I. DNA-PROTEIN CROSS-LINKING BY *CIS*-1,1,2,2-DIAMMINEDICHLOROPLATINUM (II) (CISPLATIN)

II. FORMATION OF 8-OXO-dG AND OXAZOLONE LESIONS WITH *p53* DERIVED DNA SEQUENCES FOLLOWING PHOTOOXIDATION IN THE PRESENCE OF RIBOFLAVIN

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Xun Ming

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Dr. Natalia Y. Tretyakova, Advisor

May 2011

Acknowledgements

I would first like to thank my thesis advisor, Prof. Natalia, Y. Tretyakova, for all of her support, encouragement, especially through the difficult times when things did not work out as expected. I have learned a great deal from her, and I absolutely feel that I am so lucky to have an opportunity to work in her lab.

I would also express my gratitude to my thesis committee, Dr. Stephen Hecht and my co-advisor Dr. Colin Campbell for their guidance, support, and valuable feedback throughout the past three years.

I also wish to thank those who directly contributed to the development of the work in this thesis, especially Erin Michaelson-Richie, Dr. Rachel Loeber, and Brock Matter, whose hard work and dedication laid the groundwork for this thesis. I would also like to acknowledge Matthew Song (College of Pharmacy, University of Minnesota) for his invaluable contributions to the DNA oxidation work.

I would particularly like to acknowledge Dr. Peter Villalta (Mass Spectrometry Core at the Masonic Cancer Center), Dr. Matthew Stone, Dr. LeeAnn Higgins, Dr. Lori Anderson (Center for Mass Spectrometry and Proteomics), Pratik Jagtap (Minnesota Supercomputing Institute), Jason Kuchar (Thermo Fisher Scientific), and Dr. Linda Von Weymarn (Masonic Cancer Center) for providing training and assistance with HPLC, LC-MS, MS-based proteomics, and proteomic database search analysis.

I am also grateful to the former and current members in the Tretyakova lab for providing friendship, moral support, technical support, and valuable discussions, especially Dr. Melissa Goggin, Dr. Bhaskar Malayappan, Srikanth Kotapati, Dewakar Sangaraju, Cate Tsufis, B.J. Valenta, Armanda Reigel and Susith Wickramaratne.

Finally, I wish to thank my family and friends, especially my wife, Yanrong, and my parents, Yong and Xiaoling, for their love and encouragement. I cannot finish this thesis without their support.

Dedication

This thesis was written in loving memory of my grandfather, Jixian Ming.

You will forever remain in my heart.

(1926 - 2010)

Abstract

1,1,2,2- *Cis*-diamminedichloroplatinum (II) (cisplatin) is a common anticancer drug used in treatment of solid tumors. The biological activity of cisplatin is generally attributed to its ability to form DNA-DNA cross-links by sequential platination of two nucleophilic sites within the DNA duplex. However, cisplatin also forms a variety of DNA-protein cross-links (DPCs) whose structures are not well characterized. While the biological outcomes of cisplatin-induced DPC lesions are not understood, they are hypothesized to interrupt important cellular processes such as DNA replication and transcription, potentially leading to toxicity. In the present work, a human DNA repair protein O^6-alkylguanine DNA alkyltransferase (AGT) was used as a model to investigate cisplatin-induced DNA-protein cross-linking. The normal physiological function of AGT is to repair alkylation damage by transferring O^6-alkylguanine groups from DNA to an active site cysteine residue of the protein (Cys^{145}), restoring normal guanine. Incubation of recombinant AGT protein with ^{32}P-labeled oligonucleotides duplexes in the presence of cisplatin resulted in concentration-dependent formation of DNA-protein conjugates as revealed by denaturing gel electrophoresis. Capillary HPLC-electrospray ionization mass spectrometry analysis (ESI-MS) of AGT protein treated with dG-Pt-Cl monoadduct as a model of monoplatinated DNA confirmed the ability of the protein to form multiple dG-AGT cross-links. Upon heating, dG-Pt-AGT complexes undergo platination migration from protein to the N7 position of guanine to form dG-Pt-dG cross-links. This can be explained by greater thermodynamic stability of the Pt-N bond as compared to the Pt-S bond. HPLC-ESI^+-MS/MS sequencing of tryptic peptides derived from dG-Pt-AGT complexes revealed that cisplatin-mediated cross-linking involves six different sites within this protein: Glu^{110}, Lys^{125}, Cys^{145}, His^{146}, Arg^{147}, and Cys^{150}. Among these, Cys^{145}, His^{146}, Arg^{147}, and Cys^{150} are located in the protein active site and directly or indirectly participate in alkyl transfer. Finally, HPLC-ESI^+-

MS/MS analysis of total proteolytic digests detected 1,1-*cis*-diammine-2-(5-amino-5-carboxypentyl)amino-2-(2'-deoxyguanosine-7-yl)-platinum (II) (dG-Pt-Lys) conjugates produced *via* platination of lysine residues within AGT.

To identify protein targets of cisplatin-induced cross-linking in nuclear protein extracts from human cervical carcinoma (HeLa) cells, an affinity capture methodology was combined with mass spectrometry-based proteomics and immunological detection. A total of 131 nuclear proteins were identified to form covalent DPCs in the presence of cisplatin. An estimated DNA-protein cross-linking efficiency following treatment with 50 μM cisplatin was 2-16%, depending on protein identity. HPLC-ESI$^+$-MS/MS analysis of total proteolytic digests of cross-linked proteins revealed the presence of dG-Pt-Lys conjugates. We further extended this work to characterize DNA-protein cross-linking by cisplatin in human fibrosarcoma (HT1080) cells. Following drug treatment, DPCs were isolated by a modified phenol/chloroform DNA extraction incorporating proteasome inhibitors. Proteins were released from DNA by heating and identified by mass spectrometry-based proteomics and immunological detection. Over 250 nuclear proteins were found to be captured on chromosomal DNA following treatment with cisplatin. HPLC-ESI$^+$-MS/MS analysis of total proteolytic digests revealed the formation of dG-Pt-Lys conjugates between the N7 guanine of DNA and the ε-amino group of lysine. Although cisplatin-induced DPCs spontaneously release proteins to form DNA-DNA cross-links upon heating, they appear to be stable enough under physiological conditions to inhibit DNA replication and transcription, contributing to the biological effects of cisplatin. These results indicate that clinically relevant concentrations of cisplatin induce covalent cross-links between chromosomal DNA and a large range of nuclear proteins. If not repaired, the resulting bulky DPC lesions are likely to contribute to both on-target and off-target toxicity of cisplatin.

Reactive oxygen species produced as part of normal cellular metabolism and immune response can damage cellular DNA, giving rise to promutagenic nucleobase lesions. Since the

biological impact of a given oxidative adduct is influenced by its position within gene sequence, previous studies have focused on determining the distribution of oxidative lesions along DNA sequences. However, since these studies have relied on gel electrophoresis to locate the sites of oxidative damage, they could not analyze the distribution of structurally defined nucleobase lesions and suffered from the high background of direct strand breaks induced by sugar oxidation. We now report the use of stable isotope labeling of DNA-mass spectrometry (ILD-MS) approach to map the formation of 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxo-dG) and 2,2-diamino-4-[2-deoxy-*β*-D-*erythro*-pentofuranosyl)amino]-2,5-dihydrooxazol-5-one (oxazolone) lesions along DNA sequences derived from the *p53* tumor suppressor gene. In each duplex, one of the guanine bases was labeled with [1,7,NH_2-$^{15}N_3$-2-^{13}C]-guanine which served as an isotope "tag" to enable specific quantification of guanine lesions originating from that position. Following photooxidation in the presence of riboflavin, DNA was enzymatically digested to 2'-deoxynucleosides, and the formation of 8-oxo-dG and oxazolone at each site of interest was quantified from isotope ratios obtained from capillary HPLC-ESI$^+$-MS/MS. We found that in double stranded DNA, both oxidative lesions were generated non-randomly, and their distribution was strongly influenced by the local DNA sequence. In particular, the 5' Gs in guanine repeats and guanines within ^{Me}CG dinucleotides were preferentially targeted for photooxidation in the presence of riboflavin. This can be explained by the low ionization potential of 5'-guanine at 5'-**G**G sites and/or the preferential intercalation of riboflavin at ^{Me}CG sites. Furthermore, the most frequently adducted position, $\mathbf{G_5}$ in exon 5, $\mathbf{G_4}$ and $\mathbf{G_7}$ in exon7, and $\mathbf{G_6}$ in exon 8, coincide with the known *p53* lung cancer mutational "hotspots" at *p53* codons 158 (*C**G**C*), 245 (***G****GC*), 248 (*C**G**G*), and 273 (*C**G**T*), respectively, suggesting that oxidative DNA damage may contribute to mutagenesis in the *p53* gene.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

