that the DNA sequences 5'-GG and 5'-GGG are concentrated on non-coding sequences in DNA \(^5\). It is theorized that upon formation, the hole migrates out of the coding regions to preserve the integrity of these sequences. This hypothesis is supported by experiments showing that charge transport is able to occur over 34 nm \(^5\).

A variety of guanine oxidation products have been identified \(^5\). Most studies are focused on the oxidized lesion, 7,8-dihydro-8-oxo-guanine (8OG), because it is a major
oxidative lesion and can be detected by a range of assays\textsuperscript{57}. The concentration of 8OG in uncompromised cells is estimated 1 modified base per million base pairs\textsuperscript{58}. This level corresponds to several thousands of modified molecules per cell at any given time. This value is shown to increase under cellular stress\textsuperscript{59}. Elevated 8OG levels is observed in a variety of diseases including several types of cancers\textsuperscript{60}, and other age-related diseases such as diabetes\textsuperscript{61}, heart\textsuperscript{62} and neurodegenerative diseases\textsuperscript{63}. A consequence of increased ROS and elevated 8OG level is the initiation, promotion and malignant conversion stages of carcinogenesis. A number of studies have examined the increased level of 8OG in many types of cancer including leukemia\textsuperscript{60b}, lung\textsuperscript{60c} and breast cancer\textsuperscript{60d}. Diabetic patients have been shown to have a higher concentration of 8OG in their urine\textsuperscript{64}, serum\textsuperscript{61}, blood cells\textsuperscript{65}, pancreas\textsuperscript{66}, muscles\textsuperscript{67} and kidneys\textsuperscript{68}. The occurrence of 8OG has been shown to play a critical role in the pathogenesis of neurodegenerative disorders such as, Alzheimer’s, Parkinson’s and Huntington’s Disease\textsuperscript{69}. Thus, 8OG is often used as a cellular biomarker of aging and disease\textsuperscript{70}.

8OG is moderately stable, making favorable interactions with the bases in a DNA duplex\textsuperscript{71}. However, 8OG is mutagenic. This lesion is prone to mispairing with adenine during replication, via Hoogsteen base pairing\textsuperscript{72}, resulting to a miscoding that generates a G:C→T:A transversion mutation\textsuperscript{73}, a frequent somatic mutation in lung, breast, ovarian and colorectal cancers\textsuperscript{74}. Base excision repair (BER) and mismatch repair (MMR) are employed to curtail the potential deleterious effect of this lesion\textsuperscript{72, 75}. The MutT, MutM and MutY enzymes in bacteria\textsuperscript{76}, and the corresponding human enzymes\textsuperscript{77}, MTH1, OGG1 and MUTYH, comprises the 8OG repair pathway. MutT/MTH1 hydrolyzes 8OGTP, removing it from the nucleotide pool so it cannot be incorporated by
DNA polymerases. The BER glycosylase MutM/OGG1 excises 8OG from the 8OG:C base pair to allow restoration of G:C base pair after subsequent processing of other enzymes in the BER pathway. If the action of MutM/OGG1 does not occur before replication takes place, the DNA glycosylase MuTY/MUTYH intercepts the resultant 8OG:A base pair and remove the inappropriate A. The redundant repair pathways reflect the mutagenic potential and physiological significance of this lesion.

An important characteristic of 8OG is that it has a significantly lower redox potential (0.7 V vs NHE) compared to guanine making it highly reactive towards further oxidation. Oxidation of 8OG results in the formation of a variety of stable secondary oxidation products including the hydantoin products spiroiminidihydantoin (Sp) and guanidinohydantoin (Gh), among others. These secondary oxidation products have been shown to exhibit mutation frequencies that are at least an order of magnitude higher than 8OG, causing G:C→T:A and G:C→C:G transversion mutations, and with the two diastereomers of Sp strongly blocking polymerase extension. Their ease for formation in in vitro studies suggests the presence of these lesions in cells. Sp has been detected in Escherichia coli (E. coli) cells lacking the base excision repair glycosylase Nei exposed to chromate, and recent reports indicate the occurrence of both hydantoin lesions in mammalian cells but at levels of about two orders of magnitude lower than 8OG.

8OG is a critical biomarker for aging and disease. Several methods, including RP-HPLC combined with an electrochemical detection system, GC/MS, LC/MS, and ELISA, have been employed to detect this lesion. However, its accurate determination is limited due to the problematic 8OG chemistry. Each method of
detection has yielded divergent concentrations, with some methods giving high deviations, which indicate poor precision. The fact that 8OG is easily generated in DNA, due to the ease of guanine oxidation, explains the challenges encountered in its accurate measurement in cellular DNA. This is likely due to artifactual formation of the lesion during work-up, which will result in an overestimation of 8OG level. Additionally, its facile oxidation can result also to an underestimation of its concentration. Many analytical chemists are currently working to reach a consensus\textsuperscript{57j}.

1.3 Generation and Repair of DNA-Protein Cross-links

Genomic DNA is associated with various structural, regulatory and transcription proteins. The dynamic and reversible association between DNA and protein guarantee the accurate expression and propagation of genetic information. However, various endogenous and exogenous agents including free radicals from normal cellular metabolism, UV light and ionizing radiation, chemotherapeutic agents such as cisplatin and aldehydes, and metal complexes including nickel, chromium and arsenic can induce covalent trapping of proteins on DNA, resulting in DNA-protein cross-links (DPCs)\textsuperscript{87}. Environmental toxins that promote the production of ROS significantly increase the level of DNA-protein cross-links\textsuperscript{88}. Both nuclear and mitochondrial DNA are closely associated with proteins including histones, single stranded DNA binding proteins and transcription factors. The close proximity between DNA and these proteins results to an increased effective concentration of protein compared to the
surrounding water molecules, yielding DNA-protein adducts when an oxidant is present. DPCs are distinctive in comparison to other DNA lesions in that they are extremely bulky, making them likely to impair various DNA processes including replication, transcription and repair, due to steric hindrance. These lesions have been detected at high levels when several repair pathways are compromised as in the case of Fanconi anemia. DNA-protein cross-links represent a relatively abundant form of DNA damage as shown by the data indicating that the background level of DPCs in human white blood cells ranged from 0.5 to 4.5 per $10^7$ bases, which is approximately 20% the occurrence of ubiquitous oxidative stress biomarker, 8OG which has a background level of 100 lesions per $10^7$ bases. Data concerning the repair of these lesions are minimal despite the recognition of their biological significance. One model for the repair of DNA-protein cross-links involve a replication-dependent proteolysis of cross-linked protein to a covalently attached polypeptide, followed by the removal of DNA-polypeptide adduct by nucleotide excision repair (NER).

Early work with formaldehyde (FA)-induced DNA-protein cross-links suggests the involvement of proteolytic degradation in the active removal of DPCs as the incubation of XP-A cells, which are deficient in NER, and normal cells, with lactamycin, a proteasome inhibitor, results to inhibition of DPC repair. Studies utilizing Escherichia coli nucleotide excision repair UvrABC showed the incubation of DNA covalently cross-linked with a 16-kDa protein, T4-pdg, with UvrABC proteins generates a 12-mer DNA fragment cross-linked with the protein as a major product of the reaction. The proper assembly and catalytic function
of the NER complex on DNA containing a covalently attached 16-kDa protein suggests that the NER pathway may be involved in DPC repair and at least some subset of DPCs can be removed by this mechanism without prior proteolytic degradation. The validity of the repair model suggesting that partial proteolysis of cross-linked proteins into smaller oligopeptides account for the initial step in the removal of DPCs, was tested using a polypeptide containing 16 amino acids and the 16-kDa protein, T4-pdg\textsuperscript{96}. Results for this experiment show that the incision of a DNA-protein cross-link by UvrABC, occurred at a rate approximately 3.5-8-fold slower than the rates observed for DNA-peptide cross-links. These results indicate that DNA-peptide cross-links can be efficiently incised by the NER proteins, and DNA-peptide cross-links are preferable substrates for this system relative to DNA-protein cross-links, which suggests the importance of proteolytic degradation as an initial processing step in facilitating NER. Further investigation on the roles of NER and the contribution of homologous recombination in the repair of DPCs in bacterial cells illustrate that NER repair DPCs with cross-linked proteins of sizes less than 14-kDa whereas oversized DPCs were processed exclusively RecBCD-dependent HR \textsuperscript{97}. The same group also investigated the involvement of the proteasome-like function (cytosolic ATP-dependent proteases) in the repair of DPCs in E. coli cells and suggested that they have no participation in the repair, and observed sensitivity of cells deficient in cytosolic ATP-dependent proteases is due lethal filamentation associated with ion mutation, as addition of sulA mutation suppresses the sensitivity. The SulA protein inhibits cell septation by binding to the key cell division protein. Additional
work on DPC repair and tolerance in E. coli illustrates that HR and subsequent replication restart (RR) provide the most effective means of cell survival against DPCs and elimination of DPCs from the genome relies primarily on NER, which provides a second and moderately effective means of cell survival against DNA-protein cross-links\(^9\). Studies in mammalian cell show that the upper size limit of cross-linked proteins (CLPs) amenable to mammalian NER is relatively small (8-10 kDa) compared to bacterial NER (12-14 kDa) so that NER cannot participate in the repair of chromosomal DPCs in mammalian cells\(^9\). Moreover, CLPs are not polyubiquitinated and hence are not subject to proteosomal degradation prior to NER. Induction of DPCs results in the accumulation of DNA double strand breaks in HR-deficient but not HR-proficient cells, suggesting that fork breakage at the DPC site initiates HR and reactivates the stalled fork. These results highlight the differential involvement of NER in the repair of DPCs in bacterial and mammalian cells and demonstrate the versatile and conserved role of HR in the tolerance of DPCs among species.

To determine whether translesion synthesis polymerases contribute to the damage tolerance of DPCs, the sensitivity of TLS polymerase (polB, dinB and umuDC) single, double and triple mutants to FA-induced-DPCs was examined\(^9\). All mutants did not exhibit sensitivity to formaldehyde, which led to the conclusion that bacterial TLS polymerases, Pol II, Pol IV and Pol V cannot bypass DPCs and TLS does not serve as an alternative damage tolerance mechanism for DPCs in cell survival.
1.4 Current Work on Cross-link Structure Determination

Figure 1.4. Cross-link Structures Currently Identified. Representative oxidatively-induced DNA-protein cross-link structures. Eleven structures have been identified. Single isomers are shown.

To date, twelve cross-link structures have been identified, as shown in Figure 1.4. The left panel shows guanine-based cross-links. Guanine is the most oxidation prone base and one of the reactions that it can undergo as a result of this susceptibility to oxidation is DNA-protein cross-linking. The primary oxidation product of guanine, 8OG is prone to undergoing further oxidation
resulting to hydantoin lesions, Sp and Gh following nucleophilic addition of water to oxidized 8OG. Substituting water by the side-chain of a nucleophilic amino acid like lysine in the cell, can result in DNA-protein cross-linking. Lysine can add to the C5 or C8 position, depending on whether the amino acid or the nucleobase undergoes initial oxidation to generate the cross-link\textsuperscript{100}. The left middle panel shows an adduct formed between guanine and tyrosine\textsuperscript{101}. Finally, the lowest left panel shows a cross-link between guanine and arginine that has been observed in a DNA model containing 8OG\textsuperscript{89a}. No structural information is available for this type of cross-link. Structures in the center panel are for adducts derived from aldehyde addition. Apurinic/apyrimidic (AP) sites are a ubiquitous form of DNA damage that are produced in excess of 10,000 lesions per cell per day\textsuperscript{102}, formed by spontaneous hydrolysis of normal or damaged nucleotides, upon oxidation by superoxide or hydroxyl radical, or as base excision repair intermediates\textsuperscript{103}. This lesion can react with the ε-amino group of lysine to form a Schiff base (top center panel)\textsuperscript{104}. A number of adducts are formed from the cleavage of biomolecules from other positions. For example, oxidation of the sugar moiety of DNA (dR) can lead to the scission of the backbone and subsequent liberation of a 5'-aldehyde\textsuperscript{105} that can react with a protein. These adducts are largely due to hydroxyl radical-mediated oxidation. ROS can generate aldehyde-containing metabolites such as acetaldehyde, formaldehyde and hydroxynonenals (oxidized lipids) that can mediate cross-linking between protein and DNA (middle, lower center)\textsuperscript{92, 106}. Hydroxyl radical specific thymidine-based cross-links are shown on the right panel. The first thymidine-
A tyrosine adduct was identified by Dizdaroglu et al in 1985\textsuperscript{107}. Thymidine has also been observed to cross-link to tryptophan, phenylalanine and to histone proteins, likely to the lysine residues\textsuperscript{108}.

1.5 Research Overview

1.5 A DNA-Protein Cross-links as Potential Disease Biomarkers

![Diagram of DNA-Protein Cross-links](image)

**Figure 1.5. Identification and Characterization of DNA-Protein Cross-links in Cells.** The long-term goal of this project is to eventually employ Mass Spectrometry techniques to identify whole protein adducts and to characterize DNA-protein cross-links in cells.
The distinction between diseased cell and normal cell is important in biomarker discovery. Aggressive forms of cancer are associated with elevated levels of ROS, a condition known as oxidative stress. The increased intrinsic reactive oxygen is due in part to oncogenic stimulation, increased metabolic activity and mitochondrial malfunction. Oxidation products of major biomolecules are therefore attractive biomarker candidates.

Exposure of DNA to oxidative conditions can result to a plethora of modifications on the heterocyclic DNA bases. Guanine is preferentially attacked during oxidative DNA damage due to its lower redox potential among the four natural DNA bases. While a range of guanine oxidation products have been identified, most studies are concentrated on the major oxidative lesion 8OG. This lesion is detected by a range of assays in biological samples and is a recognized biomarker of oxidative stress. Further investigation however, revealed that 8OG has an even lower redox potential than guanine, making it susceptible to further oxidation and its accurate measurement a challenge.

The plan for this project is to go beyond 8OG and analyze the formation of hyper-oxidation products like oxidative DNA-protein cross-links that may serve as better substrates for direct analysis. DNA is surrounded by and is tightly associated with a variety of proteins. Oxidation of DNA or its associated proteins results in the formation of electron-deficient species that may be subject to attack by nucleophilic residues leading to covalent cross-links. A DNA-protein cross-
link is a bulky lesion, which makes it easier to detect in cells and is therefore a more appealing biomarker candidate than 8OG.

DNA-protein cross-links have been detected in the lymphocytes welders\textsuperscript{113}, tannery workers\textsuperscript{114} and chrome-platers\textsuperscript{115} following chromium exposure and is suggested to play a role in the metals cytotoxicity and carcinogenicity\textsuperscript{115}. These lesions are currently used as biomarkers for chromium exposure\textsuperscript{116} as well as for exposure to radiation-induced stress\textsuperscript{117}. Additional data showing elevated DNA-protein cross-links under oxidative conditions suggest that these lesions can potentially be employed to fingerprint oxidative stress-related diseases\textsuperscript{118}. Identification of the participating proteins and the site of cross-linking may useful in detecting specific types of diseases like cancer at the onset.

Currently, information on the chemical structure as well as the mechanism of formation of DNA-protein cross-links is lacking. Its propensity to form in biologically relevant oxidative conditions also needs to be established. Complete characterization of DNA-protein cross-links is challenging due to the instability of these adducts\textsuperscript{89a} and the possible occurrence of multiple oxidative events within the protein. Identifying the nucleobases and amino acids involved in cross-linking will be valuable in understanding how these lesions are formed, identifying the proteins involved and aiding in the development of a method of detection and characterization of this type of lesion (Figure 1.5). Determining the propensity of DNA-protein cross-links to form relative to other prevalent and well-understood DNA-lesions like 8OG will be useful in establishing its consequences as well as its potential as disease biomarker.
1.5 B Goals of this Dissertation

This dissertation aims to identify the structures of predominant DNA-protein cross-links that form in the presence of biologically relevant oxidant species, and to determine the propensity of cross-link formation relative to other DNA oxidation products, particularly those involving guanine.

1.5 C Research Hypothesis

Figure 1.6. Cross-link formation can arise from guanine hyper-oxidation. (A) Under oxidative conditions, 2'-deoxyguanosine (dG) is oxidized to 7,8-dihydro-8-oxo-2'-deoxyguanosine (8OdG). Continued oxidative stress in the presence of a nucleophile (H₂O) leads to dSp and dGh lesions. The oxygen atom of the attacking water molecule is shown in red to indicate its position in the
product molecule (B) a nucleophilic amino acid like lysine in a protein can compete with H₂O to generate DNA-protein cross-links.

The hypothesis of this work is that DNA-protein cross-link formation involves guanine hyper-oxidation. Two guanine hyper-oxidation lesions that have recently attracted attention are the hydantoin lesions, spirominodihydantoin (Sp) and guanidinohydantoin (Gh)¹¹⁹. These lesions result from nucleophilic attack of water following oxidation of 8OG¹²⁰. Figure 1.6A shows the proposed guanine oxidation pathway, ultimately forming Sp and Gh in the presence of water as the nucleophile. In close proximity to a nucleophilic amino acid like lysine or alternatively, tyrosine, cysteine or arginine, a DNA-protein cross-link can form (Figure 1.6B).

1.5 C Research Strategy

Because DNA-protein cross-links are inherently complex, the adducts formed are investigated using model systems, from single amino acid-nucleoside models to increasingly more complex systems, including peptides and guanine, and larger systems involving Ribonuclease A and a 27-nucleotide DNA. The 27-nucleotide DNA substrate is designed to readily undergo oxidation and subsequent cross-linking by the addition of a guanine base repeat site which is known to lower redox potential⁵¹. Oxidation via several forms of reactive oxygen species including singlet oxygen (¹O₂), hydrogen peroxide (H₂O₂), hydroxyl radical (OH) and via one-electron oxidation events are tested. Each of these
oxidants can be initiated by specific chemicals. For example, hydroxyl radical is generated by iron or copper upon reaction with hydrogen peroxide both in vitro and in vivo, while singlet oxygen can be produced in vitro by irradiation with Rose Bengal.

Propensity of cross-link formation is determined by investigating cross-linking on dG or the 27-nucleotide DNA model with the peptide N-Ac-ASXF (X= amino acid of interest). Quantification of yields and product distributions are achieved by integration of product peak areas while monitoring absorbance at 240 nm, determined to be the wavelength of maximum absorbance for most guanine oxidation products, followed by normalization of peak areas using each compound’s unique extinction coefficient at this wavelength\textsuperscript{121}. The product distribution in the oligonucleotide context is deduced by calculating the cross-link yield from the product band intensity and comparing it to the total DNA oxidation, which can be determined by treatment with a base known to induce strand scission to DNA damage sites.

1.5C Overview of Chapter 2

Copper is a prevalent cellular oxidant that has been observed either free in solution or associated with genomic DNA via electrostatic interactions. It can react with the highly diffusible ROS, hydrogen peroxide and generate hydroxyl radical. It has been suggested that copper reaction with DNA results to divergent oxidation product. Unbound Cu (I) and Cu (II) mainly react with the
sugar moiety of the molecule, resulting to strand scission, while bound copper species generally result to base oxidation. Chapter 2 focuses on the investigation of the propensity of dG-lysine cross-link to form in a copper-mediated oxidative condition. Lysine is selected as the model amino acid since it possesses to a nucleophilic side-chain thought to participate in adduct formation. Additionally, histone proteins that are closely associated with genomic DNA are abundant in lysine residues making this amino acid a good candidate to represent the protein segment of the lesion. The model used is the lysine derivative, Nδ-acetyl-lysine-methyl ester, that has both the alpha amino and alpha carboxyl group of the molecule protected to prevent cross-reactivity.

This work shows the addition of a lysine derivative to guanine base in a nucleoside, an oligonucleotide and to a large DNA that occurs via oxidation by copper-generated reactive oxygen species. Under these conditions the guanidinohydantoin-lysine adduct is observed in a relative yield of 27% compared to other guanine oxidation events that did not form a cross-link. Mass spectral analysis of the adduct shows that lysine is added to the 5-position during the formation of the adduct indicating initial oxidation on the nucleoside, followed by nucleophilic attack of the ε-amino group of lysine. A fourteen-nucleotide DNA duplex was oxidized under similar conditions. Digestion shows formation of the same guanidinohydantoin-lysine nucleoside. The reaction is then examined in a 392-nucleotide DNA substrate. Oxidation in the presence of the lysine ester shows adducts as stops in a primer extension assay. Adducts predominantly form at the 5′-GGG at position 415. Six of the seven sites that showed reaction
greater than three-fold above background are guanine sites. This study demonstrates that copper can catalyze the formation of DNA-protein adducts and may form in cells with elevated copper and oxidative stress.

1.5 D Overview of Chapter 3

Chapter 3 investigates a model system using ribonuclease A and a 27-nucleotide DNA is used to determine the propensity of oxidative cross-linking to occur in the presence of oxidants. Cross-link formation is examined using four different oxidation systems that generate singlet oxygen, superoxide, and metal-based Fenton reaction, utilizing copper and iron. It is demonstrated that oxidative DNA-protein cross-linking occur in significant yields ranging from 14%, to a maximal yield of 61% in all oxidant systems when equivalent concentrations of DNA and protein are present. Because singlet oxygen generated the highest amount of cross-links, it is selected for subsequent analysis. It is shown that binding is important between DNA and protein in order for adducts to form as only proteins that bind to DNA exhibit the ability to form cross-links. Additionally, increasing the ionic strength of the reaction solution, which interferes with non-specific binding also hinders cross-link formation. Cross-linking is highly dependent on the number of guanines in a DNA sequence. Loss of guanine content in a 27-nucleotide DNA results to nearly complete loss in cross-linking. Primer extension studies shows cross-links predominantly occur at the guanine residues of a 100-nucleotide DNA. The chemical species generated are
examined using two peptides derived from the ribonuclease A sequence, N-Ac-AAAKF and N-Ac-AYKTT, which were allowed to react with dG. The cross-link products are spiroiminodihydantoin, guanidinohydantoin and tyrosyl-based adducts. Mass spectral analysis of the cross-link formed between the peptide N-Ac-AYKTT and a 5-nucleotide DNA, 5-ACGTC confirms cross-linking at guanine. Formations of tyrosine-based adducts may be competitive with the more well-studied lysine-based cross-links. The results in this work suggests that cross-links are likely present in high levels in cells since the propensity to oxidatively cross-link is high and much of the genomic DNA is coated with protein.