AGT	O^6-alkylguanine DNA alkyltransferase
Arg (or R)	arginine
BD	1,3-butadiene
BSA	bovine serum albumin
CHO	Chinese hamster ovary cell line
CID	collision induced dissociation
Cisplatin	1,1,2,2-*cis*-diamminodichloroplatinum (II)
Cys (or C)	cysteine
dA	deoxyadenosine
DEB	1,2,3,4-diepoxybutane
DFO	desferrioxamine
dG	2'-deoxyguanosine
dG-Pt-AGT	dG-AGT cross-links
dG-Pt-Cl	*cis*-1,1-diammine-2-chloro-2-(2'-deoxyguanosine-7-yl) platinum (II)
dG-Pt-dG	*cis*-1,1-diammine-2,2-bis-(deoxyguanosine-7-yl)-platinum (II)
dG-Pt-Lys	*cis*-1,1-diammine-2-(5-amino-5-carboxypentyl)amino-2-(2'-deoxyguanosine-7-yl)-platinum(II)
DMEM	Dulbecco's Modified Eagle's Medium
DNase I	deoxyribonuclease I
DPCs	DNA-protein cross-links
dT	2'-deoxythymidine

DTT	dithiothreitol
EF-1α1	elongation factor 1α1
EMSA	electrophoretic mobility shift assay
ESI	electrospray ionization
FA	Formaldehyde
FapyG	2,6-diamino-4-hydroxyl-5-formamidopyrimidine
FBS	fetal bovine serum
FDR	false discovery rate
Fen-1	flap endonuclease 1
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GC	gas chromatography
Glu (or E)	glutamic acid
GO	Gene Ontology
GSH	glutathione
Gua	guanine
Hela	human cervical carcinoma cell line
His (or H)	histidine
HMG	high mobility group protein
hMutS	human mismatch repair protein
HPLC-ESI$^+$-MS/MS	high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry
HR	homologous recombination
HSAB	Hard Soft Acid Base

HT1080	human fibrosarcoma cell line
ILD-MS	stable isotope labeling of DNA-mass spectrometry
IP	ionization potential
Ku	ATP-dependent DNA helicase subunit 2
Lys (or K)	lysine
Mechlorethamine	*bis*(2-chloroethyl)methylamine
MS	mass spectrometry
MS/MS	tandem mass spectrometry
NER	nucleotide excision repair
Oxazolone	2,2-diamino-4-[2-deoxy-*β*-D-*erythro*-pentofuranosyl)amino]-2,5-dihydrooxazol-5-one
8-Oxo-dG	8-oxo-7,8-dihydro-2'-deoxyguanosine
PARP	poly(ADP-ribose) polymerase
PBS	phosphate-buffered saline
PDE I	phosphodiesterase I
PDE II	phosphodiesterase II
PMSF	phenylmethanesulfonyl fluoride
Pol	polymerase
Ref-1	DNA-(apurinic- or apyrimidinic-site) lyase
RNS	reactive nitrogen species
ROS	reactive oxygen species
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis

Ser (or S)	serine
sgh	*S*-guanosyl-L-homocystine
SPE	solid phase extraction
SRM	selected reaction monitoring
TBS	Tris-buffered saline
TIC	total ion current
UV	ultraviolet
XRCC-1	X-ray repair cross complementing protein 1

I. Introduction

1.1 Overview of DNA-protein Cross-linking

DNA-protein cross-links (DPCs) are common DNA lesions created when proteins become irreversibly trapped on DNA as a result of exposure of cells to a variety of cytotoxic, mutagenic, and carcinogenic agents (*1*). DPCs are bulky, structurally complex DNA lesions that are expected to interfere with normal DNA-protein interactions and interrupt DNA metabolic processes such as DNA replication, repair, recombination, transcription, chromatin remodeling, etc. Despite their potential threat to living cells, our understanding of the biological effects of DPCs (i.e., cytotoxicity, mutagenesis and carcinogenesis) is incomplete, and their identities as well as atomic connectivities within DPC structures have not been well characterized. This may be because cross-linking agents generate a variety of other DNA lesions in addition to DPCs (*e.g.* DNA-DNA cross-links and monoadducts). Furthermore, different cross-linking agents induce DPCs by distinct mechanisms, leading to a variety of DNA-protein cross-linking chemistries. Therefore, a comprehensive understanding of the mechanisms of DPCs formation, repair, and their biological consequences is critical to fully realize the biological consequences of these unusual DNA lesions.

1.1.1 Types of DPCs

DPCs may be broadly classified into four groups by the DNA and protein chemistries near the cross-link and by the source of the cross-link *(2;3)*. Type 1 DPCs occur when topoisomerase I (topo I) becomes trapped at the 3'-end of a DNA single strand nick in the form of a tyrosine-phosphodiester linkage between DNA and protein. Topo I protein covalently cross-linked to the 3'-end of DNA can be removed by the tyrosine-DNA phosphodiesterase Tdp1 *(4;5)*.

Similarly, Type 2 DPCs involve the cross-linking of topo II protein to 5'-ends of a DNA double-strand break *(2;3)*. It is also reported that Type 2 DPCs took place when Spo11, a meiotic

protein, transiently cross-linked to the 5'-end of a DNA double-strand break as a result of Spo 11-mediated meiosis *(2)*. Type 2 DPCs are removed by an endonuclease that incises several nucleotides from the cross-link site *(6)*.

The formation of Type 3 DPCs involves base excision repair where 2-deoxyribonolactone is generated through the cleavage of an oxidized abasic site by DNA-(apurinic- or apyrimidinic-site) lyase (Ref-1) *(2;3)*. The resulting 2-deoxyribonolactone can become cross-linked to repair proteins via an amide linkage at the 5' end of the DNA strand *(7)*.

The fourth class of DPCs, which is also the main focus of this thesis, is probably the most common and includes DPC lesions induced by the reactions with chemical and physical agents endogenously and/or exogenously *(2;3)*. Different agents induce Type 4 DPCs by different mechanisms, leading to the formation of cross-links with a broad range of distinct chemical structures, physical conformations, and cellular consequences *(1)*. The resulting diversity in the characteristics of the Type 4 DPCs makes it difficult to fully realize the structures and biological effects of DPCs.

1.1.2 DPC Formation

Type 4 DPCs can be induced following exposure to a wide variety of cytotoxic, mutagenic, and carcinogenic agents, including environmental pollutants, industrial chemicals, chemotherapeutical drugs, physical agents such as IR or UV irradiation, and physiological metabolites *(1)*.

Among all these DPC-inducing agents, formaldehyde (FA) is perhaps the one that has been most extensively studied. FA is a genotoxic chemical widely used in many consumer products, including household chemicals, and pharmaceutical and industrial products, leading to widespread human exposure to FA in industrial and household environments *(8;9)*. Numerous studies have shown that FA is mutagenic to mammalian cells, inducing point mutations and gene deletions *(8)*. Although FA can induce several types of DNA lesions, the vast majority of FA-

induced DNA lesions are DPCs *(10;11)*. DPCs formation by FA mainly involve sequential reaction with nucleophilic sites present in DNA and proteins, such as thiol, hydroxyl, amine, and amide, in the form of Schiff base reactions *(12;13)*. *In vivo* studies with rats and monkeys revealed that the rate of DPCs formation by FA is proportional to the concentration of FA in a tissue, consequently, the DPCs formed *in vivo* have been employed as biomarkers of human FA exposure *(14-16)*. Among nuclear proteins located in close proximity to chromosomal DNA, histones are commonly found to be cross-linked to DNA in the presence of FA, including H1, H2A, H2B, H3, and H4 *(17)*. Histone-DNA cross-links are considered to be strongly correlated with FA tumorigenesis and carcinogenicity *(18)*.

Metal complexes have also been shown to induce DPCs *(1)*. Exposure to metal compounds such as nickel, arsenic, and chromium have been associated with a variety of cancers, including those of the lung, liver, bladder, intestine, mouth, skin, blood *(19)*. Generally, two mechanisms are involved in DPCs formation by metal-ions: oxidative mechanism and metal chelating mechanism *(20-22)*. Chakrabarti and colleagues studied the DPCs induced by nickel (II) compounds in isolated rat lymphocytes, and reported that it was dependent on the generation of reactive hydroxyl radicals *(23)*. Alternatively, chromate was reported to form DPCs through a chelatable form of chromium involving the side chains of cysteine, histidine, and glutamic acid within proteins and adenine, cytosine, and guanine bases within DNA *(22)*.

Exposure of cells to UV and ionizing radiation (IR) results in the generation of DPCs in addition to more common types of DNA damage, such as single strand breaks, double strand breaks, base damage or loss, and DNA-DNA cross-links *(1;24)*. IR-induced free radicals and reactive oxygen species (ROS) can cause oxidation of DNA and proteins, potentially trapping the proteins onto the chromosomal DNA *(24)*. Electrospray-ionization tandem mass spectrometry (ESI-MS/MS) analysis of IR-induced DPCs demonstrated covalent bond formation between the C-5 methyl group of thymine and either the C2 or C3 of tyrosine *(25)*. Interestingly, IR is capable

of inducing DPCs in both aerated and hypoxic conditions, but greater numbers of DPCs are formed in the absence of oxygen *(26-28)*.

Another pathway to induce DPCs *in vivo* is endogenous exposure of cells to chemicals generated physiologically, i.e. reactive oxygen species (ROS) and lipid peroxidation products *(29)*. As normal DNA-protein interactions are essential for cellular process regulation, one electron oxidation occurring at either DNA or the proximal proteins can result in electron-deficient species, initiating a free radical reaction and ultimately giving rise to covalent DPCs. For example, Burrows and colleagues have reported that 8-oxodeoxyguanosine, a product of DNA one-electron oxidation, mediates DPC formation involving lysine or tyrosine-rich motifs that provide abundant nucleophiles for cross-linking *(30;31)*.

Recently, many studies have focused on DPC formation by chemotherapeutic *bis*-electrophiles, including nitrogen mustards, *cis*-diamminedichloroplatinum (II) (cisplatin), mitomycin C, and nitrosoureas *(32-34)*. Our laboratory has examined the formation of DPCs by two representative types of *bis*-electrophiles: 1,2,3,4-diepoxybutane (DEB) *(35)* and antitumor nitrogen mustards (i.e. mechlorethamine) *(36)*. In the current work, we have investigated cisplatin-mediated DPCs formation (*see Chapters II-IV*).

1.1.3 DPC repair

Once formed and not efficiently repaired, DPCs are expected to contribute to cytotoxicity and genotoxicity due to their bulky size and helix-distorting nature *(1)*. Studies on some types of cellular DPCs revealed that these lesions can be more stable as compared to other types of DNA damage and can persist through several DNA replcation cycles *(37;38)*. Many factors may contribute to the stability and repair of DPCs, given the fact the DPCs can be induced by a variety of chemical agents *via* different mechanisms. Therefore, depending on their chemistry, size, and protein involved, these lesions may be subjected to different repair pathways.

Due to their bulky nature, DPCs are likely candidates for nucleotide excision repair (NER) *(39)*. In NER, a damaged DNA base is recognized and removed by making two incisions in the DNA backbone at precise positions 5' and 3' to the damaged site *(2)*. Because the cross-linking of proteins to DNA may impose steric hindrance which might prevent the access of NER repair enzymes to the damaged site, proteolytic degradation is proposed to play a role by breaking down the protein components of DPC to peptides, followed by NER removal of the resulting DNA-peptide lesions *(2;40)*. For example, Minko *et al* generated site-specific DPCs by covalently trapping T4 pyrimidine dimer glycosylase (T4-pdg) protein to apurinic/apyrimidinic sites within duplex DNA. *In vitro* interactions of the resulting repair substrate with bacterial NER UvrABC nucleases provided direct evidence that UvrABC could perform the repair of site-specific bulky DPC lesions by making incisions 5' and 3' from the lesion *(41;42)*. Further studies by Baker and Reardon showed that DNA-polypeptide cross-links, produced as a result of proteolytic degradation of DPCs, were removed by human NER pathways with greater efficiency, suggesting a role for proteolytic degradation in NER-mediated DPCs repair *(39;43)*.

A different repair pathway, homologous recombination (HR), may also be involved in removing some types of DPCs *(44)*. For example, Nakano *et. al*. compared the roles of NER and HR in the repair of FA-induced DPCs both *in vitro* and *in vivo* using a bacterial system *(45)*. Their results indicated that NER and HR were coordinated to repair the bulky DPC lesion, in which NER removed DPCs with cross-linked proteins less than 12 – 14 kDa, whereas the oversized DPCs were exclusively eliminated by RecBCD-dependent HR *(45)*. The same group later found that the upper size limit of DPCs amenable to mammalian NER (8 – 10 kD) was even lower than that of bacterial cells, suggesting that NER cannot participate in the repair of the majority of chromosomal DPCs in mammalian cells *(46)*. Their studies also revealed the failure of an alternative NER mechanism in mammalian cells involving the initial proteolytic degradation of DPCs to peptides by the proteasome. In contrast, HR played a major role in cellular tolerance of DPCs, especially those involving large proteins. When the replication fork

ran into the site of unrepaired DPCs, it initiated HR and reactivated the stalled fork, although the detailed mechanism of such reactivation was unclear *(3)*. Taken together, these findings support the existence of active repair processes for DPCs in both prokaryotes and eukaryotes, and such repair may require more than one pathway depending on DPC identities and structures. Further studies are still needed to better appreciate how cells repair structurally diverse DPC lesions.

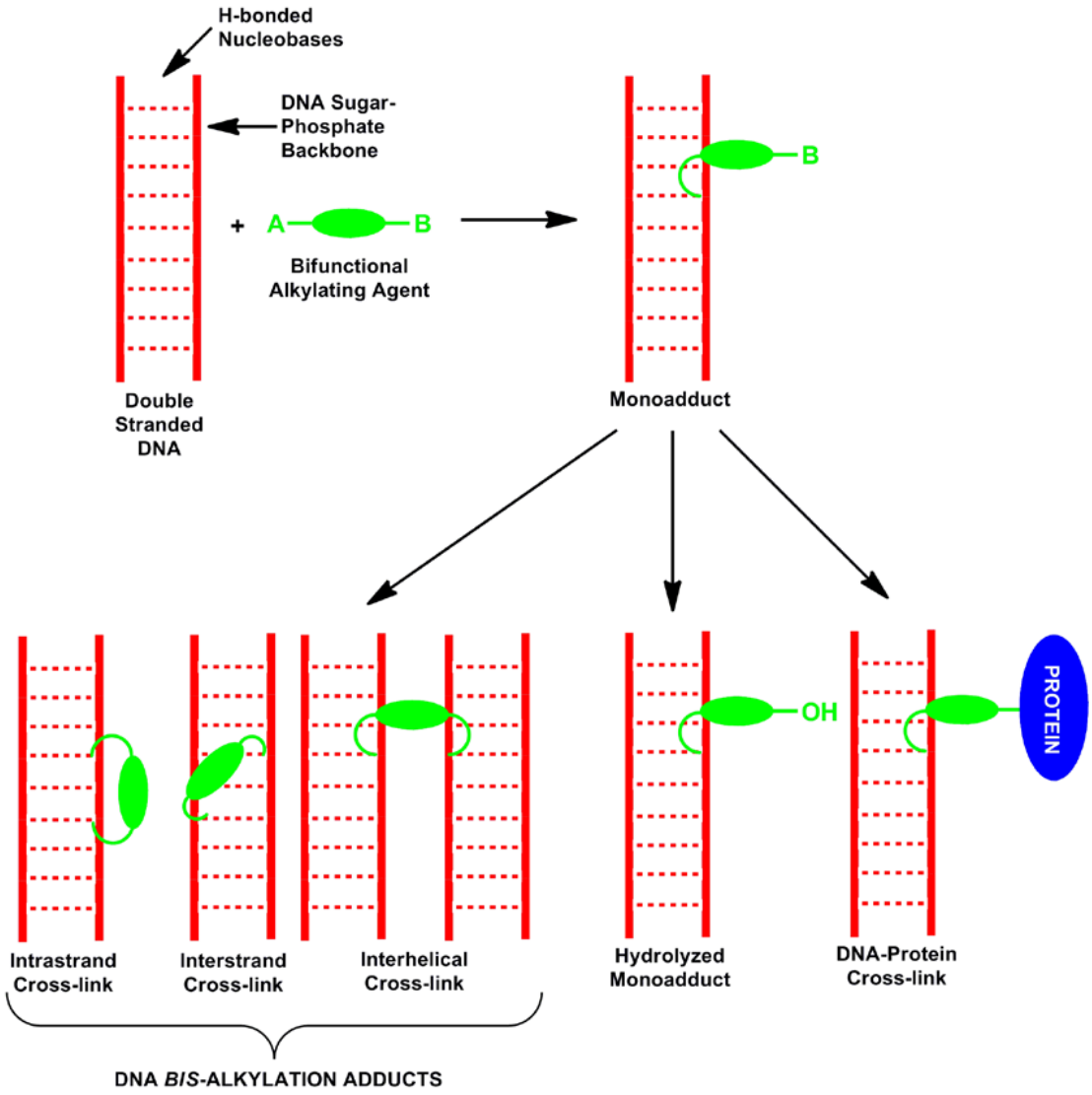
1.1.4 Biological Consequences of DPC formation