1.5 E Overview of Chapter 4

8OG is orders of magnitude more sensitive to oxidation compared to guanine based on the difference in the in redox potential. However, the secondary oxidation products are not detected in normal cells and only at levels significantly lower than 8OG in stressed cells. Chapter 4 probes the influence of highly oxidizable proteins in the oxidative process, either in participating in DNA-protein cross-link formation or in altering the other types of damage to DNA particularly on guanine hyper-oxidation. The presence of redox active amino acids, peptides and proteins results in a reduction in guanine hyper-oxidation, and leads to the buildup of 8OG. This is consistent with what is observed in cells where 8OG is the major oxidative product. To test the effect of antioxidants in DNA-protein cross-linking, the redox active amino acid derivative, N-Acetyl-
Cysteine (NAC) is added to the cross-linking reaction containing Lys and dG in the nucleoside model and ribonuclease A and the 27-nucleotide DNA model to represent the reaction in the DNA-protein model. NAC inhibits adduct formation in the nucleoside studies. In the DNA reaction, inhibition is observed only at NAC concentrations above the physiological level indicating that analogous antioxidants do not interfere with cross-linking. Product distribution studies show that DNA-protein cross-linking is competitive with other DNA oxidation events both in the nucleoside and oligonucleotide contexts, with tyrosine as the most effective cross-linker.
Chapter 2

Copper Generated Reactive Oxygen Leads to Formation of Lysine-DNA Adducts

2.1 Introduction

DNA is a reactive molecule that when modified gives rise to genetic mutations\textsuperscript{77}. It is particularly susceptible to oxidation at the guanine base to give a large abundance of lesions\textsuperscript{46}. Many guanine modifications are repaired via an intricate repair system\textsuperscript{122}. Oxidation of guanine by the cellular transition metals copper and iron is a common route for modification. These metals are known to react with hydrogen peroxide generated produced ubiquitously by several physiological processes including oxidative phosphorylation and the inflammatory cell respiratory burst. Because hydrogen peroxide is freely diffusible, it can potentially interact with nuclear DNA, in addition to mitochondrial DNA\textsuperscript{123}. Reaction via Fenton chemistry yields the potent oxidant, hydroxyl radical\textsuperscript{5a, 124}. Guanine modification is associated with increased cancer rates\textsuperscript{125}, chemotherapy drug resistance\textsuperscript{126}, and error-prone polymerization\textsuperscript{73a}. Lack of repair or oxidative stress greater than the repair capacity will leave the underlying guanine modifications intact within a cell and lead to mutations. It has been shown that guanine modifications can be used as a determinant of oxidative stress in a cell\textsuperscript{127}.

Recent work showed that excessive oxidative stress in mitochondria leads to the formation of proteolysis-dependent DNA adducts\textsuperscript{128}. These adducts are likely DNA–protein cross-links though it is currently experimentally challenging to identify these types of modifications in cells. Burrows and co-workers showed that a cross-link between guanine and lysine (Lys) can occur when oxidized by photo-oxidants like rose bengal or riboflavin and by stoichiometric one-electron
oxidants like Ir(IV)\textsuperscript{100}. If these lesions occur in cells then their formation are likely catalyzed by reactive oxygen species generated from copper or iron. It should be noted that copper-based reactive oxygen species does not generate the only other known form of DNA–protein crosslinks, a thymidine–tyrosine crosslink\textsuperscript{129}. We therefore utilized a copper–hydrogen peroxide system to induce reactive oxygen species and determined if Lys–DNA adducts can form.

Cellular oxidative stress is a complex phenomenon that occurs by many pathways. Oxidative stress occurs via one-electron, photo-initiation, singlet oxygen, hydroxyl radical, or two-electron oxidation pathways\textsuperscript{130}. Copper is a ubiquitous cellular oxidant that reacts with the DNA backbone and the nucleobase\textsuperscript{124}. Copper mediated DNA oxidation has been implicated in aging and oxidative stress related diseases including cancer\textsuperscript{131}. Furthermore, copper has been shown to be present in isolated DNA\textsuperscript{29}. It has been suggested that Cu (I) and Cu (II) species not bound to DNA mainly induce strand breaks, while DNA bound copper species cause base oxidation\textsuperscript{132}. Copper binding to DNA is shown to occur through the N7 of guanine\textsuperscript{133}, which concurs with the observed preference for guanine oxidation\textsuperscript{134}. The lesions induced by copper can lead to DNA strand breaks or mutagenesis\textsuperscript{135}. Identification of guanine–Lys adducts from copper generated reactive oxygen would mean that DNA–protein cross-link formation is competitive with the many other oxidative DNA modifications that occur. Observation of guanine–Lys adducts by common transition metals would indicate that these lesions may be formed in cells under high oxidative stress.
Two possible Lys–guanine adducts can be formed as shown in Figure 2.1. Loss of two electrons by 2'-deoxyguanosine leads to the formation 8-oxo-7,8-dihydro-2'-deoxyguanosine (8OdG). The dG oxidation product, 8OdG (0.7V vs NHE) has a more favorable redox potential than dG (1.3V vs NHE) and once formed, is readily oxidized. The net loss of four electrons has been shown to form many different lesions including addition by proteins and other nucleophiles.
The reaction can form either a spiroiminodihydantoin (Sp) or a guanidino hydantoin (Gh) derivative, as shown in Figure 2.1. Reaction with water has been shown to proceed via a 5-hydroxy intermediate that rearranges to the Sp via C6-acyl migration to form a new carbon bond with C4 as determined via low temperature NMR (Figure 2.1). The Gh lesion is thought to form directly from the intermediate by loss of CO2 at position six. Protein-based nucleophiles have been observed to be covalently bound to DNA under oxidative conditions. The guanine-Lys adduct formation is characterized in a nucleoside, oligonucleotide and a large DNA as model of this complex reaction.

This chapter shows that the oxidation of guanine in the presence of a Lys derivative, N-α-acetyl-lysine methyl ester, forms adducts. Evaluation of the reaction by HPLC and tandem MS illustrates that the reaction produces the 5-N-α-acetyl-lysine methyl ester-Gh. A small fourteen-nucleotide duplex was oxidized in the presence and absence of N-α-acetyl-lysine methyl ester. Upon digestion, a nucleoside product with the same mobility as 5-N-α-acetyl-lysine methyl ester-Gh is observed. Examination of the reaction on a large 392-nucleotide DNA substrate, shows that adduct formation occurs at guanine-rich sites and stops replication in a primer extension experiment.
2.2 Experimental

2.2.1 Materials

All experiments were carried out in at least duplicate. All chemical reagents were purchased from Sigma Aldrich unless otherwise noted. Buffers were purchased from Acros Organics. Data was quantified using 32 Karat HPLC software.

2.2.2 Methods

2.2.2.1 Copper-Mediated Lysine Addition to dG

Copper-based experiments were mixed in a 2 mL glass HPLC tube. 2'-deoxyguanosine–lysine addition experiments were carried out by mixing the nucleoside, N-α-acetyl-lysine methyl ester, sodium ascorbate, hydrogen peroxide with CuCl₂ [800 µL; 250 mM sodium phosphate, 7.5 mM 2'-deoxyguanosine, 30 mM acyLysme, 30mM hydrogen peroxide, 3 mM sodium ascorbate, and 0.5 mM copper, pH8]. Reagent concentrations were changed as indicated. The samples were directly separated by injection of 5 µL into a Beckman Coulter System Gold with an autosampler and a diode array. An Agilent Zorbax SB-C18 [5 µm, 4.6 x 150 mm] column was used for separation of the reaction products by a linear gradient of water, A, and acetonitrile, B, over several minutes [0% B for 1 minute, 14% B over 10 minutes, 90% B over 6 minutes]. Absorbance was monitored at 260 nm. Controls excluded the listed reagent.

The mass spectral identification was performed directly on a HPLC fraction that contained 0.25% acetic acid and 15% acetonitrile. The mass
spectral analysis involved infusion directly into the instrument at 5 µL/min, performed on a Thermo Fisher Scientific LTQ-FT, a hybrid instrument consisting of a linear ion trap and a Fourier transform ion cyclotron resonance mass spectrometer. The entire elutant was introduced into the LTQ-FT using the standard electrospray ionization source for the instrument with a spray voltage of 5 kV and a capillary temperature of 275 °C. Autogain control, AGC, was used and set at 500,000 with a maximum injection time of 1250 ms for FT-ICR full scans. Collision induced dissociation, MS/MS, was executed in the linear trap with an AGC setting of 10,000 and a maximum injection time of 500 ms. FT-ICR full scans were acquired in the positive ion mode at 100,000 resolving power at m/z 400. Mass accuracy errors were below 200 ppb for full scan and below 800 ppb for MS^n. The positive ion MS/MS experiments were carried out simultaneously in the linear trap portion of the instrument using helium as a collision gas, isolation widths of 2 amu, a normalized collision energy of 35, a q value of 0.250 and an excitation time of 30 ms.

2.2.2.2 Reaction on DNA Molecules

Oligonucleotides were purchased from Eurofins MWG Operon and used after a second desalting. The sequence of the 14-nucleotide DNA was 5’-TTCACCTGGCCGTCG. The 14-nucleotide DNA and its complement were mixed in equimolar concentrations (250 µL, 100 µM, and 25 mM sodium phosphate pH 8) and heated to 95 °C for 3 m. The sample was cooled to room temperature over 30 m. Adduct formation was performed in 75 µL volume as before except
the final DNA concentration was 50 µM and the reaction time was 30 minutes. After the reaction, the duplex was purified by centrispin 10 columns (Princeton Separations) according to the manufacturer's protocol. Hydrolysis was accomplished by incubation of phosphodiesterase 1 and alkaline phosphatase (225 µL, 50 mM Tris, 1 mM MgCl₂, 0.1 mM ZnCl₂, 1 mM spermine, 0.1 U phosphodiesterase, 5 U AP, pH 9.3) for 30 m. The mixture was injected directly on the HPLC.

Primer extension experiments were carried out using a 392-nucleotide DNA substrate, prepared from pUC 19 plasmid vector (Takara Bio Inc.) by a 30 cycle PCR amplification (55 °C for 30 s, 75 °C for 45 s, and 95 °C for 30 s) using a 19-nucleotide forward primer, GGCCTCTTCGCTATTACGC, starting at the nucleotide at position 287 of the vector, and a 19-nucleotide reverse primer, AATACGCAAACCGCCTCTC, starting at position 672. The DNA was purified using a Cycle Pure Kit (Omega BioTek) and eluted from the column with 150 µL of 3 mM Tris, pH 8.5 to give a final concentration of 40 µg/µL. The PCR product was treated in a similar manner except the concentration of peroxide was reduced as listed in Figure 2.5. The reaction time was reduced to 10 minutes. Reactions were carried out in 65 µL final volume. After the reaction the DNA was purified for primer extension using centrispin 10 columns. 50 µL of purified mixture was added to the primer extension mix (2x vent buffer, 200 µM dNTP, 0.5 µM M13 primer, 0.02 U/µL vent exo-DNA polymerase). A fluorescently labeled M13 primer (Licor Bio Inc.) is utilized in this primer extension experiment. The samples are subsequently cycled 3 times for primer extension (53 °C for 30
s, 72 °C for 45 s, and 95 °C for 30 s). The extended samples are then precipitated, denatured, and a 6% denaturing PAGE gel was performed. The gel is visualized using the Odyssey Infrared Imaging System (LiCor), with 169 µm resolution and the 700-channel laser source, which has a solid-state laser diode at 680 nm. Sequencing was performed by standard methods using manual sequencing of the 392-nucleotide PCR product and acyclo-terminators except the fluorescently labeled M13 primer was substituted\textsuperscript{139}.

### 2.3 Results and Discussion

#### 2.3.1 Copper-Mediated Lysine Addition to dG

The side chain of Lys is thought to add to dG in a multistep process as illustrated in Figure 2.1. Initially, dG is oxidized by the copper–hydrogen peroxide system to form 8OdG. Further oxidation in the presence of Lys can lead to formation of Sp or Gh-Lys adducts. The \( \alpha \)-amine of lysine was acetylated to avoid cross reactivity at the alpha-nitrogen. The carboxymethyl ester of Lys was used for these experiments. N-\( \alpha \)-acetyl-lysine methyl ester is represented as acyLysme. Reaction products were separated and quantified using HPLC with a C18 column and water as the mobile phase.
Figure 2.2. **Cu-Induced Lysine Addition.** Reactions were examined by HPLC with detection at 260 nm plus (black) and without (grey) Nα-acetyl-lysine methyl ester. Incubation of copper, hydrogen peroxide, ascorbate, and 2'-deoxyguanosine leads to the formation of several oxidation products (overlap of grey and black traces). Addition of Nα-acetyl-lysine methyl ester leads to the formation of a detectable product peak (lysine–dG adduct).

Figure 2.2 shows the HPLC trace for the copper generated reactive oxygen mediated reaction with dG monitored at 260 nm. A large excess of dG was used to limit total modification to less than 3%. The low overall modification lowers over-oxidation of the dG adducts. Reactions contained 0.5 mM CuCl₂, 7.5 mM deoxyguanosine, 30 mM acyLysme, 3mM sodium ascorbate and 30 mM hydrogen peroxide. After 10 hours, the reactions were injected into the HPLC to determine product distributions. Incubation without copper or without hydrogen peroxide leads to no oxidation products. When no acyLysme is present oxidation specific products are observed as a pair of peaks at 7.5 minutes and a peak at 9.5 minutes (Figure 2.2, overlapping peaks). Addition of acyLysme resulted to a new product with an elution time of 15 minutes. The product distribution was 73% oxidation products and 27% Lys-dG adduct. Importantly, the Lys-dependent
product is well-resolved and its formation competitive with other oxidation products since it comprised 27% of all adducts formed. We then set out to elucidate the structure of the Lys-dG adduct.

To determine the structure of the Lys addition product in the reaction we utilized tandem mass spectrometry. The reactions have the same concentrations as the previous section. After 10 hours, the reaction mixture was injected on the HPLC. The initial mass spectrometry of the 15-minute peak was suppressed because of co-elution by the large excess of acyLysme. To obtain mass data on the 15-minute product, the peak was collected, dried, and separated from acyLysme by HPLC using 0.25% acetic acid in the mobile phase. Acetic acid leads to the formation of cationic acyLysme and subsequent separation. The peak was then collected and a mass spectrum was taken in HPLC run buffer. MS\(^2\) was used to verify structure. It should be noted that the errors in all ion masses are below 720 ppb and, thus, small mass differences like that between O and NH\(_2\), a difference of 942 ppm, are easily distinguishable.
Figure 2.3. Product Identification Using Mass Spectrometry. (A) Tandem-mass spectrometry was used to identify the product of lysine addition. Experimental masses are in black while theoretical mass values are in grey and bold. Errors in mass are listed in the results section. MS of the addition product after the removal of a large excess of co-eluting acyLysme (inset spectrum). No other products are observed over a mass range of 250 to 650 mass units. Fragmentation of the 458 ion is consistent with the formation of a guanidino hydantoin–lysine adduct. (B) Two guanidino hydantoin–lysine isomers are possible that differ in the location of the lysine addition (5-acyLysme–Gh and 8-acyLysme–Gh). Superimposing the MS² fragmentation on each structure strongly supports that the product of lysine addition is 5-acyLysme–Gh since the 282.1562 ion can only be formed in that isomer (grey arrow). Other single fragments (outlined arrow) and double fragments (light grey outlined arrow) from the highly abundant 399 ion are observed (light grey outlined arrow).
We identified the products of acyLysme addition (Figure 2.3). Analysis of the 15-minute product led to an observed mass of 458.2358 with a calculated error of 146 ppb (Figure 2.3A). The elemental composition of this ion mass is $\text{C}_{18}\text{H}_{32}\text{N}_7\text{O}_7^+$, which is identical to a structure of Gh–acyLysme. Importantly, we do not observe any ions with a mass of 484.2152, $\text{C}_{19}\text{H}_{30}\text{N}_7\text{O}_8^+$, which is equivalent to the spiroiminodihydantoin–lysine adduct (see inset Figure 2.3). Thus, the major product of copper–hydrogen peroxide oxidation at pH 8 is Gh–acyLysme.

Further structural characterization was accomplished by MS$^2$. The position of the acyLysme could be elucidated by the MS$^2$ spectrum if appropriate fragmentation would occur as shown in Figure 2.3B. When the 458 molecular ion was fragmented five products are formed. The major product observed is loss of $\text{CH}_6\text{N}_3$, which is ejection of guanidine ($\text{C}_{17}\text{H}_{27}\text{N}_4\text{O}_7^+$, 260 ppb error, 399 ion). Loss of guanidine is consistent with the pKa of guanidine, which is 13.6. Also observed is loss of $\text{C}_9\text{H}_{18}\text{N}_2\text{O}_3$, which is ejection of acyLysme ($\text{C}_9\text{H}_{14}\text{N}_5\text{O}_4^+$, 349 ppb error, 256 ion). This fragment confirms that the adduct possesses an intact acyLysme. Loss of guanidine is so prevalent that a secondary fragmentation is observed. The two secondary fragmentation products are loss of $\text{C}_2\text{H}_5\text{N}_3\text{O}$, ejection of both guanidine and carbon monoxide ($\text{C}_{16}\text{H}_{29}\text{N}_4\text{O}_6^+$, 370 ppb error, 373 ion) and loss of $\text{C}_6\text{H}_{13}\text{N}_3\text{O}_3$, ejection of both guanidine and deoxyribose ($\text{C}_{12}\text{H}_{19}\text{N}_4\text{O}_4^+$, 770 ppb error, 283 ion). The 283 ion confirms that Lys is attached to the guanine base. The remaining fragmentation product at 282 elucidates the position of acyLysme on the ring. Importantly, we observe loss of $\text{C}_6\text{H}_{12}\text{N}_2\text{O}_4$, 

which is ejection of a urea-deoxyribose equivalent \( (C_{12}H_{20}N_5O_3^+, \text{ 610 ppb error, \text{ 282 ion}}) \) in MS\(^2\). The 282 ion fragment cannot be produced through simple fragmentation of 8-acyLysme-Gh, one of the two possible structures remaining (see structures in Figure 2.3B). Urea loss cannot occur with 8-acyLysme-Gh. For the 282 ion to be produced in 8-acyLysme-Gh a triple fragmentation must occur. The unlikely triple fragment would take place by ejection of 2'-deoxyguanosine, then ejection of the carbon monoxide at position 5, and then ejection of a single NH\(_2\) from the guanidine motif. Given that the largest ion observed is a fragmentation product that loses guanidine (399 ion), any multiple fragmentation products should involve the 399 ion. In fact, all the double fragmentation products involve loss of guanidine. Furthermore, no other triple fragmentation products are observed. In contrast, 5-acyLysme–Gh can produce the fragment in a single step by cleavage across the hydantoin ring to lose urea-deoxyribose. Thus, MS\(^2\) strongly supports that the 15-minute peak is of 5-acyLysme–Gh.

**2.3.2 Reactions on DNA Molecules**

We then set out to determine reactivity on DNA oligonucleotides. Lysine addition reactions were quantified on a 14-nucleotide DNA with the sequence 5’-TTCACTGGCCGTCG. The 14-nucleotide DNA was hybridized to its complement at a final concentration of 100 \( \mu \text{M} \). Reaction time was shortened to 30 minutes while other reaction conditions remained the same. After the reaction the duplex was purified by size exclusion chromatography and digested with a mixture of phosphodiesterase and alkaline phosphatase. The resulting nucleosides were
Figure 2.4. Reaction on a 14-Nucleotide Substrate. A 14-nucleotide sequence and its complement were oxidized by the copper oxidation system in the presence (black) and absence (grey) of Nα-acetyl-lysine methyl ester. The oligonucleotide was digested to corresponding nucleosides and analyzed via HPLC. A product is observed at 15 minutes, which has the same elution profile as 5-acyLysme–Gh.

Analysis of the digested nucleosides shows modification of the 14-nucleotide duplex. After digestion for 30 minutes, each of the four nucleosides were observed (Figure 2.4). When the 14-nucleotide duplex is oxidized by the copper–hydrogen peroxide system in the absence of acyLysme a small but noticeable peak is observed before the dG band. This small peak is similar to that observed in Figure 2.2. Addition of acyLysme resulted to the formation of a new product with an elution time of 15 minutes. The peak is well resolved from the resulting nucleosides, making identification facile. The elution time is the same as the already identified 5-acyLysme–Gh. This data shows that 5-acyLysme–Gh is also the product on oligonucleotides. The reaction was analyzed as before without any further purification.
investigated on a large DNA substrate to understand how sequence influences reactivity (Figure 2.5). A 392-nucleotide portion of the Puc19 plasmid was amplified by PCR to generate a product that overlapped nucleotides 287–678 (sequence in Figure 2.6). The PCR product includes the multiple cloning site at positions 396–454 to which known primers exist. A fluorescently labeled primer that overlaps at positions 370–389 was used for experiments. Manual sequencing of the 392-nucleotide sequence was accomplished using the fluorescent primer and terminators. The sequence listed in Figure 2.5 corresponds to the primer strand. Thus, the C-nucleotide lane interrogates G-nucleotides in the 392-nucleotide template. Lys and other small molecules were removed by size-exclusion prior to extension. DNA damage was monitored in the 130 nucleotides that range from position 396 to 526.
Figure 2.5. Addition to a 392-Nucleotide DNA. The left four lanes of the gel show manual sequencing of 392-nucleotide DNA substrate. The C-nucleotide sequencing lane corresponds to the G-nucleotide of the template. The next five lanes are the primer extension profiles when different components (shown on top) are incubated with the 392-nucleotide product. When hydrogen peroxide is excluded a large band at the top of the gel (position 678) is observed indicating little DNA damage. Addition of lysine methyl ester with no oxidant shows no
change. Incubation of the DNA substrate with hydrogen peroxide results in a decrease in intensity of the band at position nucleotide 678 but limited stops to replication. Addition of lysine and the copper induced reactive oxygen species with increasing concentrations of hydrogen peroxide results in enhanced stops in the primer extension assay and a decrease in the intensity of the band at position nucleotide 678. From position 396 to 535 seven enhanced replication stops are observed (boxes). Six of the seven boxes are damage at positions where a guanine base (black boxes) is present while other site (outlined box) is damage to other base positions.

Extension produces the full length PCR product (Figure 2.5). When no copper generated reactive oxygen species and no acyLysme is present, primer extension yields a band near the top of the gel at nucleotide position 678. The 392-nucleotide PCR product is more sensitive to hydrogen peroxide treatment than dG. To obtain amplifiable DNA the concentration of peroxide was dropped to 10 mM and the reaction time was reduced to 10 minutes. No major stops to DNA replication are observed. Addition of acyLysme affected extension yield by less than 10% since the product was purified by spin column prior to extension (compare plus and minus lysine lanes, Figure 2.5). Induction of oxidative damage to the PCR product leads to the formation of only 65% in full-length 392-nucleotide product when compared to when no oxidant is present. Addition 30 mM acyLysme with increasing concentrations of hydrogen peroxide led to formation of 55%, 49%, and 8% full-length product compared to no oxidant at 1, 3, and 10 mM hydrogen peroxide respectively. Thus, yield is dependent on the amount of oxidative stress the DNA is exposed to.
**Figure 2.6. Damage Superimposed on the Sequence of the 392-Nucleotide DNA.** The sequence of the DNA substrate is listed with position numbers in bold on the left. The position numbers refer to the Puc19 plasmid. Nucleotides where DNA damage is greater than 3-fold compared to the no hydrogen peroxide control are listed in bold and underlined. The sequences are guanine rich sites except for one.

Incubation with acyLysme causes enhanced replication stops. Addition of acyLysme leads to adducts at several nucleotide positions. Lys addition to a residue is defined as an integrated replication stop band that is 3-fold higher than the no oxidation band. Seven sites of acyLysme addition are observed. We find that acyLysme addition occurs to guanine base at six of the seven sites of damage (black boxes, Figure 2.5). The six sites of cross-linking at guanine base can be visualized by comparing the C-nucleotide sequencing lane to the sites of damage. The nucleotide composition is listed in Figure 2.6. The sites of damage are 5'-GGG, 5'-GG, 5'-GCAGG, 5'-CG, 5'-CACA, 5'-GCAGG, 5'-GTGG starting at positions 412, 421, 434, 455, 480, and 496, respectively. Several of the sites of primer extension stops had wide damage patterns like position 434. Comparing the sites of acyLysme reactivity illustrates that damage predominantly forms at guanine-rich sites. For instance, comparing all seven bands shows that 35% of the total occurs at position 412 or 5'-GGG. Therefore, the site of greatest adduct formation is located at a 5'-GGG. The triple guanine repeat is known to be an
easily oxidizable site\textsuperscript{140}. Primer extension stops at positions 421, 434, 455, 480, and 496 have an integrated relative abundance of 19, 6, 4, 16, 9, and 11\%, respectively. When no acyLysme is added, then stops to primer extension are limited. This data is consistent with the formation of lesions that do not completely stop primer extension. Careful examination illustrates that both oxidative damage and lysine adducts predominate at the same locations (compare plus and minus acyLysme lanes in Figure 2.5). Therefore, Lys addition does not change the location of damage but only enhances replication stopping at the sites.

### 2.4 Conclusion

A simple lysine model system is used to show that copper generated reactive oxygen species can induce addition to guanine base in several circumstances. These reactive oxygen species can proceed via a wide variety of pathways. Oxidation of dG by copper, ascorbate, and hydrogen peroxide increases the number of possible reactions and mimics the multitude of reactions that occur in cellular conditions. Despite these constraints, addition of Lys to dG proceeds in good yield to produce 5-acyLysme–Gh. It has been shown that 8OdG, an intermediate in the formation of 5-acyLysme–Gh (Figure 2.1), predominantly forms the Gh derivatives depending on reaction conditions\textsuperscript{141}. Based on our observation that 27\% of all oxidative guanine modification are lysine addition products, we conclude that under the appropriate cellular conditions Lys addition is competitive with ribose-based chemistry that generally
results from copper-mediated DNA oxidation. This is especially important given the recent literature that implies that Lys side chains from proteins are one of the predominant sources of DNA-protein crosslinks\textsuperscript{142}. When a protein is associated with DNA as in the case of Lys-rich histone proteins, Lys local concentration is very high. In cases of oxidative stress similar to those induced here, 5-Lys–Gh DNA–protein crosslinks are the likely product. This observation is supported by the data showing that the 392-nucleotide PCR product showed Lys addition predominantly at an easy to oxidize triple guanine repeat (Figure 2.5). Interestingly, the triple guanine repeat overlaps a protein binding site for the restriction enzyme Sma1. Sma1 recognizes 5’-GGGCCC and when bound to DNA the local Lys concentration would be high.

Copper is a potent source of reactive oxygen species\textsuperscript{71}. Most cellular copper is not the free divalent ion, but rather bound by proteins\textsuperscript{143}. Despite copper being bound to metalloproteins, like ceruloplasmin, it is still able to initiate oxidative damage\textsuperscript{144}. We show here that copper generated reactive oxygen can lead to Lys addition to dG. Since several disease states possess elevated copper levels and oxidative stress, then it is likely that the formation of these types of lesions may occur. Elevated copper is observed in patients with leukemia and in Alzheimer's Disease\textsuperscript{145}.
Chapter 3

Biologically Relevant Oxidants Cause Bound Proteins to Readily Oxidatively Cross-link at Guanine

3.1 Introduction

An excess of reactive oxygen disrupts cellular functions via deleterious modification of DNA, protein, and other biomolecules\textsuperscript{146}. Reactions centered at DNA are especially important because if they are left unrepaired, permanent mutations or cytotoxic replication stops are formed\textsuperscript{73a}. Increases in the amount of reactive oxygen have been linked to aggressive forms of certain cancers and reduced activity of tumor suppressors\textsuperscript{147}. Additionally, environmental pollutants like hexavalent chromium, a known carcinogen, induce oxidative stress via generation of reactive oxygen\textsuperscript{148}. Chromium has been shown to induce a type of lesion known as DNA–protein cross-links\textsuperscript{88, 149}. These cross-links are poorly characterized and have unknown potential to stop replication. Additionally, cross-link lesions are repaired by digestion of the protein into small peptides that could be mutagenic\textsuperscript{87a}. The lack of biochemical characterization stems from the inherent complexity, unknown stability, and lack of a model system to systematically evaluate these lesions. In this work, we identified and used an appropriate model system to begin to evaluate oxidative cross-links.

Several variants of DNA–protein cross-links are thought to occur. Despite the lack of understanding in structure, oxidative cross-links can be isolated from cells and are known to occur as mixed populations\textsuperscript{114}. One recognized class of cross-links are ternary complexes\textsuperscript{150}. These forms are mutagenic with a metal ion directly attached to DNA and the protein\textsuperscript{151}. The second class of cross-links are those induced by reactive oxygen species. This class is diverse with several structural variants proposed based on small molecule studies. One form involves bonding
between 2'-deoxythymidine and tyrosine stemming from a hydroxyl radical mediated reaction\textsuperscript{152}. Most of the remaining structures are postulated to occur between guanine and lysine or thymine and lysine\textsuperscript{100}. Lysine has been found to add to positions five and eight of guanine\textsuperscript{100, 137c}. Multiple lysine side chains can serve to sequentially add to the nucleobase\textsuperscript{137a}, while other amino acids, like tyrosine, may also be incorporated to guanine\textsuperscript{101}. In addition, it has been shown that cross-links may be derived from reaction of proteins with abasic ribose hemiacetals that form upon depurination of oxidized DNA bases\textsuperscript{153}. Because of this complex mixture of possible products, we wanted to determine both the propensity and structure of oxidative DNA–protein cross-links in a systematic manner.

We hypothesize that guanine base is the most likely addition site of oxidative cross-linking. Guanine is the most oxidation prone DNA base with its oxidation proceeding via several mechanisms\textsuperscript{56, 154}. The divergent mechanisms begin to coalesce upon the formation of oxygen (via either hydroxyl or peroxyl) addition products to either C5 or C8\textsuperscript{155}. The fate of these addition products gives rise to many of the guanine products like mutation-prone 8-oxo-7,8-dihydro-2'-deoxyguanosine (8OdG)\textsuperscript{156}. Cross-link products likely form during these oxidative events; it has been shown that guanine oxidation leads to cross-linking\textsuperscript{89a}. Oxidative DNA–protein cross-links increase concurrently with the levels of 8OdG\textsuperscript{118}. In addition to competing with guanine oxidation, cross-links can form directly from oxidation of 8OdG since it has an oxidation potential that is \textasciitilde 0.5 V lower than the 2'-deoxyguanosine (dG) radical cation\textsuperscript{157}. Major products of hyperoxidation of guanine base include guanidinohydantoin (Gh) and spiroiminodihydantoin (Sp) rings\textsuperscript{136}. A recent crystal
structure demonstrates that hydantoin lesions adopt a syn conformation with the Hoogsteen face base pairing\textsuperscript{158}. These lesions have been found to be mutagenic with insertion of purines opposite the damaged base\textsuperscript{82a, 159}. Hydantoin-based oxidative cross-links are formed when a lysine on a protein adds to DNA instead of water and we hypothesize that they may be a major type of oxidative cross-links.

It is within this context that we began to investigate the types of DNA–protein cross-links formed in order to elucidate their consequences and ultimately facilitate a more specific cellular detection. In this work, we utilize a model protein, ribonuclease A to form cross-links. Cross-link formation was examined using four different oxidation systems: the photo-oxidant rose bengal, the photo-oxidant riboflavin, and iron- and copper-based oxidation systems. Additionally, we determined that cross-links are produced both on single-stranded and on double-stranded DNA with a high dependence of cross-link yield on the number of guanines in the DNA. Cross-link between ribonuclease A and DNA was used as a substrate for primer extension. Using a LTQ-FT mass spectrometer to investigate a model reaction, we found that both lysyl and tyrosyl products are formed.
3.2 Experimental

3.2.1 Materials

All experiments were accomplished in at least duplicates. All chemical reagents were purchased from Acros Organics unless otherwise noted. Oligonucleotides were purchased from Eurofins MWG operon. The sequence of the 27-nucleotide DNA was 5’-GGGGCCCCGTCGTTTTACAACGTCGTG-3’ and its complement was 5’-CACGACGTTGTAAAACGACGGGGCCCC-3’. In studies on the effect of guanine displacement on cross-linking, guanines are methodically removed from the 5’-end of the 27-mer and replaced with adenines. The DNA with four guanines in it has the sequence 5’-AAAACCCCATCGTTTTACAACGTCGTG-3’, while the sixth, seventh, and eighth guanines were successively replaced with adenines for the DNA sequences with three, two, and one guanines, respectively. All DNAs used in these experiments were purified by gel electrophoresis, excision of the band, and overnight elution of the desired DNA band into water, followed by ethanol precipitation.

3.2.2 Methods

3.2.2.1 DNA−Ribonuclease A Oxidative Cross-Linking

Cross-linking experiments were carried out by mixing the 27-nucleotide DNA, ribonuclease A, and the photo-oxidant rose bengal [27.5 µL; 25 mM sodium phosphate, 500 µM DNA, 35 µM ribonuclease A and 100 µM photo-oxidant, pH 8] in a 1.5 mL microcentrifuge tube with the lid open and irradiating with a LED with 3 mW of radiative flux lamp at 575 nm. During photolysis, the solution is held at a distance
of 3 cm from the irradiation source for 20 min. Experiments using CuCl₂ or ammonium iron(II) sulfate based oxidation systems were accomplished in a similar manner but with 3 mM sodium ascorbate, 30 mM hydrogen peroxide, and 50 µM Cu(II) or Fe(II) and with no irradiation. PAGE (12%) was performed. Once complete, the gel was stained with ethidium bromide and visualized using the Multimage II FC UV transilluminator. Yield was calculated via integration using the manufacturer’s software. In experiments where the protein, salt, or DNA was changed, the concentrations are as described. Protein staining was accomplished according to standard protocols and visualized using the 800 channel on a Licor Imager. For double-stranded DNA experiments, the same conditions were employed except that an equimolar concentration of the 27-nucleotide DNA and its complement were added followed by heating at 95 °C for 3 min and gradual cooling to room temperature to facilitate annealing.

### 3.2.2.2 Primer Extension Experiments

Cross-linking in the presence of rose bengal is carried out as before and separated using a PAGE gel. The corresponding cross-linked species were directly visualized using a thin layer chromatography plate and cut out. Overnight elution in 1 mL of ddH₂O followed by precipitation yielded the purified cross-links at a concentration of 5 µM. A 100 µL volume of the purified cross-links was added to a 5X primer extension mix (5X vent buffer, 0.5 mM dNTPs, 500 nM primer, 0.01 U/µL vent exo-DNA polymerase). A fluorescently labeled primer (5′-IrDye700-cacgacgttgtaaaacgac) is utilized in the primer extension. The samples were
subsequently cycled 5 times for primer extension (95 °C for 3 min, 55 °C for 1 min, 72 °C for 2 min). The extended samples were then denatured, and 12% denaturing PAGE was performed. The gel is visualized using the Odyssey Infrared Imager System and the 700 laser channel source. Synthetic oligonucleotides shorter than the 27-mer, with two nucleotides systematically removed, were treated in the same manner as the purified cross-links.

Studies utilizing the 100-nucleotide sequence used a final concentration of 5 µM DNA that was greater than 95% pure. The 100-nucleotide sequence is 5’-ATGACCATGATTACGC-CAAGCTTGCATGCC TGCAGGTCGACTCTAGAG-GATCCCCGGGTACCGAGCTCGAATTCACTGGCCGTCGTTT-TACAACGTCGTG and derived from the puc19 plasmid. To limit oxidation of DNA from direct photo-oxidation, all samples contained 0.5 mM BSA. DNA cross-linking and extension was performed in a similar manner as above with the exception that the DNA was not purified after irradiation but instead was used directly for primer extension by adding 1 µL of the reaction mix to 24 µL of 1.05X extension mix. Sequencing is performed on the puc19 plasmid that has the same sequence using standard extension conditions with an appropriate acyclo-terminator added at 100 µM.

3.2.2.3 LC/MS Product Analysis

The N-terminally acetylated peptides N-Ac-AAAKF and N-Ac-AVKTT (Neo BioScience) are identical to residues 7–12 and 99–104 of the ribonuclease A sequence. Peptides were greater than 95% pure based on HPLC. Peptide–dG reactions (50 µL) were similar to the ribonuclease A–DNA reactions except that 5...
mM peptide and 5 mM dG were used. For the N-Ac-AAA FK-dG reaction, the samples were separated by injection into a Beckman Coulter HPLC with a Cosmobil Waters 5C18-PAQ [4.6 × 150 mm²] column. The gradient was 0% B for 1 min, 14% B over 35 min, 100% over 16 min with solvent A (98% H2O, 2% ACN) and solvent B (96% ACN, 4% H2O). Absorbance was monitored at 240 nm. Controls excluded the listed reagents. HPLC fractions were directly infused into the MS at 5 µL/min. The mass spectrometer is a Thermo Scientific LTQ-FT, a hybrid instrument consisting of a linear ion trap and a Fourier transform ion cyclotron resonance mass spectrometer. The entire elutant was introduced into the LTQ-FT using the standard electrospray ionization source for the instrument with a spray voltage of 5 kV and a capillary temperature of 275 °C. Autogain control, AGC, was used and set at 500 000 with a maximum injection time of 1250 ms for FT-ICR full scans. Collision-induced dissociation, MS^n, was executed in the linear trap with an AGC setting of 10 000 and a maximum injection time of 500 ms. FT-ICR full scans were acquired in the positive ion mode at 100 000 resolving power at m/z 400. Mass accuracy errors were below 200 ppb for full scan and below 800 ppb for MS^n. The positive ion MS^n experiments were performed in the linear trap portion of the instrument using helium as a collision gas, isolation widths of 2 amu, a normalized collision energy of 35, a q value of 0.250, and an excitation time of 30 ms. The same conditions are employed for the addition of the peptide AYKTT to dG and the oligonucleotide 5'-ACGTC-3', except that LC was performed online. Liquid chromatography separation of the sample components used a Waters Symmetry C18 5 µm, 2.1 × 150 mm column, Finnigan Surveyor MS pump, and Finnigan Micro AS autosampler both from Thermo
Scientific. A 200 µL/min gradient elution of water and acetonitrile took place as follows: 2–15% acetonitrile over 35 minutes.

**3.3 Results and Discussion**

![Figure 3.1](image)

**Figure 3.1. Oxidative DNA-protein Crosslink can Form Hydantion Adducts.** When oxidized, DNAs and proteins can crosslink due to the generation of strong electrophiles. DNA-protein cross-links are known to occur in cells with several structures as products. A likely manner of cross-link formation is through nucleophilic addition of amino acids in a protein to DNA following guanine oxidation to form hydantoin-based adducts.

**3.3.1 Characterizing Oxidative DNA–Protein Cross-Links Using a Model System**

This work aims to shed light on oxidative DNA–protein cross-links (Figure 3.1). To accomplish this objective, the employment of an appropriate protein and DNA sequence model system is essential. In cells, much of the DNA is bound by histones, which are lysine-rich. Our initial evaluation found that histones are not suitable as model proteins because their multimeric nature makes product identification challenging and their high pl causes poor band formation upon
electrophoretic separation. Examination of several proteins showed ribonuclease A as a fitting model protein. It is a monomeric protein with a small size. In addition, ribonuclease A binds to DNA in a nonspecific manner similar to most nonspecific DNA-binding proteins (Appendix A1). Ribonuclease A does not cleave DNA because a 2'-hydroxyl is required for nuclease activity. It also serves as a strong template for each type of oxidative DNA–protein cross-link because it possesses 10 lysine and 6 tyrosine residues, representing each reactive side chain of interest. Additionally, it has 8 cysteine, 16 serine, 10 threonine, and 4 arginine residues, which are known to be nucleophilic, as well as 3 phenylalanine residues, which have been shown to cross-link to 2'-deoxyribose as well as to thymine. An unmodified DNA sequence that can predominantly form cross-links was designed. Multiple guanines are placed in the sequence, including four consecutive guanines, which are known to lower redox potential in double-stranded DNA. In addition, a run of four thymines is located at the center of the sequence to allow for formation of thymine-based cross-links. Based on this observation, the oligonucleotide sequence, 5'-GGGGCCCGTCGTTTTACACGTCG, 27-mer, was selected as the DNA substrate.
Figure 3.2. Crosslink Visualization with DNA and Protein Stain. Ribonuclease A and a 27-mer DNA were incubated in the presence of the singlet oxygen generator rose bengal and irradiated. After which, the sample along with the controls were simultaneously run on two gels. One gel visualized DNA by ethidium bromide staining (left) and the second gel visualized protein by Coomassie Blue staining (right). In the lane containing all the reaction constituents, several bands are observed that contain both the protein and DNA components (bands denoted by asterisks). These bands are identified as DNA-protein crosslinks since exclusion of DNA, protein, or oxidant leads to loss of the product. Higher molecular weight DNA-only bands are also observed.

Oxidative cross-link formation is visualized using electrophoresis (Figure 3.2). Reactions and controls were electrophoresed on two gels simultaneously. One gel was visualized with ethidium bromide to detect DNA, while the other is stained with Coomassie Blue to detect protein. The oxidant used in Figure 3.2 is rose bengal, a singlet oxygen generator. The reaction and controls were photo-irradiated using a LED lamp emitting light at 575 nm, which corresponds to the wavelength of maximum absorption for rose bengal. Staining with ethidium bromide shows that DNA and protein irradiated in the absence of oxidant resulted in a single band at the bottom of the gel corresponding to the 27-mer (Figure 3.2 left). Piperidine treatment shows that irradiation is not generating any DNA sites susceptible to cleavage.
Integration of the total peak intensity reveals that no DNA direct strand cleavage is occurring. Addition of rose bengal to DNA in the absence of protein followed by irradiation reveals slower moving bands that are not present when no oxidant is added. These bands are likely DNA–DNA cross-links that have recently been described\textsuperscript{162}. The intensity of these cross-link bands are higher in the absence of ribonuclease A (Figure 3.2, left, lane 2). Quantification of all the DNA bands shows that the 27-mer band intensity is not reduced, indicating that photodegradation of DNA is not taking place. Irradiation of protein and DNA in the presence of rose bengal produced novel bands with higher molecular weight (Figure 3.2, asterisks). These bands were deduced to be covalent in nature since they could not be denatured. Loss of the high molecular weight bands was observed following trypsin digestion indicating that the these bands bands are due to protein associated with the DNA (Appendix A3). The bands resulting from trypsin digestion move slower into the gel than the unreacted DNA bands, which indicates that enzymatic digestion of the protein component of the cross-link resulted to a peptide bound to DNA. This is important since it indicates that these types of adducts are stable to protein digestion, which will be useful in the structural identification of these lesions. Initial optimization showed that more DNA than protein minimizes the formation of multimeric cross-links but lowers overall yield. Multimeric cross-links are secondary, higher molecular weight adducts that form from more than one DNA strand cross-linked to one molecule of protein (see next section). Concentrations of 35 µM ribonuclease A and 500 µM 27-mer promote ample cross-link formation while minimizing multimeric cross-linking. The Coomassie Blue stained gel further
confirms that the new bands consist of a protein component in addition to DNA (Figure 3.2, right). Protein incubated without oxidant and DNA showed a single band at the top of the gel, corresponding to ribonuclease A. Ribonuclease A, having an isoelectric pH of 8.68, will have a net positive charge at pH 8 and is hampered from migrating into the gel. No novel bands were observed after photo-irradiation of protein and DNA in the absence of rose bengal. Rose bengal initiated photo-oxidation of ribonuclease A without DNA results in band broadening, indicating that some protein oxidation is occurring. Photo-oxidation in the presence of rose bengal when both ribonuclease A and 27-mer are present results in new bands that migrate into the gel. The product has a net negative charge from the addition of DNA, allowing the bands to move into the gel. These bands coincide precisely with the product bands observed in the ethidium bromide visualized gel, indicating that these new lesions are made up of both DNA and protein. Exclusion of DNA, protein, or oxidant results in loss of product indicating that these are oxidative DNA–protein cross-links. The yield for the reaction relative to DNA is 7.0% ± 0.9% for total cross-link and 6.0% ± 1.7% for the primary cross-link band relative to DNA. Taking into account the molar ratio of 1:15 for protein/DNA, a yield of 6.6% is the maximum. Thus, the reaction is highly efficient for ribonuclease A.
Figure 3.3. Dependence of DNA-Protein Cross-linking on Oligonucleotide Structure and Oxidant Type. (A) The degree of oxidative crosslinking with a single stranded DNA, SS, is compared with that of double stranded, DS, DNA. The two DNA structures have a comparable tendency to form DNA-protein crosslinks. (B) Four different oxidant systems were used to generate crosslinks at equimolar concentrations of protein to DNA. Fe(II)-H$_2$O$_2$, Cu(II)-H$_2$O$_2$, riboflavin and rose bengal oxidation systems all generated high yields of crosslink despite the divergent oxidation mechanisms. Error bars are standard deviations with n=3. (C) Rose bengal induced crosslinks occur rapidly and linearly increase with irradiation time within the first twenty minutes.