Considering the extremely bulky and helix-distorting nature of DPCs, they are likely to impede DNA-involving biological processes (i.e. replication, repair, recombination, transcription, and chromatin remodeling) by mechanisms different from those for conventional DNA lesions *(1;3)*. Indeed, DPCs may physically block replication and transcription machineries and/or prevent the access of proteins required either for RNA synthesis along the template strand, for transcription, or for DNA damage recognition. However, because the vast majority of DPC-inducing agents are capable of causing other types of DNA damage such as monoadducts and DNA-DNA cross-links (Scheme 1.1), it is very difficult to directly attribute any of the observed cytotoxic, mutagenic, and carcinogenic effects to the formation of DPCs. Therefore, in order to fully comprehend the implications of DPCs involving their induction, repair, and biological consequences, it is critical to develop novel methodology which enables selective induction and/or isolation of DPC *in vivo* in the absence of the effects of other types of DNA lesions.

1.2 Detection of DPCs

Historically, several experimental methodologies have been developed to detect DPCs such as gel shift assay, comet assay, and mass spectrometry. Early studies focused on the abilities of various cross-linking agents to form DPCs, both *in vitro* and *in vivo*, and on quantitative analysis of DPC formation *(34;47)*. More recently, researchers have been more interested in exploring the identities of specific proteins which participate in DNA-protein cross-linking and

Scheme 1.1 DNA damage by *bis*-electrophiles.



elucidating the exact chemical structures of the resulting amino acid-nucleobase conjugates, by means of emerging advanced technologies such as mass spectrometry and proteomics *(48;49)*.

1.2.1 Classical Techniques Employed in the Study of DPCs

The simplest and the most widely used method for detection of DPCs *in vitro* is the gel shift assay, commonly referred to as electrophoretic mobility shift assay (EMSA) *(34;50-52)*. In this type of experiment, radiolabeled oligodeoxynucleotides are incubated with proteins of interest in the presence or in the absence of the cross-linking agent. The resulting mixture is then resolved by denaturing gel electrophresis, followed by visualization by autoradiography. These experimental conditions dissociate any non-covalent DNA-protein complexes, but preserve covalent DPCs, which can be detected on a gel due to their reduced mobility as compared to free DNA.

The gel shift assay has been successfully employed by many groups to demonstrate DPC formation involving specific proteins by a number of cross-linking agents, including cisplatin, nitrogen mustards, and DEB. For example, Brabec *et. al.* have employed this method to show that various platinum compounds were capable of cross-linking high mobility group (HMG) proteins, histone proteins, transcription factors, and replicative enzymes to DNA *(53)*. Recently, our laboratory employed a similar assay to detect cross-linking between ^{32}P-labeled double-stranded oligodeoxynucleotides and human recombinant O^6-alkylguanine DNA alkyltransferase (AGT) in the presence of nitrogen mustards *(54)*, DEB *(55)*, and cisplatin (*see Chapter II*). Although the gel shift assay provides a simple and inexpensive way to detect DPCs involving specific proteins, the main limitation of this technique is that it does not provide any information regarding the covalent structures of the DPC lesions ad cannot identify cross-linked proteins generated in biological systems (i.e. cell nuclear extracts or mammalian cells).

A biophysical technique of alkaline elution was developed in the 1970s by Kohn and colleagues to detect several types of DNA damage including single strand breaks, DNA-DNA

cross-links, and DPCs *(56)*. The principle of this methodology involves metabolically radiolabeling of cellular DNA with [^{14}C]-thymidine, followed by exposure of the resulting cells to DPC-inducing agents. DNA damage such as DPCs is then determined from the rate at which DNA from the lysed cells migrates through membrane filters under alkaline conditions (*via* scintillation counting). Unbound DNA stands pass through the filter more readily than DNA segments containing covalently-bound proteins. Since other types of DNA lesions (i.e. interstrand DNA-DNA cross-links) may show similar affinity to the filter as DPCs, sample pre-treatment with proteolytic enzymes is further required to differentiate these two types of DNA damage *(56)*. Although alkaline elution has some of the same limitations as the gel shift assay, it has been used to detect DPCs induced by several cross-linking agents (i.e. platinum compounds *(32)*, antitumor nitrogen mustards *(57)*, DEB *(50)*, and alkylnitrosoureas *(34)*) *in vivo*.

Another biophysical technique that has been successfully employed in the study of DNA-protein cross-linking in cell cultures is the comet assay, or single-cell gel electrophoresis (SCGE), developed in the late 1980s by Ostling and Johanson *(58)*. In this method, cells are exposed to a DNA-damaging agent and further subjected to γ-irradiation to produce DNA strand breaks. The cells are encapsulated into a low-melting-point agarose suspension, lysed under neutral or alkaline conditions to break open the cells and disperse all cellular components into the agarose matrix, and subjected to electrophoretic separation and visual analysis by fluorescent staining of DNA. While unmodified genomic DNA does not migrate through the agarose matrix, forming the comet's "head", the damaged DNA fragments induced by γ-irradiation elute along the gel to form comet's "tail". The length of the tail can be used as a measurement of the extent of DNA damage *(58)*. The comet assay has been used to detect DPC formation by several DNA damaging agents, such as formaldehyde and chromate *(11;59)*. However, as discussed before, many of these agents are able to form other types of DNA lesions in addition to DPCs, so the co-existence of DPCs and DNA-DNA cross-links requires further differentiation by protease pre-treatment *(58)*. In summary, the comet assay is a fast, effective, and inexpensive way of measuring DPCs in

individual cells, however, it fails to provide information about the identities of the cross-linked proteins and structures of the cross-linked conjugates, urging the development of more specific and sensitive methodologies capable of selectively isolating and characterizing DPCs.

In the l990s, Costa *et. al.* reported a novel breakthrough methodology which enabled the detection of DPCs in cell culture and *in vivo* (tissue samples) *(47)*. This method is based on co-precipitation of all cellular proteins, along with DNA-protein complexes, in the presence of potassium chloride. Cell/tissue lysates are heated in the presence of sodium dodecyl sulfate (SDS) and KCl, followed by low-speed centrifugation. The resulting DPC-containing protein pellets are subjected proteolytic digestion to release the DNA which was once bound to a protein. The resulting DNA is subsequently recovered and quantified as an indication of the amount of DNA involved in DPCs formation. Potassium-SDS precipitation has been utilized in numerous studies to detect DPCs in cells exposed to a variety of agents, including platinum, lead, chromium, and formaldehyde *(9;60)*.

1.2.2 Mass Spectrometry: Characterization of DNA-protein Cross-linking

Recent advances in the field of mass spectrometry have accelerated its full application to the identification and characterization of a variety of DNA lesions such as DPCs *(35;36;61-63)*. Numerous research groups have successfully applied mass spectrometry to elucidate the atomic connectivity in DPCs macroconjugates involving amino acid-nucleobase heteroconjugation and map the cross-linking sites within specific proteins *(48;54;55)*.

For example, Yang *et. al.* employed mass spectrometry to characterize DNA-protein cross-linking in the presence of oxanine, an oxidative lesion induced by reactive nitrogen species *(48)*. To investigate the reactivity of the amino acids that may participate in DPC formation by oxanine, lysozyme was incubated with 2'-deoxyoxanosine, followed by tryptic digestion and nanoscale liquid chromatography coupled with nanospray ionization tandem mass spectrometry (nanoLC-NSI-MS/MS) analysis of the resulting peptides. Several peptides were observed to

experience +152 Da mass shift which corresponds to the addition of depurinanted oxanine moiety, and further MS/MS analysis was conducted to assign the cross-linking sites to Lys^{13}, Lys^{97}, Lys^{116}, Ser^{85}, and Ser^{86} of the protein. Furthermore, a similar DNA-protein cross-linking experiment was conducted by incubating lysozyme with oxanine-containing calf thymus DNA, leading to oxanine-induced cross-linking modification at Lys^{116}, Ser^{85}, and Ser^{86}. These findings confirm that mass spectrometry is a valuable tool for the characterization of DPC structures *(48)*.