To determine selectivity of cross-linking for single- or double-stranded DNA, a cross-linking experiment was conducted to elucidate which DNA structure is more likely to participate. For cross-links involving double-stranded DNA, equimolar
concentrations of the 27-mer and its complement were annealed to promote double-strand formation. To ensure that equal amounts of DNA were used in both reactions, doubly concentrated 27-mer is employed for comparison. The two forms of DNA were then incubated with ribonuclease A and photo-oxidized in the presence of rose bengal. Controls including without DNA, without protein or without oxidant were used to ensure cross-link identification. During staining, all DNA was single-stranded since a denaturing polyacrylamide gel was used. Visualization with ethidium bromide (Figure 3.3A) indicated that cross-links are formed in both cases (bands denoted by asterisk). Single- and double-stranded DNA produced 7.0% ± 0.4% and 6.5% ± 0.3% DNA− protein cross-links, respectively. Again these values are at the maximum possible yield. Setting the value that forms the most product (SS) as 1.0, double-stranded (DS) DNA only lowers the yield to 92% relative to single-stranded (SS) DNA. The two-tailed p-value for comparing these two analyses is 0.09 indicating that the difference is not statistically significant. This experiment reveals that there is no significant difference in the propensity of single- and double-stranded DNA to form cross-links. This is mechanistically interesting since it is well-known that singlet oxygen reacts less efficiently with DS DNA\textsuperscript{163}, though oxidation initiated at easy to oxidize amino acid residues like tyrosine would not depend on the DS versus SS nature of DNA.

Two major forms of reactive oxygen are common in the cell: superoxide and hydroxyl radical. Several oxidants were chosen to mimic these types of oxidative stress. Superoxide is the most abundant cellular reactive oxygen species and is produced by leakage of the mitochondrial electron transport chain and certain
oxidases. Hydroxyl radical on the other hand is a potent cellular oxidant that reacts with all biomolecules. It is produced in the cell via Fenton reactions when hydrogen peroxide produced from superoxide dismutation, reacts with transition metals like iron and copper. The photo-oxidant riboflavin was used to produce superoxide, as well as for direct oxidation of substrates in a type I fashion. Copper(II), hydrogen peroxide, and ascorbate mixtures produce species that behave as singlet oxygen, hydroxyl radical, and metal bound radicals. Iron(II) produces hydroxyl radical exclusively.

To determine the effect of different cellular oxidant mimics to the amount of DNA–protein cross-links formed, equivalent concentrations of the 27-mer and ribonuclease A were allowed to react in the presence of Fe(II)–H₂O₂, Cu(II)–H₂O₂, riboflavin, or rose bengal oxidation systems (Figure 3.3B). For photo-oxidants rose bengal and riboflavin, 100 µM photo-oxidant was added to 500 µM concentrations of 27-mer and ribonuclease A, followed by irradiation. The concentration of DNA and protein was equimolar since one to one binding increases yield (Figure 3.4C). Unlike for rose bengal, the reaction mixture containing riboflavin is irradiated using a LED lamp that emits at 375 nm, corresponding to the wavelength of maximum absorption for riboflavin. PAGE analysis of the riboflavin-mediated reaction shows that the parent DNA band is smaller when no protein is present, which is indicative that strand scission is occurring. These strand scission products are observed to be less prominent when ribonuclease A is added, which is consistent with the protein being preferentially oxidized (Appendix A4 for gel image).

For Fe(II)–H₂O₂ and Cu(II)–H₂O₂ oxidation systems, experiments were
performed as before but with 50 µM Fe(II) or Cu(II), 30 mM hydrogen peroxide, and 3 mM sodium ascorbate. Samples were immediately visualized by PAGE because preliminary experiments show no noticeable increase in yield with longer incubation time. Initial optimization of reactions involving Cu(II)−H₂O₂ and Fe(II)−H₂O₂ show that lower yields of cross-links are formed at high metal cation concentrations since DNA strand scission is extensive. In addition, significant degradation of the parent DNA is observed when allowed to incubate for several days, indicating that a large portion of the oxidative events are producing strand scission (Appendix A4). The Fe(II)−H₂O₂ oxidation system behaves in a similar manner but with no detectable DNA band degradation. All reactions described are done in at least triplicate.

Comparing the level of cross-linking among the oxidants showed that the photo-oxidants rose bengal and riboflavin produce more DNA–protein cross-links compared with the metal-based oxidants, with rose bengal producing the most cross-links of up to 59.7% ± 0.5% and the Fe(II)−H₂O₂ oxidation system producing the least amount of cross-links, generating only 14.2% ± 0.6%. Riboflavin and the copper−H₂O₂ oxidation system produced 47.0% ± 0.4% and 38.3% ± 1.3% DNA–protein cross-links, respectively. These findings illustrate the ubiquity of DNA–protein cross-linking under conditions of high oxidative stress because all cellular oxidant mimics utilized in these experiments were able to produce DNA–protein cross-links.

For oxidative cross-links to be important cellular lesions, they must form rapidly upon induction of oxidative stress. To determine whether cross-links occur rapidly, ribonuclease A and the 27-mer were exposed to rose bengal at different
time periods. Cross-link formation increases linearly in the first 15 min (Figure 3.3C); thus yield is related to the amount of oxidative events. The linear relation implies that the cross-links are relatively stable and not being further degraded or oxidized in the first 15 min. The same pattern is observed with riboflavin oxidation, while the maximum yield is achieved immediately upon exposure to Cu(II) and Fe(II) oxidation systems.

The importance of binding in DNA–protein cross-link formation was assessed (Figure 3.4). SSB and BSA were selected for this purpose because previous studies show that they can both form DNA–protein cross-links. Looking at the properties of the proteins used in these experiments, ribonuclease A is a monomeric protein, composed of 128 amino acid residues, has a molecular weight of 13,700 Da, contains ten lysine and six tyrosine residues and is known to bind to both double-stranded and single-stranded DNA. SSB is composed of four identical subunits with a molecular weight of 18,900 Da and 178 residues in each subunit. SSB contains five lysine and four tyrosine residues in its sequence and is known to bind to single-stranded DNA. BSA is also monomeric, has a molecular weight of 66 400 Da with 583 residues, and has no known affinity for DNA. It contains 58 lysine and 20 tyrosine residues in its sequence.
Figure 3.4. Effect of Binding on DNA-Protein Cross-linking. To determine the importance of binding on DNA-protein crosslink formation, two experiments were conducted. (A) The tendency of RNase A, SSB, and BSA to form crosslinks are shown. RNase A and SSB utilize non-specific binding modes, while BSA does not bind to DNA. Crosslink yield for SSB and BSA relative to RNAse A is 31% and 3.4%, respectively. (B) Increasing ionic strength is known to interrupt non-specific binding between protein and DNA. Increasing concentrations of NaCl result in decreasing formation of DNA-protein crosslinks. (C) Increasing ratios of Ribonuclease A to DNA lead to greater crosslink yields.
Experiments were performed as before. The concentration of SSB calculated is based only on one subunit. The product bands were subsequently quantified and normalized so that the reaction with ribonuclease A is 1.0. The results show that the cross-link yield for SSB and BSA is 31.2% ± 5.4% and 3.4% ± 3.3%, relative to RNase A, respectively (Figure 3.4A). BSA produced a significantly lower level of DNA–protein cross-links compared with the two DNA-binding proteins suggesting that binding plays an important role in cross-link formation. To further probe the significance of binding in the formation of oxidative DNA–protein cross-links, increasing concentrations of NaCl were added to reaction mixtures of the 27-mer, ribonuclease A, and rose bengal (Figure 3.4B). Setting the cross-link level when no salt is added as 1.0, the most drastic decrease in cross-link level is observed when 0.5 M NaCl is added, as the yield is reduced by 25% ± 1.5%. When the concentration of NaCl was 1, 2.5, and 5 M, cross-linking dropped by 10% ± 2.0%, 18% ± 0.7%, and 5% ± 1.2% respectively. Thus, increasing concentrations of NaCl decrease formation of DNA–protein cross-links, and this demonstrates the importance of a close interaction between protein and DNA.

The yield relative to the amount of protein bound will further support binding as a key parameter in oxidative cross-linking. The ratio of ribonuclease A/DNA was increased from as low as 1:15 up to 2:1 (Figure 3.4C). At low ratios, most DNA is unbound and the total cross-link yield is low. As the ratio increases, total cross-link yield increases such that when protein concentration equals DNA, 60.0% ± 3.3% of the DNA is cross-linked. When there is twice as much protein as DNA, 61.0% ±
0.1% of DNA participates in DNA–protein cross-linking. At ribonuclease A/DNA ratios above 1:10, multiple cross-link bands were observed. The cross-link band migrating the farthest into the gel is identified as the primary cross-link, while higher molecular weight cross-link bands are referred to as multimeric cross-links. In addition to total cross-link quantification, primary cross-links were quantified as well to determine the effects of increased protein concentration.

Primary cross-link yield also increases with protein concentration. At lower protein concentrations, most of the products are primary cross-links, with 85% of the total cross-links due to primary cross-links when 35 µM protein is added. Addition of more protein results in more multimeric cross-links and fewer primary cross-links. Only 47% of the total cross-links are primary cross-links when there is an equivalent concentration of DNA and protein in the reaction mixture. At ribonuclease A/DNA ratios of 1:10 and 1:4, the total cross-links formed exceed the amount of protein, which is the limiting component in the reaction. For instance, 1:4 ribonuclease A/DNA ratio, with 25% protein relative to DNA results in 19% of the DNA participating in primary cross-links and a combined level of cross-links reaching 34%, which is more than the amount of protein present. This suggests that the higher molecular weight cross-links contain multiple DNA molecules bound to a single protein molecule. Considering that, at protein concentrations higher than 50% relative to DNA, not all proteins participate in cross-linking, because 60% ± 3.3% and 61% ± 0.1% of the DNA is cross-linked at 1:1 and 2:1 protein–DNA ratio, respectively. These three experiments highlight the importance of binding in oxidative cross-linking.
Figure 3.5. Guanine Content Correlates with Cross-link Yield. (Left) Ethidium bromide stained gel comparing the crosslinking yield of three 27-nucleotide DNA strands. The first strand is a mixed sequence. The second and third strands lack thymine and guanine, respectively. Loss of guanine leads to loss of product. (Right) Guanines from the 27-mer DNA model were sequentially replaced by adenine. Removal of guanines occurred from the 5'-end toward the 3'-end. The yield decreases as less guanines are present indicating that the crosslink product is likely between guanine and the protein.

3.3.2 Deducing DNA Base Participating in Cross-linking

Experiments were accomplished to determine whether oxidative cross-links occur at thymine or guanine. The formation of guanine-based cross-links by capture of protein lysyl residues would be sensitive to loss of guanine base in the sequence. Guanine bases within the 27-mer were removed and replaced with adenine (Figure 3.5). The amount of DNA and protein was held to a ratio of 15:1 and the cross-link bands were quantified. When all the guanines were substituted with adenine the yield dropped from 7.2% ± 0.5% to 0.7% ± 0.6% (Figure 3.5, left). Thymines within the 27-mer were then replaced by cytosine. In this case, no change in the overall yield was observed (Figure 3.5, left, lane 2). Guanines were systematically replaced with...
adenine from the 5′- to the 3′-end (Figure 3.5, right). The 27-mer has nine guanines. Substitution of the first five and six guanines with adenine led to a decrease in yield to 6.5% ± 0.8% and 6.2% ± 1.0%, respectively. When two or one guanine remained, the oxidative cross-link yield fell to 3.2% ± 0.5% and 2.1% ± 0.6%, respectively. The change in cross-link yield is highly sensitive to the number of guanines. Thus, the strong correlation between the number of guanines and the yield infers that cross-links are occurring predominantly at guanine sites. It should be noted that the terminal guanine repeat is not the most sensitive site of cross-linking. Rather, removal of the sixth guanine from the 5′ position most significantly decreases the cross-link yield, which indicates that this particular guanine is important for cross-linking in this DNA. This experiment was also accomplished for the Cu(II)–H₂O₂ oxidation system. Similar results are observed (Appendix A5). Even though Fenton reactions are known to induce thymine-based cross-links, we find guanine content to be essential for oxidative cross-linking in both systems.
Figure 3.6 Guanine is the Major Cross-link Site. (A) The 27-nucleotide DNA was cross-linked to ribonuclease A, and the product subsequently purified and annealed to a labeled primer. Isolation and extension of a no oxidant control (lane NO RXN) led to full extension when compared with the truncated sequences on the right lanes (C4G2, C4, C2). C4G2 lacks the last two guanosine nucleotides from the 5’-end of the 27-mer, while C4 lacks the last four guanosine nucleotides, and C2 lacks the last 6 nucleotides. Extension of the cross-linked product resulted in a stop before the poly-G tract (C4). (B) A 100-nucleotide mixed sequence DNA, rose bengal, bovine serum albumin and ribonuclease A at the listed concentrations (top) were
incubated. Lack of irradiation or ribonuclease A led to extension predominantly to the full length product. As the ribonuclease A concentration increases, new, lower molecular weight bands are observed. Boxes to the left mark the new bands. Black boxes (7 of 8 new bands) represent cross-links at guanine. The remaining white box terminated two nucleotides before a guanine.

To further elucidate the dependence of cross-linking on guanine, primer extension studies were accomplished. A primer for the 27-mer was utilized for these experiments to ensure significant replication stopping and oxidative cross-link stability during extension conditions. The 27-mer was incubated with ribonuclease A, photo-oxidized using rose bengal, and separated by PAGE, followed by cross-link purification. Analysis of the ribonuclease A–DNA cross-link band showed that it did not degrade (Appendix A6). We inferred that cross-linking at sites on the primer binding site would block primer annealing and would not be observed in the assay. The only remaining guanine site on the 27-mer is the 5′-polyguanine repeat. The cross-link was then annealed to a fluorescently labeled primer and extended (Figure 3.6) to determine whether the presence of a DNA–protein cross-link will halt replication. The location of the replication stop was determined based on comparison to migration of the unreacted 27-mer and synthetic oligonucleotides that were shorter than the 27-mer. The shorter oligonucleotides possess the intact primer binding site but have two nucleotides systematically removed. For instance, the sequence C4G2 (Figure 3.6A) lacks the 5′-GG and is 25 nucleotides long. These markers were also extended and used to locate the site where cross-linking induced a replication stop. After primer extension, samples were separated by PAGE and visualized. When no polymerase (Figure 3.6A, No ENZ) is added, a single band is observed at the bottom of the gel, indicating the unreacted primer. Addition of
polymerase to unmodified DNA gave a full-length product (Figure 3.6A, NO RXN). In contrast, extension of the cross-linked product (Figure 3.6A, DPC) led to a stop precisely one nucleotide before the polyguanosine tract. Notice that C4 and DPC migrate equally. Extension occurred but could not go past the site of cross-linking. Additionally, cross-links can be accurately identified using primer extension because they are stable toward elevated temperatures and basic conditions. Furthermore, this result underscores that these large lesions are strong replication stops: we observed no extension past the cross-link site.

A synthetic 100-nucleotide DNA was used as a substrate for oxidative cross-linking (Figure 3.6B). This substrate is a mixed sequence and therefore will determine which nucleotides are cross-linked on a large DNA. The 100-nucleotide DNA was present at 5 µM, while the level of BSA was held constant at 500 µM to limit direct DNA photo-oxidation. To these samples, increasing concentrations of RNase A were added. BSA does not induce cross-links while ribonuclease A does (Figure 3.4A); thus, when more ribonuclease A is present, more cross-links should be observed. After irradiation at 575 nm, samples were immediately extended for a single cycle and separated via PAGE. Samples were compared with thymine, guanine, and cytosine sequencing ladders. The “G” ladder corresponds to modified guanine sequences on the 100-nucleotide DNA template. The yield of full-length product is above 90% in all cases. When ribonuclease A is incubated with the 100-nucleotide template without oxidation full length products predominate because the elevated temperatures during extension facilitate its denaturation and unbinding. Incubation of DNA with BSA and subsequent oxidation led to a small amount of
damage owing to direct reaction of singlet oxygen with DNA. The concentration of ribonuclease A was changed from 100 to 200 and to 300 µM. Eight new shorter DNA bands are observed to increase in intensity with increasing concentrations of ribonuclease A (Figure 3.6B, boxes). Seven of the eight new bands are located one nucleotide from guanine (black box, compare with guanine ladder). One of the new products observed is not located at a guanine, but rather at the 5'-cytosine in the sequence 5'-GGGCC (white box). The 3-fold increase in ribonuclease A concentration causes a 4-fold increase of each polymerase stop with little variation (the increase at each site is plus or minus 30%) at the eight sites. It should be noted that each of the eight sites has a unique propensity to form cross-links, with the single non-guanine cross-link being the second most intense. This data shows that guanine is the major site of oxidative cross-linking.

3.3.3 Mass Spectral Analyses of dG-Peptide Cross-links

Due to the difficulty in assessing the structure of an intact ribonuclease A, a strategy where the product structure was determined using small molecule models was employed. Because several types of guanine-lysine cross-links are possible, two smaller peptides obtained from the ribonuclease A sequence, both containing a lysine residue, were utilized in these experiments. Furthermore, tyrosine is known to add to thymine, which is why one of the peptides selected contains a tyrosine residue. The peptide N-acetyl-AAAKF was used as a substrate in the oxidative cross-linking reaction in Figure 3.7, while N-acetyl-AYKTT was used as a substrate in Figure 3.8. The peptide N-acetyl-AAAKF is only capable of forming lysyl-type
cross-links and the peptide N-acetyl-AYKTT can form each of the three types of oxidative cross-links: those between lysine and guanine base, those formed by addition of lysine to depurinated nucleotides, and those between tyrosine and thymine bases[153]. The two peptides comprise residues 7–12 and 99–104 of ribonuclease A, respectively. Because binding is an important factor in cross-link formation, the concentrations of both peptide and nucleoside were increased to 5 mM as lower concentrations resulted in much lower yields. The yield was below 5% for both peptides despite the one-to-one ratio and the elevated concentration. The peptide was incubated with the nucleoside and irradiated as before. Products were analyzed by mass spectrometry. The mass spectrometer is a hybrid instrument that consists of a linear ion trap for low resolution spectra along with tandem mass spectrometry (MS^n) and a Fourier transform ion cyclotron resonance mass spectrometer for high resolution spectra in addition to high mass accuracy. Thus, an introduced sample has both spectra acquired. The low-resolution MS samples are sampled faster and, therefore, have a more accurate TIC. The high-resolution instrument is used to obtain elemental compositions. Mass accuracy for high-resolution mass spectrometry has errors below 200 ppb for full scan and below 800 ppb for MS^n. The MS^n fragmentation is accomplished in the linear ion trap and analyzed in both modes. Therefore, MS^n also has high resolution with high mass accuracy spectra to accurately identify fragments. It is important to note that many guanine hyper-oxidation products are not stable in acid. Hydantoins are modified under acidic conditions; thus the addition of acid can potentially affect results[165]. We optimized LC and MS conditions in water with no acid added since we have found
that low pH, even during direct injection, does lead to severe degradation.

Figure 3.7. LC-MS Analysis of Rose Bengal Oxidized Adduct of dG and the Peptide N-Ac-AAAKF. (A) A peptide containing a single cross-link active lysine and dG were utilized to assess if high-resolution, multi-dimensional MS could be used to establish the location of a cross-link. The HPLC-UV trace of the reaction and the controls are shown. The reaction chromatogram is in black while the controls lacking N-Ac-AAAKF, dG, rose bengal or light are in grey. The 10.5 and 24-minute peaks were unique to the reaction, denoted by B and C respectively. The products were isolated and analyzed by mass spectrometry. (B) Mass spectral analysis of the 10.5-minute peak revealed that it has an observed mass and elemental composition consistent with formation of a spiroiminodihydantoin. MS² analyses show exclusive fragmentation at guanine. MS³ can be used to determine the site of crosslinking based on the y²⁺ and b⁴⁺ ions. (C) The 24-min peak shows similar results except that more DNA fragmentation is seen in MS². The 24-min peak is a 5-lys-guanidinohydantoin crosslink based on elemental
composition and MS$^3$ fragmentation. All fragments (grey) are acquired using high resolution MS with errors below 1 ppm.

The reaction of dG with N-acetyl-AAAKF led to formation of hyper-oxidized guanine derivatives (Figure 3.7). An example of a HPLC-UV trace is shown in Figure 3.7A. Two new prominent product bands are observed when the reaction chromatogram (black) is compared with the controls, which excluded a reaction component. Control reactions include exclusion of the peptide, dG, rose bengal, or irradiation (gray traces). The clearly observable products migrate at 10.5 and 24 min (denoted by B or C). The reaction trace shows that these products are not in low abundance and that LC/MS is not being used to detect a minor product. The reaction products were analyzed. The MS spectra for the 10.5 min peak is shown in Figure 3.7B. The 10.5 min product has an observed mass of 830.3792 that corresponds to an elemental composition $\text{C}_{36}\text{H}_{52}\text{N}_{11}\text{O}_{12}$ with an error of 44 ppb. Because our conditions do not use acid, a strong sodium adduct is observed with an error of 202 ppb. The 10.5 min product was fragmented in MS$^2$. The ion trap spectra are shown, while the high-resolution MS is in the appendix. A single major fragment was observed with a mass of 714.3316 (error = 349 ppb) or a loss of $\text{C}_5\text{H}_8\text{O}_3$. The 714.3 fragment is due to the loss of 2′-deoxyribose indicating adduct formation is between the guanine and the peptide. Further fragmentation during MS$^3$ can be used to accurately assign the location of modification on a peptide. The seven most abundant ions during MS$^3$ fragmentation are shown. The peptide fragments $b_4^+$,$y_4^+$,$y_2^+$,$y_3^+$, and loss $\text{H}_2\text{O}$ all have a mass that includes the hyperoxidized guanine base. The loss of $\text{H}_2\text{O}$ likely derives from the carboxyl terminus.
Additionally, $y_4^+$ is observed to also lose the C-terminal phenylalanine. The ion at 549.3 could have been the $b_4^+$ ion, 549.2528, or loss of guanine, 549.3031. High resolution MS showed only 549.2527, determining the ion to be $b_4^+$. Only a single modification of guanine is observed with a loss of CHNO. This ion stems from ring cleavage at either C8−N7 or C6−N1. Importantly, the $y_2^+$ and $b_4^+$ show that the guanine must be added to the lysyl residue since the other peptide positions are lost without the modified guanine base. The mass and fragmentations identify the lesions as lysyl-spiroiminodihydantoin lesions, but the isomeric forms cannot be determined.

Figure 3.8. LC-MS Identification of Rose Bengal Oxidized Adduct Between dG and the Peptide N-Ac-AYKTT. (A) A peptide containing both a tyrosine and
lysine were reacted with dG. Selected ion chromatograms from the reaction are shown. The top two chromatograms are N-acetyl-AYKTT (black) and dG (grey). The total ion chromatogram (middle) shows the starting materials as the most abundant ions. Monitoring of ions with masses of 880 (black) and 906 (grey) show several peaks eluting within one minute (bottom). (B) Analysis of the 17.7-minute peak shows masses of 880.4158 and 906.3951. Experimental masses are in black while theoretical masses are in grey. (C) MS$^2$ analysis of the 880 peak show that the lesion is guanidinohydantoin. (D) The 906 peak can either be a lysine-spiroimino-dihydantoin or a tyrosine-guanine crosslink. The b$^{2+}$ ion bisects the tyrosine and lysine revealing that tyrosine is modified.

The product formed using the peptide N-acetyl-AYKTT and dG is then determined. This peptide can be used to compare the propensity of a tyrosine–guanine and lysyl–guanine to form in a single peptide. The LC/MS of the reaction is shown in Figure 3.8A. Two slowly migrating products were observed with elemental compositions of C$_{37}$H$_{58}$N$_{11}$O$_{14}^{+}$ and C$_{38}$H$_{56}$N$_{11}$O$_{15}^{+}$ and a mass of 880.4158 and 906.3951, respectively. The largest error in all cases was 276 ppb. We looked for a cross-link formed from an abasic nucleoside and lysine that has been recently observed$^{104}$. The product would have an elemental composition of C$_{33}$H$_{53}$N$_{6}$O$_{13}^{+}$ with a mass of 741.37, which we did not detect by selective ion monitoring. Additionally, we also did not observe any thymine–tyrosine oxidative cross-links when a similar reaction with 2′-deoxythymidine was performed. The total ion chromatogram shows the most abundant ions to be unmodified peptide N-Ac-AYKTT and dG (Figure 3.8, compare top two traces). When 880.4 and 906.4 were monitored, we found several bands all eluting within 1 min (compare traces). This data is consistent with formation of both conformational isomers and stereoisomers of hydantoin rings. The MS$^2$ of 880.4 reveals that it is a guanidinohydantoin–peptide adduct with the major product having the peptide appended on C5 of Gh via lysine
addition. The 906.4 ion was 10 times more abundant than the 880.4 ion. The 906.4 ion can be either a lysine–spiroiminodihydantoin or a tyrosine–guanine(ox) cross-link. The two most abundant ions that result when the 906 ion was fragmented were the 790.4 ion, which corresponds to the loss of 2'-deoxyribose, which confirms that the adduct is between guanine and the peptide, and the 888.4 ion, which is due to the loss of H₂O (Figure 3.8D). In order to determine the amino acid involved and to identify the adduct structure, MS³ of the 790.4 ion was obtained. The 15 most abundant ions during MS³ fragmentation are shown. Two of the resulting ions are due to guanine modifications. The 722.4 peak corresponds to the loss of C₂N₂O, arising from the removal of C6–N1–C2. This fragment can occur for both the lysine–spiroiminodihydantoin and the tyrosine–guanine(ox) adduct. The loss of CHNO is also observed, which can arise from elimination of either C8–N7 or C6–N1, also possible for both lysine–spiroiminodihydantoin and the tyrosine–guanine(ox) adduct. The peptide fragmentations are b₄⁺,b₃⁺,b₂⁺, and y₄⁺, in addition to y₄⁻TT, b₄⁻CO₂, and b₄⁻C₂H₄O. Loss of one or two H₂O molecules, alone, along with CO₂ or from the b₄⁺ ion, is also observed. The H₂O molecules can come from the two threonine residues, while the CO₂ molecule can come from the carboxyl terminal. The C₂H₄O lost from b₄⁺ is likely from the remaining threonine residue. The b₃⁺,y₄⁺, and y₄⁻TT together imply that either lysine or tyrosine is adding to guanine, since the b₃⁺ ion is due to the loss of the two threonine residues adjacent to lysine, the y₄⁺ ion is due to the loss of the N-acetyl-alanine next to tyrosine, and the y₄⁻TT ion is due to the loss of both. The b₂⁺ ion on the other hand is indicative of tyrosine substitution to guanine since the observed mass, which is 442.1469 (error
−64 ppb), is 165.1 mass units greater than the predicted mass when guanine is not included. This mass corresponds to the mass of the hyper-oxidized guanine; thus the b2+ ion is N-acetyl-AY-guanine(ox). A 497.2 peak, which is a y3+ ion for lysine substitution to guanine, is found but is buried in the background. These data confirm that tyrosine modification is occurring but does not rule out lysine modification since the ions may be differentially detected by MS ionization. These data show that oxidative cross-linking is facile, and product characterization shows similar trends. First, the MS2 generally cleaves at dG leaving a peptide and a small portion of the DNA used to identify the site of cross-linking. Second, these data illustrate that both tyrosine−guanine(ox) and lysine−guanine(ox) readily form in good yields.
Figure 3.9. LC-MS Analysis of Rose Bengal Oxidized Adduct Between 5’-ACGTC-3’ and the Peptide N-Ac-AWKTT. The peptide N-Acetyl-AWKTT is reacted with a 5-nucleotide DNA containing the four natural DNA bases. Results displaying linear ion trap MS² are shown. (A) Selected ion chromatograms for the reaction. (B) Analysis of the 12.6 minute peak shows a doubly charged mass of 1036.30 corresponding to the Gh adduct. (C) Analysis of the 11.5 minute peak shows a doubly charged mass of 1049.29 consistent
with the mass of Sp addition to the peptide. Almost exclusive DNA fragmentation (through w-ions) was observed with \( w_3^{2-} \) and \( w_2^{2-} \) identifying the cross-link site to be on guanine.

To extend the DNA model, the peptide N-Ac-AYKTT, possessing two cross-link active amino acids is reacted with a 5-nucleotide DNA with the sequence 5'-ACGTC-3', which contains the four natural DNA bases, in the presence of rose bengal. The reaction is monitored by LC-MS in both the positive and negative ion modes. The results displaying linear ion trap MS\(^2\) in the negative ion mode is shown in Figure 3.9. The selected ion chromatograms for the reaction are shown in Figure 3.8A. The total ion chromatogram is shown at the top. Monitoring of ions with a mass of 623 (orange) identifies the peptide as the most abundant ion, eluting at 9 minutes. Monitoring ions with masses 1036 (green) and 1049 (blue) corresponding to doubly charged Gh and Sp adducts, respectively shows several peaks for each of the two masses, with the largest 1036 peak eluting at 12.6 minutes and the largest 1049 peak eluting at 12.6 minutes. This is consistent with isomers where either lysine or tyrosine can form adducts with the DNA, which will exhibit identical masses if the cross-link site is on guanine, in addition to the conformational isomers that can form with the hydantoin rings. Almost exclusive DNA fragmentation (through w-ions) is observed, as shown in Figures 3.9A and 3.9B. Figure 3.9A shows the MS\(^2\) fragmentation of the peak with m/z 1036, which is consistent with the doubly charged Gh adduct. The most abundant ions are shown. The peak with m/z of 1419 shows the loss of TC and CHNO, while the 1462 peak is the loss of TC, both peaks showing no loss of the peptide mass. The \( w_2^{2-} \) ion correspond to the
mass of ACG and the peptide, which indicates that the peptide is bonded on either of these three bases. The $w_3^2$ ion corresponds to the loss of adenosine without the peptide, suggesting that it is likely bound to either on C or G. The $w_4^2$ ion is consistent with the loss of AC without the peptide, which supports cross-linking at guanine. Additionally, the peak with m/z 1310 corresponds to the loss a mass corresponding to Gh and the peptide further illustrating cross-linking at guanine. The peak with m/z 1551 is due to singly charged $w_3$ ion.

Figure 3.9B shows MS$^2$ fragmentation of the 1049 peak, which has a mass identical to the predicted mass of doubly charged Sp adduct. The most abundant peaks are shown. Again, the $w_2^2$ ion shows the loss of TC, indicating that the peptide is likely adducted to A, C or G. The $w_3^2$ ion is due to the loss of AC, suggesting that the adduct is with G, T or C. Put together, these data indicates that the peptide is bound on G. Again, the 1310 peak shows loss of Sp and the peptide, supporting adduct formation at guanine. Mass errors for the parent ions are less than 300 ppb. Since w-ions are produced, and no peptide fragmentation is observed, leads to the conclusion that cross-link structure is a strong influence on the MS/MS fragmentation patterns. Analysis of a short segment of DNA and a peptide can identify the base involved in cross-linking, while digestion of the DNA to the nucleoside level may be important in identifying the amino acid involved in cross-linking as this is shown promote fragmentation on the peptide.
3.4 Conclusion

We show that oxidation-induced DNA–protein cross-links have a high propensity to form. Earlier studies on DNA-protein cross-links utilized oxidation-prone 8OdG-substituted oligonucleotides, the presence of which increases the propensity for oxidative cross-linking at a single site on the DNA. This study shed light on oxidative cross-links by determining key parameters in the reaction. The first key observation is that unmodified DNA can readily lead to DNA-protein cross-link reactions in high yields of up to 61%. Despite the very different oxidation mechanisms of the four oxidants used they all have the potential to form oxidative cross-links. These oxidants are classic mimics of cellular oxidative stress. A key and underappreciated factor in successful cross-linking is binding. Addition of salt or co-incubation with non-binding proteins severely reduces cross-link efficiency. A second key observation is that these lesions likely derive from guanine modification. We show by substituting guanine with adenine in the 27-nucleotide DNA and from primer extension studies that these products are forming at guanine sites. Structural analysis shows that guanine forms guanidinohydantoin, spiroiminodihydantoin, and an aromatic addition product from guanine hyper-oxidation.

The story from the amino acid components is much different. We were able to isolate both tyrosine- and lysine-based cross-links. Due to its more favorable oxidation potential, tyrosine oxidation leads to formation of tyrosine-based cross-link even in the presence of an adjacent lysine. This indicates that the formation of these tricyclic tyrosyl adducts can compete with the more well-known lysine-based cross-links. Mechanistically, it appears that lysine cross-links stem from DNA oxidation as
we observe the formation of 5-lys, which is diagnostic of DNA oxidation, while tyrosine adducts stem from protein oxidation. These divergent reaction products may have little consequence when a whole protein is attached to DNA since all cross-links are likely strong replication stops for replicative polymerases. In fact, DNA–protein cross-linking mediated by the antitumor trans-[PtCl2(E-iminoether)2] was shown to constitute a fairly strong replication block for exonuclease-deficient Klenow fragment and RT HIV-1, although a small amount of translesion synthesis was observed. A family polymerase, DNA polymerase α, trans-lesion synthesis may be involved in cellular tolerance of DNA–protein cross-link. Polymerase α can bypass exceptionally large lesions whose linkages are through the DNA-major groove but not the minor groove. This polymerase can polymerize past extremely large major groove lesions including DNA–protein and DNA–DNA cross-links. It is established that cross-links are repaired by digestion of the protein into small peptides. At the stage of a DNA–peptide cross-link the exact structure of a cross-link will be more important because the smaller lesions may show differential mutagenicity since these lesions are easier to bypass. The high yields of product formation illustrate that oxidative cross-links are under-appreciated as DNA lesions. Oxidative cross-links may be present at high relative levels in cells since the propensity to oxidatively cross-link is high and so much of the genomic DNA is coated with protein.
Chapter 4

DNA-Protein Interactions Prevent Hyper-oxidized DNA and Allows Buildup of 7,8-dihydro-8-Oxo-Guanine
4.1 Introduction

Oxidative DNA damage has been implicated in cancer, aging and neurological disorders\(^{168}\). Oxidative damage to DNA can be mediated by reactive oxygen species (ROS), including superoxide radical anion, singlet oxygen, hydrogen peroxide and hydroxyl radical, which are present in the cell as a consequence of endogenous reactions, or from exogenous sources, including ionizing radiation\(^{5b,169}\). Oxidation of DNA not only results in small modifications like oxidation of the guanine base but also a set of large lesions like DNA-protein cross-links (DPC). In this chapter, we explore the influence that amino acids and proteins play in altering the types of DNA lesions both large and small.

Figure 4.1. Proteins Inhibit Singlet Oxygen Hyper-Oxidation of Guanine. Singlet oxygen oxidation of guanine in DNA can result to the formation of hyper-oxidized adducts including spiroiminodihydantoin (Sp) and oxazolone (Z), which are known to be strong mutagens in the cell. Genomic DNA is closely associated with a variety of proteins performing structural, regulatory and repair functions. Close association of proteins with DNA increases the effective concentration and allowing for the formation of DNA-protein cross-links (DPC) under oxidative stress conditions. We hypothesize that the activity of bound proteins is not only to participate in DPC formation but can largely influence the oxidative process. A
consequence of this interaction can be the inhibition of these hyper-oxidation event, allowing the primary guanine oxidation product, 8-oxo-7,8-dihydro-2'-deoxyguanosine (8OdG) to persist in biological systems.

We hypothesize that the activity of bound protein is not only to participate in DPC formation but can significantly influence the oxidative process (Figure 4.1). It has been estimated that the endogenous level of DPCs in human white blood cell range between 0.5 to 4.5 in $10^7$ bases$^{92}$, with this amount being about one order of magnitude less than 8-oxo-7,8-dihydro-2'-deoxyguanosine (8OdG). DPCs are known to accumulate with age in mouse organs and correlate with 8OdG, likely because in the presence of amines, 8OdG oxidation produces hydantoin-amine adducts, relevant in cross-link formation between DNA and lysine-rich proteins like those present in histones$^{118,170}$. Previous work from our group and others, have demonstrated that in vitro, bound proteins readily form cross-links to DNA when exposed to a variety of biologically relevant oxidizing agents$^{89,137b,171}$. This rapid formation owes to the large effective concentration of protein side chains that can compete with solvent to react with DNA-bound radicals and intermediates. In a preceding study, we have shown that several bound proteins form adducts with DNA with exceptionally high yields using biologically relevant ROS forms$^{89c}$. Critical works by Burrows, Greenberg, and Stemp show that a variety of proteins, substrates, and oxidant systems can accomplish the cross-linking reactions$^{89a,89b,137b,171-172}$. This chapter addresses not only DPC formation but also how proteins alter smaller oxidation products. For example, a bound protein with redox active amino acids can change the types and amounts of DNA oxidation products. In addition, the quasi-stable$^{173}$
DPCs can be fragmented giving a different DNA lesion with an altered mutational spectrum. Given that the protein is such an important variable we are exploring how it influences an oxidation reaction.

This chapter focuses on the oxidative consequences to guanine. Guanine is preferentially attacked during oxidative DNA damage owing to its low redox potential (1.3 V vs NHE) among the four DNA bases. Despite it being relatively hard to oxidize guanine is present at such high concentrations that it is able to react with ROS based oxidative equivalents in vivo. Reaction of guanine with singlet oxygen leads to initial formation of an endoperoxide that subsequently forms 8OdG. An important characteristic of 8OdG is that it has an even lower redox potential (0.7 V vs NHE) than dG, making it highly reactive towards further oxidation. A number of in vitro studies have shown that 8OdG within a double helix is highly prone to subsequent oxidative damage, forming what are termed hyper-oxidation products including spirominodihydantoin (Sp), guanidinohydantoin (Gh), Imidazolone (Z) and its hydrolysis product, oxazolone (Z). Recently, the occurrence of these secondary oxidation products have also been explored in the G-quadruplex folds of the human telomere sequence. These guanine hyper-oxidation products however, have limited detection in vivo with one case from E. coli, and recent reports indicating occurrence in mammalian cells, with the background level of these lesions determined to be 100 times lower than the parent 8OdG in the liver and colon tissue of Rag2−/− mice. This is in contrast to 8OdG, which has been detected ubiquitously using variety of methods. One possibility is that upon formation, the
concentration of 8OdG is sufficiently low that it is not able to compete (i.e. the net $\Delta G$ is positive), with easy to oxidize protein side chains despite the reduced potential. Thus, we propose that the presence of easy to oxidize protein side chains prevent 8OdG from further oxidation, resulting to its buildup.

In this chapter we focus on singlet oxygen (${}^{1}\text{O}_2$). This reactive oxygen has been shown to exclusively target the guanine moiety of nucleosides, oligonucleotides and isolated DNA, which is the focus of these investigations\textsuperscript{174}. Singlet oxygen may be efficiently generated in cells by a range of biological enzymes, in addition to exposure of biological systems to visible or ultra-violet light in the presence of appropriate sensitizers via the Type II photosensitization mechanism that is accounted by the triplet energy transfer to molecular oxygen\textsuperscript{179}. Cells contain a range of chromophores that can act as sensitizing agents for singlet oxygen formation\textsuperscript{30}. Singlet oxygen has also been postulated to be an important intermediate in a range of biological systems formed by a variety of peroxidase enzymes\textsuperscript{31}. Furthermore, we have shown from previous work that singlet oxygen generates the most DNA-protein cross-links when compared to Type I photo-oxidation and hydroxyl radical generating systems. This makes singlet oxygen an ideal oxidant for this investigation.

In this work, we identified the singlet oxygen-mediated oxidation products of dG with and without the highly oxidizable protein, BSA, to be spiroiminodihydantoin (Sp) and oxazolone (Z), with Sp being the major product. Addition of select proteins reduces hyper-oxidation events, resulting to the buildup of 8OdG. We show that electron-rich highly oxidizable amino acids and
proteins protect 8OdG from oxidation, with the level of 8OdG preservation, roughly correlating with the amino acids oxidation rate constant$^{180}$. The presence of the antioxidant N-Acetyl Cysteine (NAC) diminishes dG-lysine adduct formation but only affects DNA-protein cross-linking at concentrations exceeding physiological levels. Tyrosine is observed to be the most efficient cross-linker, both in the nucleoside and oligonucleotide studies, allowing cross-linking capacity that is competitive with other oxidation events. Tyrosine containing peptides do not discriminate between single- and double stranded DNA in adduct formation, while double-stranded DNA is the preferred substrate for lysine cross-linking.

4.2 Experimental

4.2.1 Materials

All experiments are accomplished in at least triplicates. All chemical reagents are purchased from Acros Organics unless otherwise noted. Oligonucleotides were purchased from Eurofins MWG operon. The sequence of the 27-nucleotide DNA is 5’-GGGGCCCCGTCGTTTTACAACGTCGTG-3’ and its complement is 5’-CACGACGTTGTAAACGACGGGGCCCC-3’. The N-terminally acetylated peptides N-Ac-AYKTT, N-Ac-AFKTT, N-Ac-ASYK and N-Ac-ASKF were obtained from Neo Bioscience. Proteins were obtained from Sigma Aldrich.
4.2.2 Methods

4.2.2.1 Nucleoside Experiments

Nucleoside oxidation experiments were carried out by mixing dG and the photo-oxidant rose bengal with or without an amino acid, peptide or protein of interest [100 µL; 5 mM dG, 10 mM amino acid or peptide, or 500 µM protein and 100 µM photo-oxidant, 25 mM sodium phosphate, pH 8] in a 1.5 microcentrifuge tube with the lid open and irradiating with a LED with 3 mW of of radiative flux lamp at 375 nm. During photolysis, the solution was held at a distance of 3 cm from the irradiation source for 4 hours. During irradiation, the samples were submerged in water at 22°C to limit evaporation. The samples were directly separated by injection of 2 µL into a Beckman Coulter System Gold with an autosampler and a diode array. A Phenomenex Synergi 4U C18 [4µm, 150 x 4.6 mm] reverse phase column was used for resolving nucleoside oxidation samples without antioxidant. The gradient was 0% B for 1 minute and 95% B over 15 minutes with solvent A (98% H₂O, 2% ACN) and solvent B (96% ACN, 4% H₂O). The product of the reaction was pre-purified by collecting the unique peak that elutes at the void volume of the RP-HPLC run. The fraction was then concentrated 10X via low drying rate speedvac for ~4 hours and re-suspended in 0.1% acetic acid (Hypercarb column mobile phase), followed by injection into a Thermo Scientific Hypercarb [5 µm, 30 x 2.1 mm] column to resolve the polar components. The gradient was 0% for 8 minutes and 100% B over 2 minutes with 0.1% acetic acid for solvent A and methanol for solvent B. In the presence of
amino acid, peptide or protein, a Cosmolis 5C18-PAQ [5µm, 250 x 4.6 mm] reverse phase column was used, with a gradient of 0% B for 1 minute, 14% B over 35 minutes and 100% B over 5 minutes, with the same solvent composition. Absorbance was monitored at 240 nm. Controls excluded the listed reagents. Data was quantified using a 32 Karat HPLC software.

4.2.2.2 Oligonucleotide Experiments

Oligonucleotide studies were carried out by mixing the 27-nucleotide DNA, ribonuclease A or peptide, and rose bengal [20 µL; 500 µM DNA, 500 µM ribonuclease A or 10 mM peptide, 100 µM photo-oxidant, 25 mM sodium phosphate, pH 8] in a 1.5 mL microcentrifuge tube with the lid open and irradiating with a LED with 3 mW of of radiative flux lamp at 375 nm. During photolysis, the solution was held at a distance of 3 cm from the irradiation source for 20 minutes for protein analyses and 4 hours for peptide studies. During irradiation, the samples were submerged in water to limit evaporation. For product distribution analyses, the samples were mixed with water and piperidine [100 µL; 100 µM DNA, 100 µM ribonuclease A, or 2 mM peptide, 1 M piperidine] and heated at 90°C for 30 minutes, followed by ethanol precipitation. PAGE (12% for protein and 20% for peptide studies) was performed. Once complete, the gel was stained with ethidium bromide and visualized using Multimage II FC UV transilluminator. Yield was calculated via integration using the manufacturer’s software. For double-stranded DNA experiments, the same conditions were employed except than an equimolar concentration of the 27-
nucleotide DNA and its complement were added followed by heating at 95°C for 3 minutes and gradual cooling to room temperature to facilitate annealing.

4.3 Results and Discussion

4.3.1 Nucleoside Studies

Nucleoside studies are focused on dG as the important target of oxidation in DNA. The formation and resistance to further oxidation of its major primary oxidation product, 8OdG in the presence of relevant oxidizable amino acids and select peptides and proteins is explored. The results of these analyses will attempt to respond to the prevailing question of why 8OdG persists in biological systems despite its ease of oxidation. In this section, we quantify 8OdG and guanine hyper-oxidation products when dG is oxidized via the singlet oxygen forming photosensitizer rose bengal with and without the presence amino acids, peptides, or proteins.
Figure 4.2. Singlet oxygen oxidation of 2'-deoxyguanosine (dG). The reactions are examined by HPLC detection at 240 nm (A). The black trace represents the reaction when all components are present. To identify the product of this reaction, controls (grey, from top to bottom) where light, rose bengal, or dG is removed from the reaction. Irradiation of dG and rose bengal results to the formation of a product peak at 1.7 minutes. (B). The 1.7 product peak was isolated and the components separated using a hypercarb column. Three distinct peaks are observed at 1.7, 2.4 and 5.9 minutes. (C and D) MS analysis of the isolated peaks indicates that the 1.7 and 2.4 min peaks are diastereomers of dSp (m/z 300) and the 5.9 min peak is dZ (m/z 247). (E) Product distribution analysis with and without BSA shows that the presence of the protein reduces level of dG hyper-oxidation by 32% and that the ratio of dSp to dZ product in both reactions with 52% of the product formed without and 56% with BSA. This indicates that close association with proteins in the cell reduces hyper-oxidation events while distribution of dG hyperoxidation products by singlet oxygen is unaffected by the protein.
Figure 4.2A shows an example of HPLC chromatograms following photo-oxidation of dG with rose bengal. Product formation is detected by HPLC using first C18 reverse phase. To identify the product of this reaction, controls were used, where irradiation (top), rose bengal (middle, top) and dG (middle, bottom) were taken out of the reaction. The 5.5-minute peak is dG. After rose bengal induced oxidation, a unique peak at 1.7 minutes is detected when all reaction components are present (bottom). The amount of dG consumed is 66 ± 4% after 4 hours of irradiation at 375 nm. Values are obtained from three independent experiments. Aside from the prominent peak at 1.7 minutes, no other product peak is detected, indicating that the entire product is merged into a single highly polar band.