Mass spectrometric analysis can also be used to provide direct information about the structural relationships between nucleic acids and the non-covalently associated proteins using photochemical cross-linking. For example, Doneanu and colleagues photochemically cross-linked human replication protein A (hRPA) to 30-mer oligonuclotide dT_{30} by UV irradiation, and subsequently employed a combination of matrix-assisted laser desorption and ionization time of flight (MALDI-TOF) MS and nanoLC-NSI-MS/MS analyses to determine amino acid residues that reside in the hRPA's DNA-binding domain *(64;65)*. Following UV-irradiation, SDS-PAGE separation, and in-gel tryptic digestion, three peptides were identified to contain cross-links to DNA, including A^{235}TAFNEQVDKFFPL^{248}, I^{264}ANKQFTAVK^{273}, and V^{383}SDFGGR^{389}. Further exhaustive C-terminal proteolysis coupled with MALDI-TOF-MS was used to localize the sites of DNA binding at or near the Phe^{238}, Phe^{269}, and Phe^{386} residues, in agreement with an X-ray crystallographic profile that already exists for hRPA's DNA-binding domain. These findings indicated that DNA-protein cross-linking combined with mass spectrometry can be successfuly used to study non-covalent DNA-protein interactions. Data acquired can complement or amend the results obtained from site-directed mutagenesis and X-ray crystallography *(65)*.

1.2.3 Proteomics Technologies: Identification of Proteins Which Form DPCs

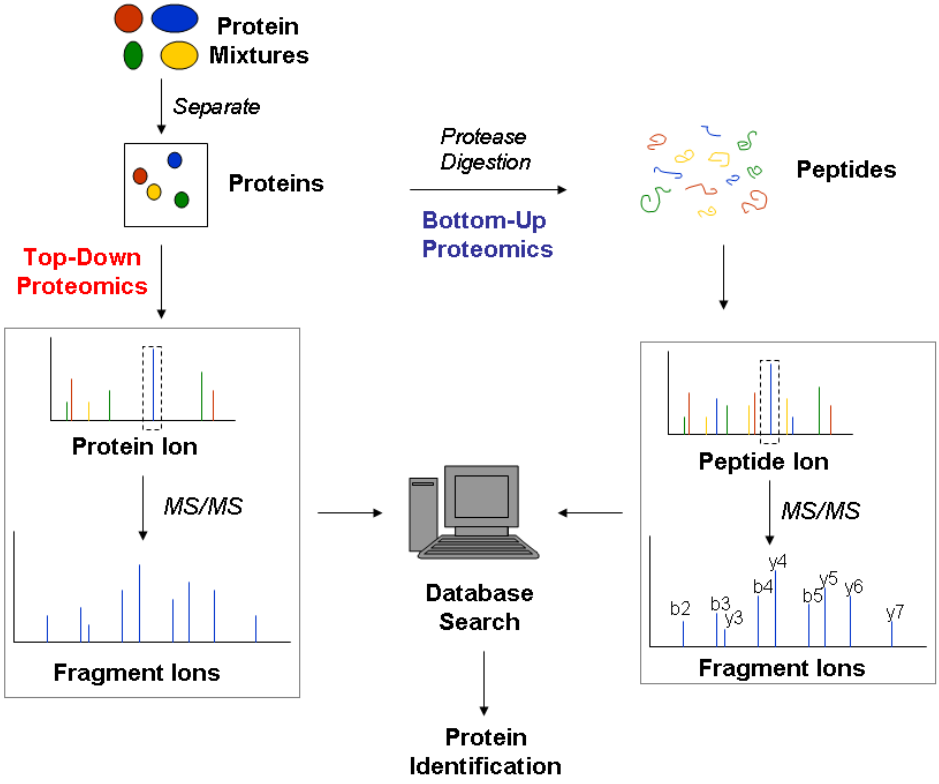
As the research interest of DNA-protein cross-linking increasingly lies in the identities of the cross-linked proteins in more complex biological systems (i.e. nuclear extracts or different cell lines), mass spectrometry-based proteomics have been widely utilized for the identification

and quantitation of proteins which participate in DNA-protein cross-linking *(49)*. In regard to characterization of cross-linked proteins using MS-based proteomics, two possible strategies can be performed: the so-called top-down and bottom-up approaches (Scheme 1.2). In a typical top-down experiment, the cross-linked protein mixtures are separated by gel electrophoresis or liquid chromatography *(66)*. The proteins are then ionized by electrospray ionization (ESI) and trapped in a Fourier transform ion cyclotron resonance (FTICR) or quadrupole ion trap mass spectrometer, followed by fragmentation via collision-induced dissociation (CID), electron-capture dissociation (ECD), or electron-transfer dissociation (ETD). Cleavage between every possible position within the protein provides a mass ladder, which can be further used to match against theoretical values in a database to identify the proteins present in the original samples *(66)*.

The bottom-up proteomics is the dominant approach in proteomics today, where protein mixtures are digested to peptides prior to mass spectrometric analysis *(67)*. The major advantage of the bottom-up method lies in high-throughput protein identification because peptides produced by tryptic digestion allow efficient MS analyses in combination with increasing sequence coverage and data-matching. Recently, the complementary use of top-down and bottom-up approaches have been reported by numerous groups to study post-translational modifications, characterize DNA-protein or protein-protein interactions, and identify disease biomarkers.

Our laboratory combined affinity capture methodology with mass spectrometry-based proteomics and immunological detection for the analysis of DPCs formed *in vitro* (nuclear protein extracts) exposed to *bis*-electrophilic cross-linking agents such as the antitumor agent mechlorethamine *(36)*. Nuclear extract protein from human cervical carcinoma (Hela) and CHO cells were incubated with biotinylated oligodeoxynucleotides in the presence of mechlorethamine to induce DPC formation. DNA containing covalently attached proteins was captured on streptavidin beads, washed to remove any non-covalently bound proteins, and released from the DNA backbone in the form of protein-guanine conjugates by thermal hydrolysis. Following SDS-PAGE separation and in-gel tryptic digestion of cross-linked proteins, the resulting peptides were

Scheme 1.2 Experimental scheme for mass spectrometry-based proteomics: Top-Down and Bottom-Up Proteomics.



analyzed by HPLC-ESI-MS/MS. SEQUEST database searching revealed that 53 Hela and 15 Chinese hamster ovary (CHO) cells nuclear proteins were trapped on the DNA by mechlorethamine. The proteins participating in DPC formation included architectural structural proteins, chromatin-regulatory proteins, proteins related to transcription and translation, DNA replication and repair proteins, and those involved in the cell cycle.

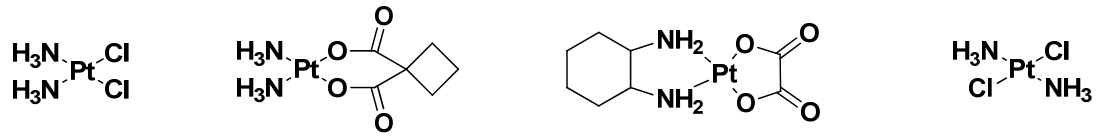
More recently, our laboratory investigated *in vivo* DNA-protein cross-linking by mechlorethamine in human fibrosarcoma (HT1080) cells *(35)*. Following drug treatment, DPCs were isolated by a modified phenol/chloroform DNA extraction. Proteins were further released from DNA by thermal hydrolysis and identified by mass spectrometry-based proteomics. Thirty eight nuclear proteins participated in DNA-protein cross-linking in the presence of mechlorethamine.

1.3 Cross-linking Agent Examined in this Work

As described above, our laboratory has previously employed mass spectrometry to study DNA-protein cross-linking both *in vitro* and *in vivo* by two *bis*-alkylating agents: DEB and mechlorethamine *(35;36)*. We now extend these studies to cisplatin-mediated DNA-protein cross-linking within human recombinant proteins, nuclear protein extracts, and cultured mammalian cells (*see Chapters II-IV*).

Platinum-based compounds, such as cisplatin, carboplatin, and oxaliplatin (Figure 1.1), are efficient antitumor agents widely used in the treatment of solid tumors of the brain, head and neck, ovary, testicle, and bladder *(68)*. Interestingly, transplatin, the isomer of cisplatin, is clinically ineffective (Figure 1.1). It is generally accepted that the biological activity of cisplatin is due to its ability to form DNA adducts, e.g. 1,2-intrastrand DNA-DNA cross-links (1,2-GG or 1,2-AG intrastrand, 90% of total lesions), 1,3-intrastrand adducts between purine bases separated by another base (1,3-GG or 1,3-AG, 5-10% of total lesions), interstrand DNA-DNA crosslinks (2%), and monofunctional adducts (3% of total lesions) *(69;70)*. However, cisplatin is also

Figure 1.1 Chemical structures of platinum-based antitumor drugs and the inactive cisplatin isomer, *trans*-diamminedichloroplatinum (II) (transplatin).