The highly polar band at 1.7 min was resolved using an appropriate column. Upon rose bengal photo-oxidation it was found that three hyperoxidation products are generated. The peak at 1.7 min was collected and injected using a Hypercarb column, which is designed for the separation of more polar molecules. The merged peak resolved into three products, eluting at 1.7, 2.4 and 5.9 minutes as shown in figure 4.2B. Product assignments are accomplished using known absorbance spectra and MS analysis. Absorbance spectra can be found in appendix A8. Figures 4.2C and 4.2D show the mass spectra obtained for each of the isolated peaks. It is determined that the 1.7 and 2.4 minute peaks have identical mass spectra, with m/z 300 and m/z 184 as the most prominent peaks. This indicates that these two peaks are the two diastereomers of dSp as the m/z 300 is consistent with protonated dSp, while m/z 184 coincides with the
mass of protonated Sp. Figure 4.2D shows the mass spectrum for the 5.9 minute peak. The 247 m/z is consistent with protonated dZ. Fragmentation of the 247 peak resulted in the formation of a peak with m/z 131, which corresponds to the loss of deoxyribose, indicating that this molecule is a modified nucleoside (appendix A8). The experiment was initially carried with irradiation at 560 nm and showed the same product distribution pattern as with the 375 nm irradiation but at a much lower yield (appendix A9). Product distribution analyses are conducted with and without BSA to determine how the presence of a readily oxidizable protein alters the relative amounts of products formed and consequently, the resulting mutational spectrum (Figure 4.2E). The analysis is accomplished by integration of the product peak areas while monitoring absorbance at 240 nm, followed by normalization of peak areas using each compound’s unique extinction coefficient at this wavelength\(^{121}\). Addition of BSA resulted in the reduction of the hyper-oxidized peak area by 32\%, with 45 ± 3\% of dG undergoing hyper-oxidation, as opposed to 66 ± 4\% without BSA. This indicates that the presence of the protein results in the inhibition of the hyper-oxidation reaction. In the absence of BSA, 48 ± 2\% of the total hyper-oxidation product is dZ and 52±1.2\% is dSp, while 44 ± 4\% of the product is dZ and 56 ± 1.1\% is dSp when BSA is present. The major product for both reactions is dSp with p-values between dSp and dZ of 0.00182 and 0.00022 with and without BSA, respectively, which is in agreement with the findings of Ye and co-workers, indicating that Sp is the major secondary oxidation product when guanosine is exposed to rose bengal generated singlet oxygen at neutral and basic pH\(^{81a}\). The
1.7 minute dSp peak constitutes 44 ± 2% of the total dSp product, while the 2.4 minute peak comprises 56 ± 1% when no protein is added. With BSA, the 1.7 minute peak make up 47 ± 1%, while the 2.4 minute peak consists 53 ± 2% of the total dSp present. The later eluting dSp peak constitutes the greater part of the product in both conditions. Normalizing total hyper-oxidation products to 100% and comparing the percentages of each product for samples with and without protein added, a p-value of 0.18273 is obtained. This indicates that, though the presence of protein reduces dG hyper-oxidation events, it does not significantly alter the relative amounts of hyper-oxidized products.

Figure 4.3. Easily Oxidized Amino Acids and Proteins Enhance 8OdG and Limit Formation of Hyper-Oxidation Products. The listed amino acids (10 mM)
were co-incubated with dG and oxidized. The reactions were examined by HPLC with detection at 240 nm. An example cysteine chromatogram is shown (A). Controls (grey, listed from bottom to top) were exclusion of dG, cysteine, and rose bengal. Absorbance profile (inset) reveals that unique peak at 5.7 min is 8OdG. Analyses of 8OdG levels on amino acids (B) and the peptide ASXF (X represent listed amino acids) (C) reveal that easily oxidized amino acids, cysteine and tyrosine results to the highest conversion of dG to 8OdG. The presence of proteins significantly increases the level of dG to 8OdG conversion compared to peptides and amino acids (D). The percentage of dG to 8OdG conversion is listed. RNase A results to 5.4% conversion, while BSA, cytochrome and histone generate ~5X more 8OdG. Error bars represent standard deviation from three independent experiments.

Free amino acids increase 8OdG and decrease hyper-oxidized dG. Figure 4.3A in black shows a sample chromatogram for a photo-reaction containing 10 mM Cys, 5 mM dG, and 0.1 mM rose bengal. Controls include reactions (Figure 4.3A, grey, from bottom to top) lacking dG, Cys, and rose bengal. Addition of Cys results in formation of a sizable 8OdG peak when compared to standard. A calibration curve was used to quantify the amount in moles of 8OdG. Data are analyzed as moles 8OdG over total nucleoside (i.e. the percent 8OdG). The results with various amino acids are listed in Figure 4.3B. No 8OdG was observed when Leu, Arg, or when no amino acid was added. Addition of Lys and Phe results in 1.0 ± 0.1% and 1.0 ± 0.2% percent 8OdG. Inclusion of Tyr in the reaction results to 2.6 ± 0.3% percent 8OdG, while Cys generates the most 8OdG with 7.7 ± 0.4% 8OdG. Of the amino acids tested, Tyr and Cys are the easiest to oxidize having oxidation potentials of ~0.8 and 0.7 V vs NHE, respectively. Lys has an oxidation potential of 1.1 V vs NHE. This suggests that addition of easy to oxidize amino acids cause a buildup in 8OdG.
concentration and prevents it from undergoing further oxidation. Furthermore, of the amino acids tested, only Lys and Tyr formed adducts with dG.

Peptides were then examined systematically. A similar experiment is conducted with the peptide N-Ac-ASXF, where X is one of the amino acids of interest: Leu, Arg, Lys, Tyr and Cys to determine if a similar effect can be observed when the amino acid is incorporated into a peptide. Since the amino acid Leu possess an alkyl side-chain that is not easily oxidizable, the peptide N-Ac-ASLF serves as the negative control. To offset the effect of Phe, which is initially shown to also impede 8OdG oxidation, the moles of 8OdG generated by N-Ac-ASLF (negative control) is subtracted from that of the peptides. Figure 4.3D shows the levels of 8OdG generated in this context. The moles of 8OdG resulting from peptide addition is significantly lower than that of the corresponding amino acids, but a similar pattern can be observed with the two oxidizable amino acids Tyr and Cys forming the highest level of 8OdG, having 1.1 ± 0.2% and 0.9 ± 0.2%, respectively. The p-value for the two peptides is 0.2058 indicating that the difference is not statistically significant, unlike that of the amino acid context where the presence of Cys resulted to a significantly higher level of 8OdG. Amino acids incorporated into a peptide are less effective in protecting 8OdG from further oxidation. The level of 8OdG is reduced by 58% between the amino acid and the peptide context for Tyr. The most significant reduction in 8OdG level is observed in the case of Cys with 88% decrease in the nucleoside level when the peptide is present. This is likely due to steric constraints introduced by incorporation into peptide and, likely, slower diffusion,
which can impede amino acid interaction with oxidants. Thus a diminished capacity to compete with 8OdG, which has fewer steric restrictions, is observed.

Proteins increase 8OdG concentration during nucleoside oxidation. Structurally diverse proteins, rich in oxidizable amino acids, Cys and Tyr, and those that can potentially form close associations with DNA are selected. The proteins RNase A, BSA, Cytochrome c and Histone were tested (Figure 4.3D). RNase A is a Lys, Cys and Tyr-rich protein used in previous work and has been shown to bind to DNA in a non-specific manner, ultimately forming cross-links. Bovine serum albumin is acidic and Cys and Tyr-rich. Histones are basic, Lys-rich and are closely associated with nuclear DNA. Cytochrome c is also basic, Lys-rich, and can interact electrostatically with DNA. Additionally, the proteins BSA and Cytochrome c have been identified to exhibit antioxidant properties.

Due to solubility limitations, the amount of protein added is 0.5 mM for all reactions to allow systematic variation. Figure 4.3D shows the relative levels of 8OdG formed in these reactions. RNase A generates 5.0 ± 2.0 % 8OdG relative to nucleoside, and Histone, Cytochrome c and BSA forming 22 ± 1, 24 ± 1 and 25 ± 1%, 8OdG respectively. All proteins tested were able to shield 8OdG from further oxidation. Cytochrome c and BSA, the two antioxidant proteins tested, generate a significantly higher level of 8OdG compared to RNase A and Histone, with a p-value of 0.0228 between Histone and Cytochrome c. Cytochrome c and BSA exhibit similar capacity to inhibit 8OdG oxidations with a p-value of 0.2001. The 8OdG levels in the presence of proteins are at least 5X that of the peptides tested. This data underscores the importance of proteins as antioxidants in
biological systems, protecting the organism’s genetic material from oxidative damage.

**Figure 4.4. Comparison of Oxidation Product Distributions when Peptides or Amino Acids are Added.** Addition of peptides or amino acids alters total dG oxidation (A) oxidation and the relative level of guanine hyperoxidation products (B) when total dG conversion is held at 50%. The distribution of 8OdG, dG hyperoxidation, cross-links, and peptide/AA direct oxidation is shown (C). The peptide N-Ac-AYKTT forms the most cross-links with dG likely due to the easy to oxidize tyrosine, while dG hyper-oxidation events are more favored in the presence of lysine. This indicates that the DNA oxidation products are likely dependent on the surrounding molecules or associated proteins. Additionally, DNA-protein cross-linking by tyrosine and lysine is competitive with other DNA oxidation events, with a minimum level of 11.6% of the total dG oxidation when lysine is present. Error bars represent standard deviation from three independent experiments.
Changes from cross-link active peptides were then analyzed. Lys and Tyr were used in product distribution analysis since initial survey of nucleophilic amino acids show that these are the only amino acids that can form adducts with dG following oxidation with singlet oxygen. In addition, the peptides N-Ac-AYKTT and N-Ac-AFKTT were also used in the nucleoside context. The peptide N-Ac-AYKTT possesses both Lys and Tyr in the sequence, allowing for the determination of the effect of two cross-link active amino acids on cross-link yield. The amount of dG remaining after four hours of irradiation with or without a peptide or amino acid is determined (Figure 4.4A). In the absence of peptide or amino acid, 66 ± 4% of dG is consumed (34% dG remaining). Addition of the peptides N-Ac-AYKTT and N-Ac-AFKTT, results to the reduction of the dG peak to 53 ± 1 for both peptides. Addition of Tyr and Lys result in a dG peak reduced to 72 ± 6% and 53 ± 1%, respectively, of the peak size when no oxidant is added. Since the dG hyper-oxidation product distribution has little change (see Figure 2E) the intensity of the resulting 1.7 minute HPLC peak can be used to quantify the levels of hyper-oxidized dG (Figure 4.2A). Values given are integrations measured and normalized to 100 when no peptide is present. (Figure 4.4B). Addition of 10 mM N-Ac-AYKTT, results to a hyper-oxidation product peak area that is 31 ± 5% of the area when no peptide is added. Considering that 66 ± 4% of dG is converted when no peptide is present, it is deduced that 20 ± 3% of dG is hyper-oxidized in the presence of N-Ac-AYKTT (Figure 4.4C). The presence of the peptide, N-Ac-AFKTT and the amino acids Tyr and Lys, results to a dG
hyper-oxidation of 31 ± 5, 15 ± 2 and 40 ± 1%, respectively. The presence of a peptide or amino acid results in a significant reduction in hyper-oxidation product, which is more prominent when Tyr is present. This indicates that some of the oxidative equivalents are absorbed by redox active peptides as should be expected. As discussed before, 8OdG is one of the products of the reaction. The amount of 8OdG generated by each reaction is determined using a calibration curve plotting known 8OdG concentration against the area of the peak. The peptides N-Ac-AVKTT and N-Ac-AFKTT result to 0.9 ± 0.04 and 0.8 ±0.05% of dG converted to 8OdG. The presence of Tyr results to the highest level of 8OdG with 2.5 ± 0.3% conversion, while Lys generates 1.0 ± 0.1% 8OdG. The amount of cross-link generated is deduced by subtracting the amount dG remaining, dG converted to 8OdG and the amount of hyper-oxidized dG from the known concentration of dG prior to oxidation. Of the molecules tested, the peptide N-Ac-AVKTT produces the highest amount of cross-links as 26 ± 1% of dG is involved in adduct formation with the peptide. This is likely due to the presence of two available sites of cross-linking in the sequence. Both the peptide N-Ac-AFKTT and Tyr generate 14 ± 1 and 14 ± 3% cross-links, respectively, while the addition of Lys results to 7 ± 1% adduct formation. Because the peptides N-Ac-AVKTT and N-Ac-AFKTT and the amino acids, Tyr and Lys can be oxidized competitively with dG, the difference between the amount of dG hyper-oxidation products when no peptide/amino acid is present and the products when a peptide/ amino acid is added (hyper-oxidation product, 8OdG and cross-link), is taken as the dG oxidative equivalents absorbed by
these molecules. The peptides N-Ac-AYKTT and N-Ac-AFKTT and the amino acid, Lys are comparably oxidized with 19 ± 1, 18 ± 1 and 19 ± 1%, respectively. Tyr absorbed the highest level of oxidative equivalents with 36 ± 1%. The most abundant product when the peptide N-Ac-AYKTT is added to dG is the peptide-dG adduct, followed by hyper-oxidation, which is comparable to peptide oxidation, and low levels of 8OdG. The hyper-oxidation of dG is the most prominent product when the peptide N-Ac-AFKTT is present, followed by peptide oxidation, adduct formation and low levels of 8OdG. When Tyr is present, most of the oxidative equivalents is absorbed by the amino acid. Cross-linking is competitive with dG hyper-oxidation and an 8OdG level of ~2.5X higher compared to the other molecules tested. In the presence of Lys, dG hyper-oxidation is the most prevalent, followed by Lys oxidation, cross-linking, and finally, low levels of 8OdG. A reduction in cross-link level when N-Ac-AFKTT is used compared to N-Ac-AYKTT indicates that the hydroxyl group in Tyr is important in initiating cross-linking.

4.3.2 Oligonucleotide Studies

Studies on the oligonucleotide context are conducted to determine efficiency of cross-linking and hyper-oxidation product distributions on a realistic substrate. The oligonucleotide was a 27-nucleotide double stranded DNA. The sequence of strand one is 5’-GGGGCCCCGTCGTTTTACAACGTCGTG-3’, while the sequence of strand two is 5’-CACGACGTTGTAAAACGACGGGGCCCC-3’. In order to visualize the product we initiated this study with the peptide N-acetyl-AYKTT since this peptide have been shown to form adducts with guanine via
both Tyr and Lys and form appreciable amounts of cross-links with the model DNA. Hyper-oxidized guanine and cross-links are alkali labile\textsuperscript{89a,124} and we take advantage of this characteristic to determine product distribution between cross-linking and dG hyper-oxidation events. In addition, yield of cross-linking is compared for single and double stranded DNA to determine the preferred substrate for each amino acid of interest.

Tyrosine cross-link efficiency is competitive with other oxidation events in a dsDNA substrate. The peptides N-Ac-ASXF (where X is Leu, Arg, Lys, Tyr and Cys) are tested and compared against N-Ac-AYKTT to identify those that can form cross-links with the DNA model and select the appropriate peptides for the product distribution analyses. The peptide lengths are tailored so that cross-link formation between the peptide and DNA model can be readily determined via electrophoretic mobility shift assay as the product size allows for facile separation from un-reacted DNA. The peptides containing Arg, Cys, Lys and Tyr were found to form adducts with the DNA model with yields of 1.5 ± 0.1% for Arg and Cys, and 10.7 ± 2 % and 17.4 ± 0.4%, for Lys and Tyr, respectively. Due to the considerably lower level of adduct formation between the DNA model and the peptides containing Arg and Cys, only the peptides N-Ac-AYKTT, N-Ac-ASYF and N-Ac-ASKF are used for the product distribution studies (Appendix A10 for gel image). The product distribution is deduced by calculating the cross-link yield from the product band intensity following detection by ethidium bromide. This yield is then compared with total alkali labile DNA oxidation, which can be revealed by piperidine treatment. Alkali labile sites are surrogates for hyper-
oxidized dG and cross-links. Many oxidative lesions, excluding 8OdG, are alkali labile and addition of hot piperidine results to DNA strand scission\textsuperscript{124}.

\textbf{Figure 4.5. Oxidation Product Distribution in the Oligonucleotide Context.} The product distribution is deduced by comparing total DNA modification, revealed by hot piperidine treatment, with cross-link yield. The sample gel image (A) shows bands following cross-linking reaction of the 27-nucleotide DNA model and N-Ac-AYKTT, with and without the addition of hot piperidine. Reaction in the...
absence of hot piperidine shows a slowly migrating band corresponding to the DNA-peptide adduct (lane 1). Addition of hot piperidine to the reaction mixture results to the generation of shorter oligonucleotides, corresponding to the bands that migrate faster into the gel compared to the unmodified DNA (lane 2). Exposure to light without rose bengal did not result to cross-links (lane 3) or strand scission (lane 4) when piperidine is added as shown by the absence of the adduct and modified DNA bands. Product distribution analysis (B) shows that the peptide N-Ac-AYKTT forms the most adduct with DNA, followed by N-Ac-ASYF and N-Ac-ASKF. The highest level of DNA modification not resulting to cross-linking occurs when N-Ac-ASKF is added. Comparison of cross-link yield between single-stranded (ss) and double-stranded (ds) DNA, reveals that ds DNA produce more cross-links than ss DNA, with the peptide N-Ac-ASKF (p-value = 0.002).

Figure 4.5A shows a sample gel image for the reaction with the peptide N-Ac-AYKTT. Lane 1 shows cross-linking reaction in the absence of piperidine, while lane 3 shows a control lacking oxidant. The slower moving band unique to the reaction is identified as the cross-link. Addition of piperidine (lanes 2 and 4), results to strand scission and an accompanying reduction in the parent DNA band intensity when oxidant is added (lane 2), while no strand scission is observed in the absence of oxidant (lane 4). Figure 4.5B shows the result of the product distribution analysis. The percentage of hyper-oxidized DNA is calculated by subtracting the amount of DNA cross-linked from the amount of alkali labile DNA. The oxidative equivalent absorbed by the peptide is calculated by deducting the amount of DNA consumed with the peptide from that consumed without the peptide. The total amount of DNA consumed is due to DNA-peptide adducts and DNA oxidation that did not result to cross-linking. When no peptide is added, all DNA consumed is due to oxidation. As in the nucleoside context, DNA-peptide adduct is the primary product formed when the peptide N-Ac-AYKTT is present, with 26±2% of the DNA participating in cross-linking, 22±1% is
hyper-oxidized and 37±1% of the oxidative equivalents is absorbed by the peptide. Of the three peptides tested, N-Ac-AYKTT also generates the most cross-links, followed by N-Ac-ASYF with 17±0.4% of the DNA cross-linked and N-Ac-ASKF with 11±2%. This indicates that the presence of two cross-link active amino acids in a peptide allows for more possibilities for cross-linking. In addition, results from both the nucleoside and oligonucleotide contexts suggest that Tyr is a better cross-linker than Lys. The presence of N-Ac-ASYF results to 37±2% of the DNA consumed, which is the lowest among the three peptides, 20±2% of this is due to DNA oxidation while 17±0.4% is due to cross-linking. The peptide absorbed the highest level of oxidative equivalents with 47±2%. Addition of the peptide N-Ac-ASKF results to 36±2% DNA oxidation, the highest among the three and the lowest amount of cross-links. DNA oxidation when N-Ac-ASKF is present is competitive with peptide oxidation. Comparison of cross-link level between single- and double-stranded DNA in Figure 4.5C shows that there is no significant difference in yield with the two forms of DNA with the peptides, N-Ac-AYKTT (p-value = 0.8) and N-Ac-ASYF (p-value = 0.1), while double-stranded DNA produces more cross-links with the peptide N-Ac-ASKF (p-value 0.002), compared the single-stranded DNA. Due to its high side-chain pKa (10.53), the dominant form for lysine is likely protonated at pH 8, indicating that electrostatic interaction may play a role in the increased cross-link yield.
Figure 4.6. N-Acetyl-Cysteine Does Not Inhibit Cross-linking of a Protein Bound to DNA at Physiological Concentrations. (A) Changes in dG-lysine adducts. Rose Bengal mediated dG oxidation in the presence of lysine leads to the formation of 5-lysine-spiroiminodihydantoin and limited 8OdG (black trace). Upon addition of 10 mM N-acetyl-cysteine the spirominodihydantoin product is lost and formation of 8OdG is restored. (B) In contrast cross-linking of a 27-nucleotide DNA and RNase A via rose bengal is resistant to disruption by N-acetyl-cysteine. Only upon of non-physiological concentrations above 40 mM are statistically significant changes in cross-linking observed. Error bars represent standard deviation from three independent experiments.

Antioxidant N-Acetyl Cysteine (NAC) also inhibits dG lysine adduct formation when equivalent amounts of NAC and lysine are present. Figure 4.6A shows a representative chromatogram for the singlet oxygen-mediated reaction of 10 mM N-α-acetyl-lysine-methyl-ester (N-Ac-Lys-OMe) with 5 mM dG. Absorbance profile comparison and mass spectral analysis of the 3.8 minute
peak shows that it is 5-Lys-Sp (appendix A11). Low levels of 8OdG are observed at 5.2 minutes. To determine how the presence of easily oxidizable amino acids affect cross-linking, 10 mM NAC, equivalent to the N-Ac-Lys-OMe concentration is added to the reaction. Figure 4.6B shows the chromatographic trace for this reaction. Addition of NAC resulted to loss of Lys-dG adduct formation (note the absence of 3.8 minute peak), while an increase in the 8OdG peak intensity is evident. This suggests that the presence of an antioxidant inhibits 8OdG oxidation, and consequently, the subsequent nucleophilic addition that ordinarily forms cross-links. A similar experiment is carried out for a cross-linking reaction between RNase A and the 27-nucleotide DNA model (Figure 4.6C). Increasing concentrations of NAC was added to the reaction and the effect on the cross-link level is determined. NAC concentrations from 0-80 mM, is added to the reaction. The cross-link level for reaction with no NAC added is normalized to 1.0 and compared with the products of reactions where NAC is added. At 10 and 20 mM NAC, no observable reduction in cross-linking is observed. Addition of 40 & 80 mM NAC results to ~40% and 86% reduction in cross-linking, respectively. This indicates that elevated levels of antioxidants can protect the system from DNA-protein cross-linking, while physiological antioxidant concentrations (less than 20 mM) have no observable effect in cross-link levels.
4.4 Conclusion

Proteins are the major components of most biological systems comprising approximately 68% of the dry weight of cells and tissues\(^{180}\). Some of these proteins are associated with genomic DNA performing functions including structural, regulatory and repair. Under conditions of oxidative stress, one consequence of this close contact is the formation of covalent adducts between a protein and the nearby DNA resulting in DNA-protein cross-links. DNA-protein cross-links are extremely bulky lesions that are likely to impair various DNA processes including replication, transcription and repair\(^{90b, 98}\). In the last two chapters, structures of guanine adducts with lysine have been identified to involve hydantoin structures, which are products of dG hyper-oxidation. Several research groups are attempting to detect hydantoin based lesions in biological systems, with little success until recently, where dSp and dGh are detected in the liver and colon of Rag2\(^{-/-}\) mice, but at levels that are 100 times lower than the primary dG oxidation product, 8OdG\(^{85}\). Binding of proteins with DNA likely results to a high effective concentration enabling the protein to compete with water for the reaction with the nucleobases and allowing for the formation of covalent bonds between DNA and protein. The low detection of these lesions, in addition to the effective protection of antioxidant proteins in biological systems, which inhibits their formation, is likely because most of the hydantoin lesions are participating in DNA protein cross-links.