Cisplatin **Carboplatin** **Oxaliplatin** **Transplatin**

capable of forming ternary DNA-platinum-protein cross-links (DPCs) (Scheme 1.3) *(32;71;72)*. For example, Zwelling *et al.* employed alkaline elution methodology (see above) to detect DPC formation in L1210 mouse leukemia cells treated with cisplatin *(32)*. This observation was subsequently confirmed by other groups that utilized similar biophysical methods for detecting cisplatin-induced DPCs *(71)*. For instance, cisplatin has been reported to form DPCs involving HMG proteins, histones, and cytokeratins *(73)*. Furthermore, DNA-protein cross-linking has been shown to play an important role in the cytotoxicity of cisplatin *(74)*. However, little is known about the identities of other nuclear proteins that participate in cross-linking to DNA in the presence of cisplatin. Furthermore, the covalent structures of cisplatin-mediated DPCs and their cellular abundance have not been determined. Thus, the ability of cisplatin to induce DPCs both *in vitro* and *in vivo* was investigated in our laboratory (*see Chapters II-IV*).

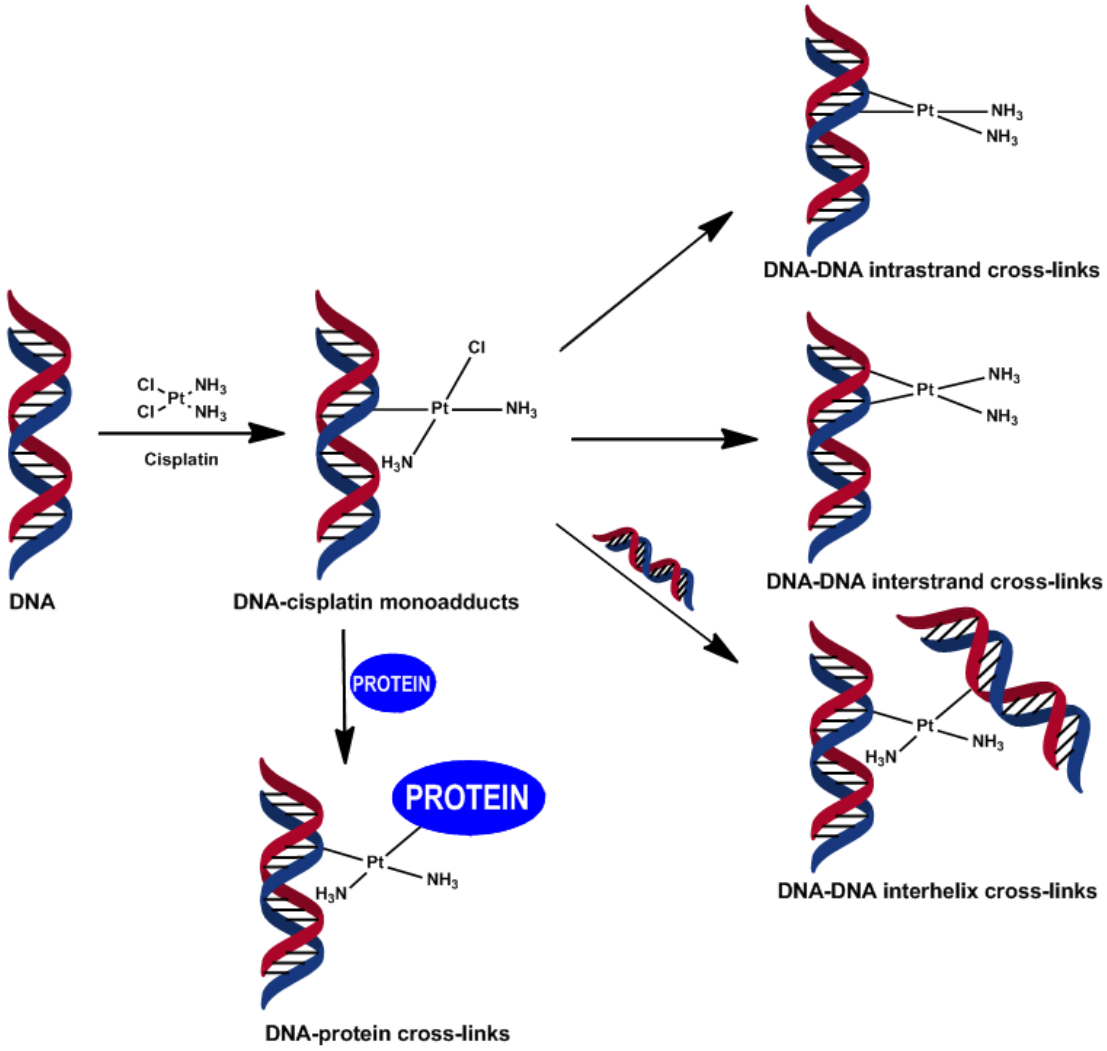
1.4 Biological Importance of DNA Oxidation

Reactive oxygen and reactive nitrogen species (ROS and RNS), such as peroxide (H_2O_2), superoxide ($O_2^{\bullet-}$), nitric oxide ($NO^{\bullet}$), peroxynitrite ($ONOO^-$), nitrosoperoxycarbonate ($ONOOCO_2^-$) and the hydroxyl radical ($^{\bullet}OH$) *(75)*, are produced in tissues as a result of normal aerobic metabolism, immune response, and inflammation *(76-78)*. ROS and RNS play a dual role in a cell as they are required for certain cellular processes, but can damage cells when overproduced. At low concentrations, ROS/RNS can be involved in cellular signal transduction and immune response *(77)*. In contrast, overproduction of ROS/RNS can lead to oxidative degradation of certain cellular biomolecules, contributing to toxicity and cancer *(75)*.

1.4.1 Chemistry of DNA Oxidation: One-electron Oxidation

Oxidative degradation of DNA has received much attention due to its potential role in aging, cancer, and in some neurodegenerative diseases *(76-78)*. Guanine bases have the lowest

Scheme 1.3 Formation of bifunctional DNA adducts by cisplatin.



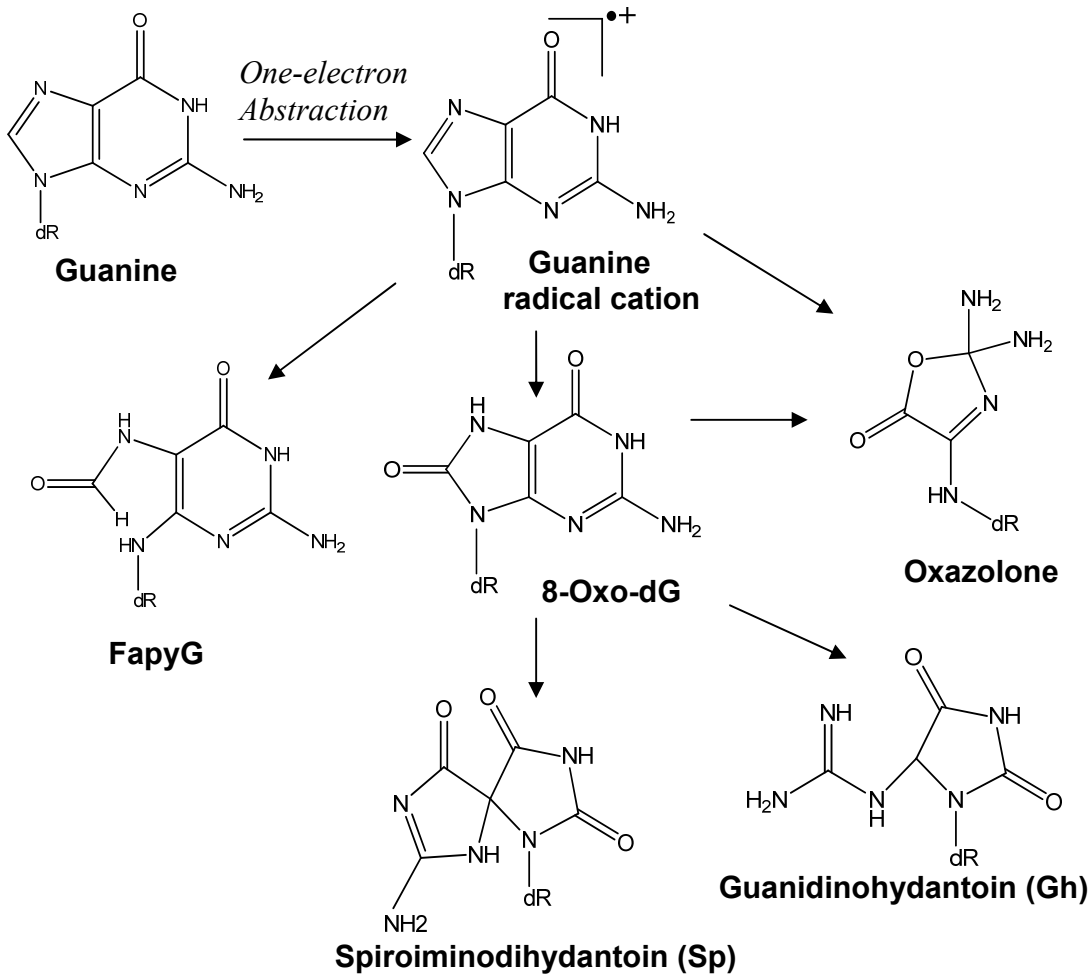
oxidation potential among DNA bases and are usually the most easily oxidized *(79;80)*. Guanine oxidation gives rise to a variety of oxidation products, including 8-oxo-7,8-dihydro- 2'-deoxy-guanosine (8-oxo-dG), spiroiminodihydantoin (Sp), guanidinohydantoin (Gh), 2,6-diamino- 4-hydroxyl-5-formamidopyrimidine (FapyG), and 2,2-diamino-4-[(2-deoxy- *β*-D-*erythro*-pentofuranosyl)amino]-5(2*H*)-oxazolone (oxazolone) (Figure 1.2) *(81;82)*. Formation of these oxidative products is highly dependent on the nature of the oxidizing species, generally involving three possible mechanisms: direct electron transfer from the guanine base, addition of hydroxyl radicals ($^{\bullet}$OH), and [4+2] cycloaddition of singlet oxygen *(81)*.

As for the one-electron oxidation mechanism, the initial step involves one-electron abstraction of guanine to form a guanine radical cation ($G^{\bullet+}$), usually induced by one-electron oxidants such as riboflavin *(83)*. The resulting electron migrates along the π-stack of the B-DNA duplex until it reaches sites with relative lower oxidation potential, where the radical cation irreversibly reacts with water or molecular oxygen *(84)*. Direct evidence of electron transfer along DNA duplexes was observed by studies using synthetic double-stranded oligodeoxynucleotides with photo-oxidizing agents covalently attached to the 5'-ends *(84;85)*. Following photo-activation, a cationic hole was initially injected into the π-stack, however, oxidative damage was found at sites located up to 200 Å away from the tethered oxidant at the 5'-ends, suggesting that the cationic hole produced by oxidizing agents can migrate long distances through the π-stack of a DNA duplex until being irreversibly trapped.

1.4.2 DNA Sequence Context as a Determinant in Guanine One-electron Oxidation

Ab initio calculations by Saito and coworkers indicated that the ionization potentials (IP) of G are dependent on DNA sequence context, with the lowest guanine IP occurring at the 5'-guanine in runs of several Gs (*e.g.* 5'-**G**G-3' or 5'-**G**GG-3') *(86)*. Later, the Dedon group quantified riboflavin-induced damage at each guanine in 5'-^{32}P-labeled oligodeoxynucleotides

Figure 1.2 Guanine oxidative products following one-electron oxidation.



containing guanine in all possible three-base sequence contexts (5'-XGY-3'). Guanine lesions were converted to single strand breaks, and the resulting fragments were resolved on sequencing gels. The relative amount of oxidative lesions originating from each guanine was then plotted against their ionization potential (IP). A strong inverse correlation between the IP and the extent of oxidation has been consistently confirmed by numerous studies, suggesting that sequence-specific IPs were the major determinants of guanine's reactivity towards one-electron oxidation in DNA duplexes *(87;88)*.

1.4.3 Unusual Sequence Selectivity in Guanine Oxidation

It is tempting to further predict the hotspots for DNA oxidation toward one-electron oxidation simply based on the assumption that guanines with the lowest IP will be most readily oxidized as a result of charge migration. However, Margolin and co-workers observed a totally different DNA oxidation behavior when they examined the guanine reactivity toward another one-electron oxidant nitrosoperoxycarbonate ($ONOOCO_2^-$) arising from inflammation *(87;88)*. By using the same gel electrophoresis method, they analyzed the $ONOOCO_2^-$ reactivity efficiency with guanine in all possible three-base sequence contexts (5'-XGY-3'). Contrary to their observations from riboflavin-induced oxidation, $ONOOCO_2^-$ preferentially oxidized guanines with the highest IP in GC-containing context *(87)*. For example, the highest frequency of $ONOOCO_2^-$-induced damage occurred in A**G**C and T**G**C sequences bearing the highest guanine IP. Generally, an increase in reactivity of guanine oxidation was in parallel to increasing IP. Furthermore, their results indicated that the GC motif was the determinant rather than IP in $ONOOCO_2^-$-induced oxidation, suggesting that, although both $ONOOCO_2^-$ and riboflavin share the initial one-electron abstraction step to form the radical cation, sequence-dependence of guanine oxidation should be considered as a function of multiple factors, such as the nature of the oxidant, kinetics and stability of the resulting radical cation, solvent accessibility, and its subsequent chemical behaviors *(87)*.

1.4.4 DNA Sequences and Oxidizing Agent Examined in this Work

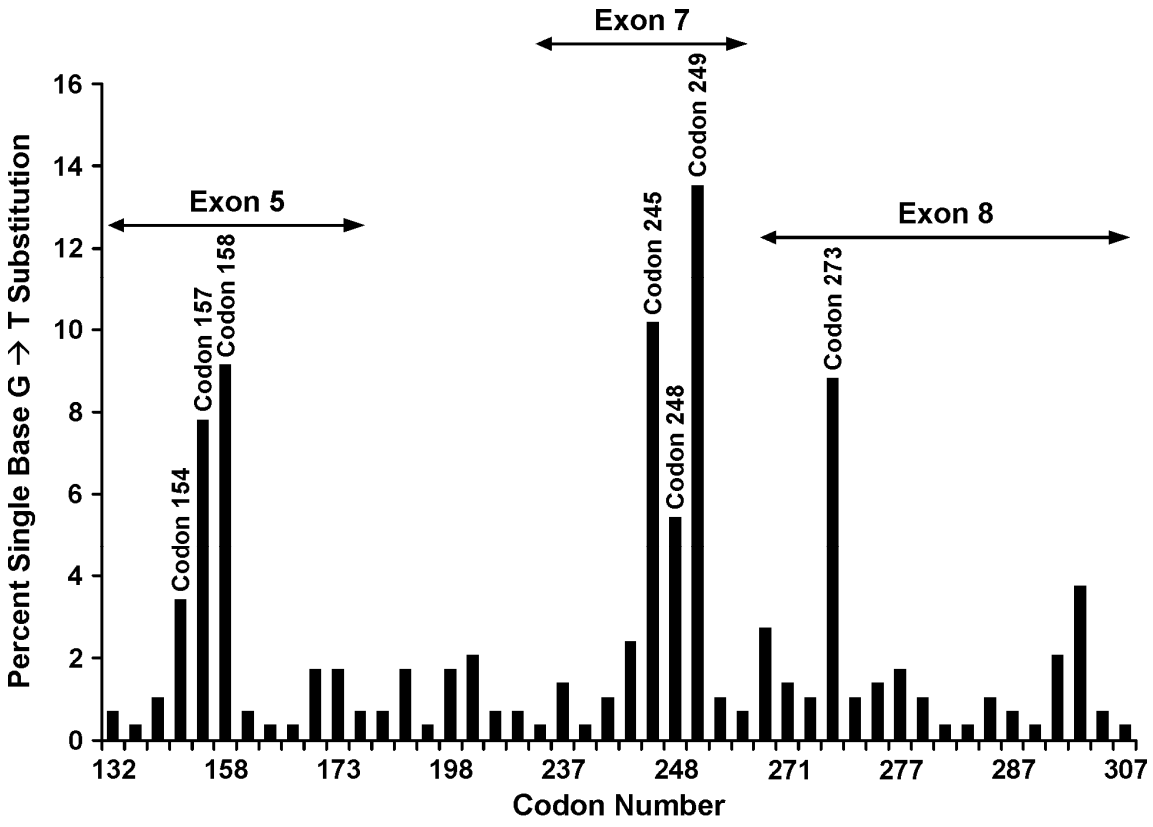
One of the major targets for DNA adduct formation and mutations in smoking- associated lung cancer is the *p53* tumor suppressor gene *(89;90)*. The *p53* tumor suppressor gene is involved in multiple important cellular processes, including control of cell cycling, gene transcription, chromosomal segregation, DNA repair and apoptosis *(91)*.