This chapter explores the possibility that in addition to DNA-protein cross-links, another consequence of close protein-DNA interaction during oxidative condition
is the persistence of 8OdG in the system. Our findings indicate that 8OdG is protected from further oxidation in the presence of highly oxidizable amino acid side chains, with the 8OdG level roughly correlating with the amino acids’ oxidation rate constants. Experiments on dG oxidation in the presence of proteins concur with the results on the amino acid level, showing 8OdG conservation for all proteins tested. The anti-oxidant amino acid, N-Acetyl-Cysteine (NAC) inhibit adduct formation between dG and Lys and was unable to interrupt cross-linking between DNA and RNase A except when a large excess of NAC is present. This is likely due to the inhibition of 8OdG oxidation, which is postulated to initiate adduct formation with a nucleophilic amino acid. The reduced effectiveness of NAC to inhibit cross-linking between the DNA and RNase A is likely due to the close electrostatic association between the DNA and the protein, which increases the effective concentration between guanine and the cross-linking amino acid. Thus, efficient adduct formation occurs each time a suitable 8OdG is oxidized. Cross-linking experiments in the nucleoside level show that Tyr forms cross-links with dG with twice as much efficiency as lysine. The presence of both Lys and Tyr in a peptide results to the highest level of adduct formation with a cross-link yield greater than the sum of the yields for each individual amino acid. The presence of Tyr in the system allows for cross-linking efficiency that is competitive with dG hyper-oxidation. In contrast to the results in the nucleoside studies, the cross-link yield for N-Ac-AKYTT is significantly lower than the sum of the yield for N-Ac-ASKF and N-Ac-ASYF. Competitive yields for cross-linking and DNA oxidation is observed for peptides
containing Tyr, while DNA oxidation is the major product formed when lysine is present. Furthermore, Tyr-containing peptides do not discriminate between single- and double-stranded DNA in adduct formation, while double-stranded DNA is the preferred substrate for the lysine containing peptide.

The amino acids Cys and Arg were unable to form adducts with dG in the nucleoside context, but low levels of cross-linking are observed when the peptides N-Ac-ASCF and N-Ac-ASRF are reacted with the 27-nucleotide DNA model in the presence of singlet oxygen. Arginine forming adducts with DNA substrates containing 8OdG have been previously reported but no structural information has been determined\textsuperscript{89a}. These data indicate that these two amino acids can potentially form adducts with DNA in biological systems. A logical next step will be to identify the participating nucleobase and adduct structure.

Many proteins that bind to DNA contain electron-rich amino acids in their binding site that are preferably attacked by oxidants\textsuperscript{185}. We show successful preservation of 8OdG in the presence of highly oxidizable amino acids and proteins. This lesion is only mildly mutagenic, but it can be readily oxidized to Sp and Gh, which have been shown to exhibit mutation frequencies that are at least an order of magnitude higher than 8OG\textsuperscript{82}. This suggests that proteins may act as sacrificial anodes, protecting DNA from damage as some of these amino acids have redox potentials\textsuperscript{181} that are lower than that of guanine, the most oxidizable DNA base, and 8OG. Additionally, we demonstrate that DNA-protein cross-link yields can compete with other DNA oxidation events, with Tyr as the more efficient cross-linker. This is important since Tyr is present in the active sites of
some proteins that bind to DNA including single-stranded DNA binding proteins and topoisomerases$^{185a, b, 186}$. We illustrate the possibility that proteins have divergent roles during oxidative insult, one in forming deleterious covalent cross-links with DNA, and another in protecting 8OdG from undergoing oxidation, salvaging the cell from the mutational pressure imposed by its oxidation products.
Chapter 5
Future Directions
5.1 Characterization of DNA-Protein Cross-links in More Complex Systems

Figure 5.1. DNA-Protein Cross-link Characterization in Complex Systems. LC-ICPMS and LC-MS/MS methods can be used to identify cross-link site and structure following isolation and enzymatic digestion of protein and DNA.

Mass spectrometry has advantages in the analysis of cross-links in that, it has the ability to analyze high molecular weight complexes, both the intact and the enzyme-digested products. Mass spectrometry also has a high sensitivity for peptides and oligonucleotides. Despite the advantages of mass spectrometry in analyzing cross-links, identifying the site of cross-linking is challenging due to the low abundance of cross-linked species and the high background of non-cross-linked components. ICP-MS is a powerful analytical tool for element specific analysis because the signal response is in direct proportion to the element...
concentration in the sample, and the ionization source is not sequence or matrix-dependent\textsuperscript{188}. The latest developments on ICP-MS allows for the detection for elements that are previously difficult to detect, such as phosphorus, which will be utilized in the next step of this work.

The next step will be to develop a method for structural characterization and establishing cross-link site in a DNA protein model employing inductively coupled plasma (ICP), accurate mass and tandem mass spectrometry techniques. The proposed scheme for this work is shown in Figure 5.1. We have shown oxidative DNA-protein cross-links can be separated from un-cross-linked DNA and protein by polyacrylamide gel electrophoresis, and the adduct can be isolated by excising the portion of the gel containing the cross-link into a solvent of choice. An enzymatic digestion on the protein portion of the cross-link can be conducted and the peptide-oligonucleotide adduct can be selectively detected and purified via ICP-MS by monitoring for elemental phosphorus present in the phosphodiester linkage of the DNA\textsuperscript{189}. Following isolation by ICP-MS, the DNA portion of the adduct can be enzymatically digested to yield a nucleoside-peptide adduct. Identified adducts using small molecule model system will aid in narrowing down candidate lesions. LC-MS/MS analysis will reveal the identity of the peptide, and consequently the participating protein following data base searching and analysis. A successful method of lesion detection and identification will allow for the determination of relevant proteins in a human cell culture model system and eventually be applied in the identification of ROS-related disease biomarkers. We currently have a method for cross-link isolation.
and quantification in the cell. This method can be modified to assess the feasibility of the proposed method\textsuperscript{190}.
References:


94. Quievryn, G.; Zhitkovich, A., Loss of DNA-protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteosome function. *Carcinogenesis* 2000, **21**(8), 1573-1580.


110. (a) Lee, K. F.; Chung, W. Y.; Benzie, I. F., Urine 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxodG), a specific marker of oxidative stress, using direct,
isocratic LC-MS/MS: Method evaluation and application in study of biological variation in healthy adults. *Clinica chimica acta; international journal of clinical chemistry* 2010, 411 (5-6), 416-22; (b) Loft, S.; Danielsen, P.; Lohr, M.; Jantzen, K.; Hemmingsen, J. G.; Roursgaard, M.; Karotki, D. G.; Moller, P., Urinary excretion of 8-oxo-7,8-dihydroguanine as biomarker of oxidative damage to DNA. *Arch Biochem Biophys* 2012, 518 (2), 142-150.


120. (a) Kornyushyna, O.; Burrows, C. J., Effect of the oxidized guanosine lesions spiroiminodihydantoin and guanidinoxydantoin on proofreading by Escherichia coli DNA polymerase I (Klenow fragment) in different sequence contexts. *Biochemistry* 2003, **42** (44), 13008-18; (b) Niles, J. C.; Wishnok, J. S.; Tannenbaum, S. R., Spiroiminodihydantoin and guanidinoxydantoin are the dominant products of 8-oxoguanosine oxidation at low fluxes of peroxynitrite: Mechanistic studies with O-18. *Chemical Research in Toxicology* 2004, **17** (11), 1510-1519; (c) Hailer, M. K.; Slade, P. G.; Martin, B. D.; Rosenquist, T. A.; Sugden, K. D., Recognition of the oxidized lesions spiroiminodihydantoin and guanidinoxydantoin in DNA by the mammalian base excision repair glycosylases NEIL1 and NEIL2. *DNA repair* 2005, **4** (1), 41-50.


139. Merino, E. J.; Barton, J. K., Oxidation by DNA charge transport damages conserved sequence block II, a regulatory element in mitochondrial DNA. *Biochemistry* 2007, 46 (10), 2805-2811.


147. (a) Vafa, O.; Wade, M.; Kern, S.; Beeche, M.; Pandita, T. K.; Hampton, G. M.; Wahl, G. M., c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: A mechanism for oncogene-induced


157. (a) Stover, J. S.; Ciobanu, M.; Cliffl, D. E.; Rizzo, C. J., Chemical and electrochemical oxidation of C8-arylamine adducts of 2'-deoxyguanosine. *J Am Chem Soc* 2007, 129 (7), 2074-81; (b) Hall, D. B.; Holmlin, R. E.; Barton, J. K.,
Oxidative DNA damage through long-range electron transfer. *Nature* 1996, **382** (6593), 731-5.


164. Frelon, S.; Douki, T.; Favier, A.; Cadet, J., Hydroxyl radical is not the main reactive species involved in the degradation of DNA bases by copper in the presence of hydrogen peroxide. *Chemical Research in Toxicology* 2003, **16** (2), 191-197.


168. (a) Beckman, K. B.; Ames, B. N., Oxidative decay of DNA. *The Journal of biological chemistry* 1997, **272** (32), 19633-6; (b) Cooke, M. S.; Olinski, R.; Evans, M. D., Does measurement of oxidative damage to DNA have clinical significance? *Clinica chimica acta; international journal of clinical chemistry* 2006, **365** (1-2), 30-49.


185. (a) Mijakovic, I.; Petranovic, D.; Macek, B.; Cepo, T.; Mann, M.; Davies, J.; Jensen, P. R.; Vujaklja, D., Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. Nucleic Acids Res 2006, 34 (5), 1588-96; (b) Grozav, A. G.; Willard, B. B.; Kozuki, T.; Chikamori, K.; Micluta, M. A.; Petrescu, A. J.; Kinter, M.; Ganapathi, R.; Ganapathi, M. K., Tyrosine 656 in topoisomerase IIBeta is important for the catalytic activity of the enzyme: Identification based on artifactual +80-Da modification at this site. Proteomics 2011, 11 (5), 829-42; (c) Kaneda, K.; Sekiguchi, J.; Shida, T., Role of the tryptophan residue in the


Appendices
A1. Ribonuclide A Binding to DNA. To illustrate RNAse A binding to the 27-mer model DNA, DNA migration in a non-denaturing PAGE gel by itself or with the addition of RNAse A are compared. In the absence of RNAse A (left), the 27-mer DNA is able migrate into the gel but is hampered from migrating when an equimolar concentration of the protein is added (right). This indicates that the DNA has some affinity to RNAse A, which due to its large size and positive charge is impeded from moving into the gel.
A2. Piperidine Test for DNA Damage. Piperidine is added to DNA samples, with and without irradiation as well as with rose bengal and no irradiation. No DNA strand breaks are observed for all samples indicating that DNA damage does not occur from irradiation alone. Addition of rose bengal without irradiation also did not generate any piperidine labile sites.
A3. Trypsin Digestion of DNA-Protein Cross-links. Trypsin digestion of DNA-protein cross-links followed by gel electrophoresis with ethidium bromide detection shows loss of high molecular weight product band (RB and RF Rxn lanes). The bands resulting from trypsin digestion (lanes 2 and 4; Trypsin Digested) move slower into the gel than un-reacted DNA bands (bands at the bottom of the gel), indicating that enzymatic digestion resulted to a peptide bound to DNA.
A4. DNA Strand Scission as By-Product of DNA-Protein Cross-linking. No strand scission is occurring for rose bengal oxidation as a similar pattern is observed for the unreacted DNA band (upper left gel, bottom), with and without the presence of the oxidant. Comparing the unreacted DNA bands with riboflavin as oxidant shows a streak below the bands only when the oxidant is present indicating that DNA strands shorter than the 27-mer DNA are present, which is indicative of strand scission (upper right gel). Allowing the reactions to proceed for 72 hours in the presence if Cu(II)-H₂O₂ oxidation system result to DNA band degradation, particularly when no protein is added (lower right gel) while very little DNA band degradation is detectable when Fe(II)-H₂O₂ oxidation system is used (lower left gel).
A5. Sensitivity of Cu(I)-H$_2$O$_2$-Generated DPCs to Loss of Guanine. As with rose bengal oxidation, the Cu-H$_2$O$_2$ oxidation system is also determined to be sensitive to loss of guanine from the sequence. A 3.6 +/- 0.6% yield is obtained when all 9 guanines are present in the sequence. The most dramatic yield reduction is observed when the first six guanines are removed were the yield is reduced from 3.1 +/- 0.6% when the first five guanines were replaced to 1.9 +/- 0.4% when the sixth guanine is substituted for adenine. This implies that the sixth guanine may have an important role in crosslinking for this particular sequence. The yield drops to 0.2 +/- 0.1% when all guanines are exchanged with adenine.
A6. Stability of Intact 27-Mer-Ribonuclease A Oxidative Cross-link. A crosslink reaction was purified by electrophoresis. The portion of the gel was then excised (~1cm by 1cm) and eluted into 800 μL of water. After 24 hours the solution was filtered and 20 μL was used for electrophoresis. A reaction standard is shown on the right. Though dilute no unreacted DNA or any other bands are observed.

**Figure 7 (B1)**

**Sp-N-acetyl-AAAKF • H⁺**  
\[ \text{C}_{36}\text{H}_{52}\text{N}_{11}\text{O}_{12}^+ \]  
Theor. = 830.37914  
Error = 44 ppb

**Sp-N-acetyl-AAAKF • Na⁺**  
\[ \text{C}_{36}\text{H}_{51}\text{N}_{11}\text{O}_{12}\text{Na}^+ \]  
Theor. = 852.36109  
Error = 202 ppb
Figure 7 (B2)

Sp-N-acetyl-AAAKF MS²

830 →

N-acetyl-A – A – A – K – F

y³⁺ (with sugar loss)

b⁴⁺ (with sugar loss)

y⁴⁺ (with sugar loss)

y²⁺ (with sugar loss)

- sugar

- sugar, H₂O

- sugar cleav,

- C₃H₆O₃

- H₂O

Figure 7 (B2)
<table>
<thead>
<tr>
<th>Ion Mass</th>
<th>Elemental Composition</th>
<th>Error (ppb)</th>
<th>LTQ Rel. Abund.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>714.3315</td>
<td>C₂ᵢHᵢ₄Nᵢ₆Oᵢ₉⁺</td>
<td>-349</td>
<td>100</td>
<td>sugar loss</td>
</tr>
<tr>
<td>549.2527</td>
<td>C₂ᵢHᵢ₃Nᵢ₆Oᵢ₇⁺</td>
<td>-74</td>
<td>9.3</td>
<td>bᵢ⁺ - sugar</td>
</tr>
<tr>
<td>601.2840</td>
<td>C₂ᵢHᵢ₃Nᵢ₆Oᵢ₇⁺</td>
<td>-51</td>
<td>2.1</td>
<td>yᵢ⁺ - sugar</td>
</tr>
<tr>
<td>665.3002</td>
<td>C₆H₆N₃O₁₈⁺</td>
<td>143</td>
<td>1.9</td>
<td>b₄⁺</td>
</tr>
<tr>
<td>459.2099</td>
<td>C₂ᵢH₂N₂O₈⁺</td>
<td>168</td>
<td>1.6</td>
<td>y₂⁺ - sugar</td>
</tr>
<tr>
<td>530.2468</td>
<td>C₂ᵢH₂N₂O₈⁺</td>
<td>-277</td>
<td>1.2</td>
<td>y₁⁺ - sugar</td>
</tr>
<tr>
<td>436.2052</td>
<td>C₁ᵢH₂₀N₉O₅⁺</td>
<td>271</td>
<td>1.1</td>
<td>y₁⁺ - sugar &amp; phenylalanine</td>
</tr>
<tr>
<td>696.3210</td>
<td>C₂ᵢH₁₂N₁₁O₅⁺</td>
<td>-279</td>
<td>1.0</td>
<td>sugar &amp; H₂O loss</td>
</tr>
<tr>
<td>812.3679</td>
<td>C₃ᵢH₁₀N₁₃O₃⁺</td>
<td>-811</td>
<td>1.0</td>
<td>H₂O loss</td>
</tr>
<tr>
<td>740.3475</td>
<td>C₃ᵢH₁₀N₁₄O₄⁺</td>
<td>96</td>
<td>0.8</td>
<td>(sugar cleavage)</td>
</tr>
</tbody>
</table>
Figure 7 (B3)

Sp-N-acetyl-AAAKF $\text{MS}^3$


(714 is sugar loss)

$\text{y}_4^+$ - F
$\text{y}_2^+$ - NH$_3$
$\text{y}_3^+$ - F
$\text{y}_2^+$ - NH$_3$
$\text{y}_3^+$ - F
$\text{y}_2^+$ - NH$_3$
$\text{a}_4^+$ - NH$_3$
$\text{a}_4^+$ - NH$_3$

- O=C=NC(NH$_2$)=NH - F
- O=C=NC(NH$_2$)=NH - F
- O=C=NC(NH$_2$)=NH - F
- O=C=NC(NH$_2$)=NH - F
- HN=O, H$_2$O
- HN=O, H$_2$O
- HN=O, H$_2$O
- HN=O, H$_2$O

546.25270
601.28403
629.30414
653.31520
671.32593
696.32101
714.33@cid25.00
<table>
<thead>
<tr>
<th>Ion Mass</th>
<th>Elemental Composition</th>
<th>Error (ppb)</th>
<th>LTQ Rel. Abund.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>549.2527</td>
<td>C_24H_24N_6O_3^+</td>
<td>-219</td>
<td>100</td>
<td>b_3^+</td>
</tr>
<tr>
<td>436.2050</td>
<td>C_24H_24N_6O_3^+</td>
<td>-119</td>
<td>33.4</td>
<td>y_5^+ - phenylalanine</td>
</tr>
<tr>
<td>601.2840</td>
<td>C_24H_24N_6O_3^+</td>
<td>-151</td>
<td>29.5</td>
<td>y_5^+</td>
</tr>
<tr>
<td>459.2098</td>
<td>C_24H_24N_6O_3^+</td>
<td>-93</td>
<td>25.8</td>
<td>y_5^+</td>
</tr>
<tr>
<td>530.2468</td>
<td>C_24H_24N_6O_3^+</td>
<td>-239</td>
<td>19.3</td>
<td>y_5^+</td>
</tr>
<tr>
<td>365.1679</td>
<td>C_24H_24N_6O_3^+</td>
<td>-104</td>
<td>11.7</td>
<td>y_5^+ - phenylalanine</td>
</tr>
<tr>
<td>696.3210</td>
<td>C_24H_24N_6O_3^+</td>
<td>-322</td>
<td>10.0</td>
<td>H_2O loss</td>
</tr>
<tr>
<td>671.3259</td>
<td>C_24H_24N_6O_3^+</td>
<td>-83</td>
<td>8.1</td>
<td>HNN=C-O loss</td>
</tr>
<tr>
<td>521.2578</td>
<td>C_24H_24N_6O_3^+</td>
<td>-126</td>
<td>6.7</td>
<td>a_3^+</td>
</tr>
<tr>
<td>249</td>
<td>too low for FT</td>
<td></td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td>442.1832</td>
<td>C_24H_24N_6O_3^+</td>
<td>-280</td>
<td>3.1</td>
<td>y_2^+ - NH_3</td>
</tr>
<tr>
<td>464.2251</td>
<td>C_24H_24N_6O_3^+</td>
<td>-62</td>
<td>2.6</td>
<td>m/z 629 - phenylalanine</td>
</tr>
<tr>
<td>504.2313</td>
<td>C_24H_24N_6O_3^+</td>
<td>27</td>
<td>2.5</td>
<td>a_5^+ - NH_3</td>
</tr>
<tr>
<td>629.3041</td>
<td>C_24H_24N_6O_3^+</td>
<td>-75</td>
<td>2.2</td>
<td>O=C=N-C(NH_2)=NH loss</td>
</tr>
<tr>
<td>266</td>
<td>too low for FT</td>
<td></td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>348</td>
<td>too low for FT</td>
<td></td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>558.2785</td>
<td>C_24H_24N_6O_3^+</td>
<td>458</td>
<td>1.6</td>
<td>y_2^+ - HNCO</td>
</tr>
<tr>
<td>416</td>
<td>too low for FT</td>
<td></td>
<td>1.6</td>
<td>possible y_2^+ - HNCO</td>
</tr>
<tr>
<td>653.3152</td>
<td>C_24H_24N_6O_3^+</td>
<td>-338</td>
<td>1.5</td>
<td>HNCO &amp; H_2O loss</td>
</tr>
<tr>
<td>294</td>
<td>too low for FT</td>
<td></td>
<td>1.4</td>
<td>possible y_2^+ - phenylalanine</td>
</tr>
<tr>
<td>584.2576</td>
<td>C_24H_24N_6O_3^+</td>
<td>100</td>
<td>1.3</td>
<td>y_2^+ - NH_3</td>
</tr>
<tr>
<td>583.2731</td>
<td>C_24H_24N_6O_3^+</td>
<td>-781</td>
<td>0.6</td>
<td>y_2^+ - H_2O</td>
</tr>
</tbody>
</table>

(714 is sugar loss)
Theor. = 804.39988
Error = 101 ppb

Gh-N-acetyl-AAA
C$_{35}$H$_{54}$N$_{11}$O$_{11}^+$

Theor. = 826.38182
Error = 510 ppb

Gh-N-acetyl-AAA
C$_{35}$H$_{53}$N$_{11}$O$_{11}$Na$^+$

Figure 7 (C1)
Figure 7 (C2)

**Gh-N-acetyl-AAA\textsubscript{K}F**

\[ 804 \rightarrow \]

**MS\textsuperscript{2}**
<table>
<thead>
<tr>
<th>Ion Mass</th>
<th>Elemental Composition</th>
<th>Error (ppb)</th>
<th>LTQ Rel. Abund.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>549.3032</td>
<td>$\text{C}<em>{26}\text{H}</em>{41}\text{N}<em>{6}\text{O}</em>{7}^+$</td>
<td>192</td>
<td>100</td>
<td>protonated peptide</td>
</tr>
<tr>
<td>628.3204</td>
<td>$\text{C}<em>{29}\text{H}</em>{42}\text{N}<em>{9}\text{O}</em>{7}^+$</td>
<td>411</td>
<td>14.1</td>
<td>sugar &amp; ((NH$_3$ + HNCO) or urea) loss</td>
</tr>
<tr>
<td>745.3517</td>
<td>$\text{C}<em>{34}\text{H}</em>{49}\text{N}<em>{8}\text{O}</em>{11}^+$</td>
<td>333</td>
<td>7.6</td>
<td>HN=C(NH$_2$)$_2$ loss</td>
</tr>
<tr>
<td>719.3725</td>
<td>$\text{C}<em>{35}\text{H}</em>{53}\text{N}<em>{10}\text{O}</em>{10}^+$</td>
<td>338</td>
<td>6.4</td>
<td>O=C=N-C(NH$_2$)$_2$ loss</td>
</tr>
<tr>
<td>671.3262</td>
<td>$\text{C}<em>{30}\text{H}</em>{43}\text{N}<em>{10}\text{O}</em>{8}^+$</td>
<td>335</td>
<td>1.0</td>
<td>sugar &amp; NH$_3$ loss</td>
</tr>
<tr>
<td>761.3942</td>
<td>$\text{C}<em>{34}\text{H}</em>{53}\text{N}<em>{10}\text{O}</em>{10}^+$</td>
<td>243</td>
<td>0.9</td>
<td>HN=C=O loss</td>
</tr>
<tr>
<td>787.3736</td>
<td>$\text{C}<em>{35}\text{H}</em>{51}\text{N}<em>{10}\text{O}</em>{11}^+$</td>
<td>382</td>
<td>0.5</td>
<td>NH$_3$ loss</td>
</tr>
<tr>
<td>688.3527</td>
<td>$\text{C}<em>{30}\text{H}</em>{46}\text{N}<em>{11}\text{O}</em>{8}^+$</td>
<td>371</td>
<td>0.4</td>
<td>sugar loss</td>
</tr>
</tbody>
</table>
Gh-N-acetyl-AAAKF \text{MS}^3
804 \rightarrow 745
(745 is \text{HN=C(NH}_2)_2 \text{ loss})

Very low - noise observed as false ions at m/z 494.73, 537.93, 617.98, 683.92, 704.72, etc.

Figure 7 (C3)
<table>
<thead>
<tr>
<th>Ion Mass</th>
<th>Elemental Composition</th>
<th>Error (ppb)</th>
<th>LTQ Rel. Abund.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>580.2727</td>
<td>C$<em>{23}$H$</em>{36}$N$<em>{6}$O$</em>{6}$</td>
<td>305</td>
<td>100</td>
<td>b$_4^+$</td>
</tr>
<tr>
<td>490.2299</td>
<td>C$<em>{23}$H$</em>{32}$N$<em>{6}$O$</em>{6}$</td>
<td>662</td>
<td>27.3</td>
<td>y$_2^+$</td>
</tr>
<tr>
<td>632.3043</td>
<td>C$<em>{29}$H$</em>{42}$N$<em>{6}$O$</em>{6}$</td>
<td>739</td>
<td>25.9</td>
<td>y$_4^+$</td>
</tr>
<tr>
<td>629.3044</td>
<td>C$<em>{29}$H$</em>{41}$N$<em>{6}$O$</em>{6}$</td>
<td>417</td>
<td>25.5</td>
<td>sugar loss</td>
</tr>
<tr>
<td>561.2672</td>
<td>C$<em>{26}$H$</em>{37}$N$<em>{6}$O$</em>{6}$</td>
<td>910</td>
<td>23.6</td>
<td>y$_3^+$</td>
</tr>
<tr>
<td>464</td>
<td>too low for FT</td>
<td>21.4</td>
<td>possible b$_4^+$ - sugar</td>
<td></td>
</tr>
<tr>
<td>727.3419</td>
<td>C$<em>{34}$H$</em>{49}$N$<em>{6}$O$</em>{10}$</td>
<td>1283</td>
<td>17.7</td>
<td>H$_2$O loss</td>
</tr>
<tr>
<td>516</td>
<td>too low for FT</td>
<td>8.4</td>
<td>possible y$_4^+$ - sugar</td>
<td></td>
</tr>
<tr>
<td>374</td>
<td>too low for FT</td>
<td>5.9</td>
<td>possible y$_3^+$ - sugar</td>
<td></td>
</tr>
<tr>
<td>445</td>
<td>too low for FT</td>
<td>2.1</td>
<td>possible y$_3^+$ - sugar</td>
<td></td>
</tr>
</tbody>
</table>
Appendix A8. High Resolution Mass Spectral Analysis of N-Ac-AYKTT-dG Adduct (for figure 3.8). For methods, see materials and methods section of Chapter 3.

Figure 8 (B)

- dG(ox)-N-acetyl-AYKTT • H⁺
  - C₃₈H₅₆N₁₁O₁₅⁺
  - Error = -96 ppb

- Gh-N-acetyl-AYKTT • H⁺
  - C₃₇H₅₈N₁₁O₁₄⁺
  - Error = -105 ppb

- dG(ox)-N-acetyl-AYKTT • Na⁺
  - C₃₈H₅₆N₁₁O₁₅Na⁺
  - Error = 116 ppb
Figure 8 (C)

Gh-N-acetyl-AYKTT MS²

- sugar, NH₃
- sugar, HN=C=O, H₂O
- C₂H₄O₃
(sugar cleav.)
- sugar, NH₃ or - sugar, urea
- sugar, HN=C=NH
- sugar, HN=C=O
- sugar, HN=C(NH₂)₂
- sugar

AYKTT • H⁺

625.31929
679.34105
705.32038
721.36283
747.34219
764.36861
790.38434
809.35791
826.38432
844.39521
862.40564
880.4200

- H₂O
- 3H₂O
- HN=C=O
- HN=C=O, 3H₂O
- C₂H₄O₃
(sugar cleav.)
<table>
<thead>
<tr>
<th>Ion Mass</th>
<th>Elemental Composition</th>
<th>Error (ppb)</th>
<th>LTQ Rel. Abund.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>764.3686</td>
<td>C_{32}H_{50}N_{11}O_{11}^+</td>
<td>41</td>
<td>100</td>
<td>sugar loss</td>
</tr>
<tr>
<td>747.3421</td>
<td>C_{32}H_{49}N_{10}O_{11}^+</td>
<td>215</td>
<td>40.0</td>
<td>sugar &amp; NH_3 loss</td>
</tr>
<tr>
<td>837.4101</td>
<td>C_{30}H_{49}N_{10}O_{13}^+</td>
<td>97</td>
<td>24.5</td>
<td>HN=C=O loss</td>
</tr>
<tr>
<td>826.3843</td>
<td>C_{31}H_{52}N_{11}O_{11}^+</td>
<td>111</td>
<td>13.6</td>
<td>3H_2O loss</td>
</tr>
<tr>
<td>862.4056</td>
<td>C_{31}H_{50}N_{10}O_{13}^+</td>
<td>327</td>
<td>10.7</td>
<td>H_2O loss</td>
</tr>
<tr>
<td>809.3579</td>
<td>C_{31}H_{49}N_{10}O_{11}^+</td>
<td>285</td>
<td>10.0</td>
<td>3H_2O &amp; NH_3 loss</td>
</tr>
<tr>
<td>625.3192</td>
<td>C_{28}H_{45}N_{9}O_{10}^+</td>
<td>194</td>
<td>8.3</td>
<td>AYKTT • H^+</td>
</tr>
<tr>
<td>527.2361</td>
<td>C_{24}H_{31}N_{8}O_{6}^+</td>
<td>137</td>
<td>0.8</td>
<td>b_3^+ with sugar loss</td>
</tr>
<tr>
<td>444.2630</td>
<td>C_{24}H_{34}N_{10}O_{6}^+</td>
<td>649</td>
<td>0.6</td>
<td>b_3^+ - NH_3 with sugar loss</td>
</tr>
<tr>
<td>586.2619</td>
<td>C_{27}H_{38}N_{8}O_{6}^+</td>
<td>-31</td>
<td>0.5</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 8 (C)**

Gh-N-acetyl-AYKTT MS²
Figure 8 (D1)

dG(ox)-N-acetyl-AYKTT MS²

906 →

- sugar
- sugar, H₂O
- sugar, H₂O
- C₃H₆O₃
- sugar cleav.
- 3H₂O
- H₂O
- 2H₂O

b₄⁺ - 3H₂O
b₃⁺
(b with sugar loss)

b₂⁺
(with sugar loss)

570.24200

790.34780

m/z

Relative Abundance

790.34789
888.38470
852.36364
816.36375
852.36364
888.38470
570.24200
722.34684
773.32128
125.00
245.00
<table>
<thead>
<tr>
<th>Ion Mass</th>
<th>Elemental Composition</th>
<th>Error (ppb)</th>
<th>LTQ Rel. Abund.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>790.3478</td>
<td>C&lt;sub&gt;19&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;11&lt;/sub&gt;O&lt;sub&gt;12&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>59</td>
<td>100</td>
<td>sugar loss</td>
</tr>
<tr>
<td>888.3847</td>
<td>C&lt;sub&gt;19&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;11&lt;/sub&gt;O&lt;sub&gt;14&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>87</td>
<td>5.2</td>
<td>H&lt;sub&gt;2&lt;/sub&gt;O loss</td>
</tr>
<tr>
<td>852.3636</td>
<td>C&lt;sub&gt;26&lt;/sub&gt;H&lt;sub&gt;30&lt;/sub&gt;N&lt;sub&gt;11&lt;/sub&gt;O&lt;sub&gt;12&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>172</td>
<td>3.6</td>
<td>3H&lt;sub&gt;2&lt;/sub&gt;O loss</td>
</tr>
<tr>
<td>671.2896</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;10&lt;/sub&gt;O&lt;sub&gt;8&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>15</td>
<td>1.8</td>
<td>b&lt;sub&gt;1&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt; with sugar loss</td>
</tr>
<tr>
<td>870.3742</td>
<td>C&lt;sub&gt;26&lt;/sub&gt;H&lt;sub&gt;30&lt;/sub&gt;N&lt;sub&gt;11&lt;/sub&gt;O&lt;sub&gt;13&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>209</td>
<td>1.8</td>
<td>2H&lt;sub&gt;2&lt;/sub&gt;O loss</td>
</tr>
<tr>
<td>816.3637</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;11&lt;/sub&gt;O&lt;sub&gt;12&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>315</td>
<td>1.4</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;2&lt;/sub&gt;O loss (sugar cleavage)</td>
</tr>
<tr>
<td>863.3895</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;10&lt;/sub&gt;O&lt;sub&gt;8&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>216</td>
<td>1.4</td>
<td>HN=C=O loss</td>
</tr>
<tr>
<td>773.3212</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;10&lt;/sub&gt;O&lt;sub&gt;14&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>-18</td>
<td>1.2</td>
<td>sugar &amp; NH&lt;sub&gt;3&lt;/sub&gt; loss</td>
</tr>
<tr>
<td>772.3375</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;11&lt;/sub&gt;O&lt;sub&gt;11&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>313</td>
<td>1.2</td>
<td>sugar &amp; H&lt;sub&gt;2&lt;/sub&gt;O loss</td>
</tr>
<tr>
<td>722.3468</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;10&lt;/sub&gt;O&lt;sub&gt;8&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>82</td>
<td>1.0</td>
<td>net C&lt;sub&gt;5&lt;/sub&gt;H&lt;sub&gt;10&lt;/sub&gt;O&lt;sub&gt;2&lt;/sub&gt; loss</td>
</tr>
<tr>
<td>570.2420</td>
<td>C&lt;sub&gt;19&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;11&lt;/sub&gt;O&lt;sub&gt;7&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>138</td>
<td>1.0</td>
<td>b&lt;sub&gt;1&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt; with sugar loss</td>
</tr>
<tr>
<td>747.3422</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;10&lt;/sub&gt;O&lt;sub&gt;11&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>242</td>
<td>1.0</td>
<td>sugar &amp; HNCO loss</td>
</tr>
<tr>
<td>733.3053</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;10&lt;/sub&gt;O&lt;sub&gt;8&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>123</td>
<td>1.0</td>
<td>b&lt;sub&gt;1&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt; - 3H&lt;sub&gt;2&lt;/sub&gt;O</td>
</tr>
<tr>
<td>787.3370</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;11&lt;/sub&gt;O&lt;sub&gt;13&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>147</td>
<td>0.6</td>
<td>b&lt;sub&gt;4&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>686.2893</td>
<td>C&lt;sub&gt;20&lt;/sub&gt;H&lt;sub&gt;26&lt;/sub&gt;N&lt;sub&gt;10&lt;/sub&gt;O&lt;sub&gt;10&lt;/sub&gt; &lt;sup&gt;+&lt;/sup&gt;</td>
<td>94</td>
<td>0.3</td>
<td>b&lt;sub&gt;3&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>442</td>
<td>too low for FT</td>
<td>0.2</td>
<td>0.2</td>
<td>b&lt;sub&gt;2&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt; (Sp on Y with sugar loss)</td>
</tr>
<tr>
<td>677</td>
<td>too low for FT</td>
<td>0.1</td>
<td></td>
<td>y&lt;sub&gt;B&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt; with sugar loss</td>
</tr>
<tr>
<td>558</td>
<td>too low for FT</td>
<td>tr.</td>
<td></td>
<td>b&lt;sub&gt;1&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt; (Sp on Y)</td>
</tr>
<tr>
<td>349</td>
<td>too low for FT</td>
<td>tr.</td>
<td></td>
<td>y&lt;sub&gt;B&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt; (Sp on Y with or without sugar loss)</td>
</tr>
<tr>
<td>331</td>
<td>too low for FT</td>
<td>tr.</td>
<td></td>
<td>y&lt;sub&gt;2&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt; - H&lt;sub&gt;2&lt;/sub&gt;O (Sp on Y with or without sugar loss)</td>
</tr>
</tbody>
</table>
Solivio_16-20O_20111006-R01 #249-998  RT: 2.71-11.25  AV: 37  NL: 1.15E4
F: FTMES + p ESI Full ms3 906.40@cid25.00  790.35@cid25.00 [215.00-915.00]

Figure 8 (D2)

**dG(ox)-N-acetyl-AKYTT MS³**

906 → 790

(790 is sugar loss)

**b⁴⁺**

- 627/627/628
- b⁴⁺ - CH₂=CHOH
- b⁴⁺ - CO₂
- b⁴⁺ - HΝ=C=O

**b₂⁺**

- y⁴⁺ - TT

**b₃⁺**

- 570.24194
- 628.26396
- 635.26854
- 652.23142

**b⁴⁺**

- O=C=N-CN
- H₂O

**b⁴⁺ - H₂O**

- 2H₂O
- CO₂

**y⁴⁺**

- O=C=N-CN
- H₂O, CO₂
- H₂O, CO₂
- -2H₂O
- HΝ=C=O

- CH₂=CHOH
- CO₂
- H₂O, CH₂=CHOH
- b⁴⁺ - 2H₂O
- b⁴⁺ - H₂O

**b⁴⁺ - H₂O, CH₂=CHOH**

671.28659
<table>
<thead>
<tr>
<th>Ion Mass</th>
<th>Elemental Composition</th>
<th>Error (ppb)</th>
<th>LTQ Rel. Abund.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>671.2895</td>
<td>C_{20}H_{29}N_{10}O_{9}^+</td>
<td>-15</td>
<td>100</td>
<td>b_4^-</td>
</tr>
<tr>
<td>772.3374</td>
<td>C_{32}H_{49}N_{11}O_{11}^+</td>
<td>157</td>
<td>47.0</td>
<td>H_2O loss</td>
</tr>
<tr>
<td>570.2419</td>
<td>C_{22}H_{26}N_{9}O_{9}^-</td>
<td>33</td>
<td>45.0</td>
<td>b_4^-</td>
</tr>
<tr>
<td>457.1942</td>
<td>C_{22}H_{26}N_{9}O_{9}^-</td>
<td>-6</td>
<td>44.7</td>
<td>y_4^- with TT (C_8H_{16}N_{2}O_5) loss</td>
</tr>
<tr>
<td>722.3468</td>
<td>C_{31}H_{48}N_{11}O_{11}^+</td>
<td>27</td>
<td>40.8</td>
<td>O=C=N-CN loss</td>
</tr>
<tr>
<td>635.2685</td>
<td>C_{26}H_{34}N_{14}O_{10}^-</td>
<td>109</td>
<td>39.2</td>
<td>b_4^- - H_2O</td>
</tr>
<tr>
<td>442.1469</td>
<td>C_{13}H_{20}N_{7}O_{6}^-</td>
<td>-64</td>
<td>25.2</td>
<td>b_4^- (Sp on Y)</td>
</tr>
<tr>
<td>754.3268</td>
<td>C_{33}H_{44}N_{11}O_{10}^-</td>
<td>128</td>
<td>16.8</td>
<td>2H_2O loss</td>
</tr>
<tr>
<td>653.2790</td>
<td>C_{29}H_{37}N_{10}O_{8}^-</td>
<td>-8</td>
<td>15.7</td>
<td>b_4^- - H_2O</td>
</tr>
<tr>
<td>728.3475</td>
<td>C_{32}H_{46}N_{11}O_{9}^-</td>
<td>70</td>
<td>14.6</td>
<td>H_2O &amp; CO_2 loss</td>
</tr>
<tr>
<td>677.3002</td>
<td>C_{24}H_{34}N_{10}O_{10}^-</td>
<td>111</td>
<td>14.0</td>
<td>y_4^-</td>
</tr>
<tr>
<td>609.2529</td>
<td>C_{25}H_{38}N_{10}O_{9}^-</td>
<td>147</td>
<td>13.1</td>
<td>b_4^- with H_2O &amp; CH_2=CH=O loss</td>
</tr>
<tr>
<td>747.3421</td>
<td>C_{33}H_{46}N_{11}O_{10}^-</td>
<td>201</td>
<td>10.6</td>
<td>HN=C=O loss</td>
</tr>
<tr>
<td>301.1043</td>
<td>C_{13}H_{20}N_{7}O_{3}^-</td>
<td>-17</td>
<td>9.7</td>
<td>Sp(no sugar)-O-C=CH=CH=NH_2^+</td>
</tr>
<tr>
<td>710.3369</td>
<td>C_{20}H_{29}N_{10}O_{9}^-</td>
<td>149</td>
<td>6.1</td>
<td>2H_2O &amp; CO_2 loss</td>
</tr>
<tr>
<td>627.2634</td>
<td>C_{27}H_{33}N_{10}O_{9}^-</td>
<td>151</td>
<td>5.6^-</td>
<td>b_4^- with CH_2=CH=O loss</td>
</tr>
<tr>
<td>627.2996</td>
<td>C_{26}H_{33}N_{10}O_{9}^-</td>
<td>-128</td>
<td>5.6^-</td>
<td>b_4^- with CO_2 loss</td>
</tr>
<tr>
<td>628.2839</td>
<td>C_{24}H_{32}N_{9}O_{8}^-</td>
<td>261</td>
<td>5.4</td>
<td>b_4^- with HN=C=O loss</td>
</tr>
<tr>
<td>329.0992</td>
<td>C_{13}H_{17}N_{9}O_{8}^-</td>
<td>-60</td>
<td>4.2</td>
<td>Sp(no sugar)-O-C=CH=CH=NH_2^+</td>
</tr>
</tbody>
</table>

* = merged together due to LTQ’s lower resolution

Figure 8 (D2)

dG(ox)-N-acetyl-AYKTT MS^3

906 —— 790 ❍

(790 is sugar loss)
Appendix A9. Absorbance and Mass Spectra for the Three dG Hyper-Oxidation Products Shown in Figure 4.2 B. The absorbance profiles for the 1.7, 2.4 and 5.9 minutes peaks in figure 4.2B are shown (A). The profiles for the 1.7 and 2.4 minute peaks match the known profiles for spiroiminodihydantoin (Sp), while the absorbance profile for the 5.9 minute peak matches the known profile for oxazolone (Z). The structures, elemental compositions and exact masses for each adduct are shown in B. Mass spectral analysis (C) confirms adduct identity.

**A**

- Peak C; 1.7 mins
- Peak C; 2.4 mins
- Peak D

**B**

- dSp
- dZ

Chemical Formula: C_{10}H_{14}N_{5}O_{6}^+

Exact Mass: 300.25045

Chemical Formula: C_{8}H_{15}N_{4}O_{5}^+

Exact Mass: 247.10370

**C.**

- MS^*_{1.7m} 300 MH+
- 184 M-dRH+

**D.**

- MS^*_{5.9m} 247 MH+
- 131 M-dRH+
- 247 M-CO_2H+
- 203
Appendix A10. Hypercarb Column Chromatograms for dG Hyper-Oxidation Products when Irradiated with Rose Bengal at 560 nm (black trace) and 375 nm (grey trace).
Appendix A11. Peptides Containing Cys and Arg Form Adducts with the dsDNA Model but at a Much Lower Yield Compared to Tyr Containing Peptides. A peptide model N-Ac-ASXF (where X is Leu, Arg, Lys, Tyr or Cys) was tested to determine adduct formation with the dsDNA model. From previous work, the peptide N-Ac-AKYTT is determined to generate high yields of cross-links with the DNA model and used as a positive control. The gel image (20% PAGE) shows reactions and controls for the peptides containing Tyr, Cys, Leu and Arg. Leu does not possess a cross-link active side-chain and is used as a negative control. Adducts are identified as a band unique to the reaction and are denoted by red arrows. The peptides containing Cys and Arg are able to form adducts with DNA but at a much lower yield compared to Tyr containing peptides.
Appendix A11. Mass Spectral Analysis of 3.8 Minute Peak in Figure 4.6. The unique peak at 3.8 minutes in the reaction in the absence of the antioxidant N-Acetyl-cysteine is analyzed by mass spectrometry showing an ion with m/z 484, corresponding to the protonated mass of the dSp-Lys adduct. Other ions detected have m/z of 203 and 368, which corresponds to the protonated masses of N-Ac-Lys-OMe (Adduct minus dSp) and Sp-Lys (adduct minus deoxyribose), respectively. MS/MS analysis of the 484 peak shows a fragment with m/z 368, which corresponds to the mass of the adduct, minus deoxyribose, confirming that the molecule is a modified nucleoside.