The *p53* gene encodes a nuclear protein that functions as a cell cycle checkpoint to respond to DNA damage and to assist DNA repair processes prior to DNA replication and mitosis, thus limiting the occurrence of mutations *(91)*. Mutations in the *p53* gene are likely to produce other gene products that are unable to bind DNA, thus blocking DNA repair as well as apoptosis, which further leads to a loss of control over cell growth and ultimately the formation of tumors *(92)*.

The *p53* tumor suppressor gene is mutated in many human cancers, including smoking-induced lung cancer *(93)*. The majority of the *p53* base substitution mutations observed in smoking-associated lung tumors are G → T transversions and G → A transversions frequently occurring at *p53* exons 5, 7, and 8 *(94)*. The *p53* tumor suppressor gene is mutated in over 40% of human lung cancer and these mutations are found more frequently occurring in smokers than non-smokers, suggesting that the *p53* tumor suppressor gene might be a potential target of tobacco carcinogens *(92;95)*. Interestingly, mutations in the *p53* tumor suppressor gene are not randomly distributed, but rather are preferentially located in exons 5-8 at endognenously methylated CG dinucleotides (Figure 1.3) such as codons 154 (*G**G**C* → *G**T**C*), 157 (***G**TC* → ***T**TC*), 158 (*C**G**C* → *C**T**C*), 245 (***G**GC* → ***T**GC*), 248 (*C**G**G* → *C**T**G*), 249 (*A**G**G* → *A**T**G*), and 273 (*C**G**T* → *C**T**T*) *(92;94;95)*.

One possible mechanism for the increased mutagenesis at ^{Me}CG sites within the *p53* gene involves targeted binding of metabolically activated tobacco carcinogens to these sequences. For example, previous studies in our laboratory revealed that bulky DNA adducts of polycyclic

Figure 1.3 Distribution of G → T mutations in human lung cancer mapped along exons 5-8 of the *p53* gene. The frequency of distribution of G → T mutations was obtained from the AACR *p53* mutation database (2010). Cancer from nonsmokers and from occupational exposure was excluded.

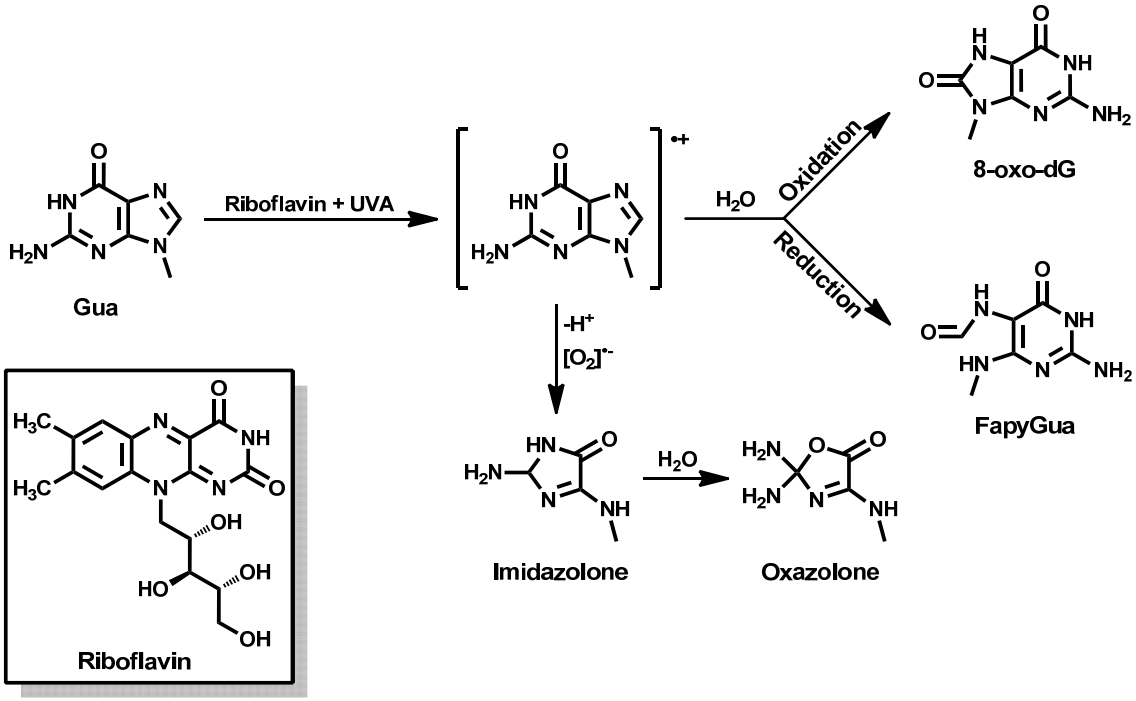


aromatic hydrocarbons present in tobacco smoke are preferentially formed at endogenously methylated CG dinucleotides *(96-99)*. Our recent study indicates that PAH diepoxides preferentially form physical complexes with ^{Me}C:G base pairs, leading to increased adduct yields at these sites *(100)*. However, other constituents of tobacco smoke such as reactive oxygen species may also target ^{Me}CG dinucleotides, contributing to mutagenesis.

Oxidative DNA damage is a likely contributor to the pathogenesis of lung cancer *(101)*. The increased cellular load of oxidative DNA lesions in the tissues of smokers likely plays a role in lung tumor induction. Both 8-oxo-dG and oxazolone have been shown to cause G → T transversion mutations in both prokaryotic and eukaryotic cells *(102)*. Furthermore, cytosine methylation has been demonstrated to facilitate guanine oxidation at the base-paired position *(103)*. Therefore, we hypothesized that oxidative damage may contribute to G → T transversions observed at the lung cancer mutational hotspots within endogenously methylated CG dinuclotides at *p53* exons 5, 7, and 8. Therefore, we hypothesized that oxidative damage may contribute to G → T transversions observed at lung cancer mutational hotspots within endogenously methylated CG dinuclotides at *p53* exons 5, 7, and 8.

The goal of the present investigation was to analyze the distribution of specific riboflavin-mediated photooxidation products (8-oxo-dG and oxazolone) along DNA duplexes (Scheme 1.4) derived from the frequently mutated regions of the *p53* gene within exons 5, 7, and 8 (*see Chapter V*).

Scheme 1.4 Formation of oxidative guanine adduct by one-electron oxidation in the presence of riboflavin.



II. Cross-linking of the DNA Repair Protein O^6-Alkylguanine DNA Alkyltransferase to DNA in the Presence of Cisplatin

2.1 Introduction

Reversible DNA-protein interactions play an important role in normal cell function. Nuclear protein binding to regulatory sequences within DNA controls DNA replication, gene expression, and mediates responses to DNA damage *(104-106)*. Any interruptions of these dynamic interactions can have serious consequences for cell viability and genetic stability *(1)*. For example, proteins can become covalently trapped on chromosomal DNA as a result of exposure to physical and chemical agents such as formaldehyde *(16)*, ionizing radiation *(24)*, and anticancer drugs *(34;107;108)*. The resulting irreversible DNA-protein cross-links (DPCs) are expected to block normal DNA-protein interactions, potentially contributing to toxicity, cancer, and neurodegenerative diseases *(1;38)*.

The role of DPC lesions in biological effects of *bis*-alkylating drugs is not well understood due to their inherent complexity and the propensity of these agents to induce other types of DNA damage. For example, the anticancer drug 1,1,2,2-*cis*-diamminedichloroplatinum (II) (cisplatin) is capable of forming DNA-DNA cross-links and monoadducts *(109)* in addition to DPCs *(110)* (Scheme 1.3). While DNA-DNA cross-linking by cisplatin is well-characterized and thought to be responsible for its anticancer mechanism *(111)*, only limited information is available about the corresponding DNA-protein cross-links. DPCs are estimated to constitute < 1% of total DNA damage following exposure to cisplatin *(112)* and constitute a highly heterogenous and complex type of DNA damage, making it difficult to evaluate their role in the cytotoxic and mutagenic effects of platinum drugs *(32;113)*.

O^6-Alkylguanine DNA alkyltransferase (AGT) protein is an important repair protein that removes promutagenic DNA O^6-alkylguanine lesions formed as a result of exposure to

chemotherapeutic drugs and environmental toxins *(114)*. AGT transfers the O^6-alkyl group from O^6-alkylguanines in DNA to an active site cysteine residue within the protein active site (Cys^{145}), thus restoring normal guanine *(115)*. Crystal structures of AGT-DNA complexes reveal that during the alkyl transfer reaction, the alkylated nucleotide is flipped out of the DNA base stack to enter the AGT active site *(115)*.

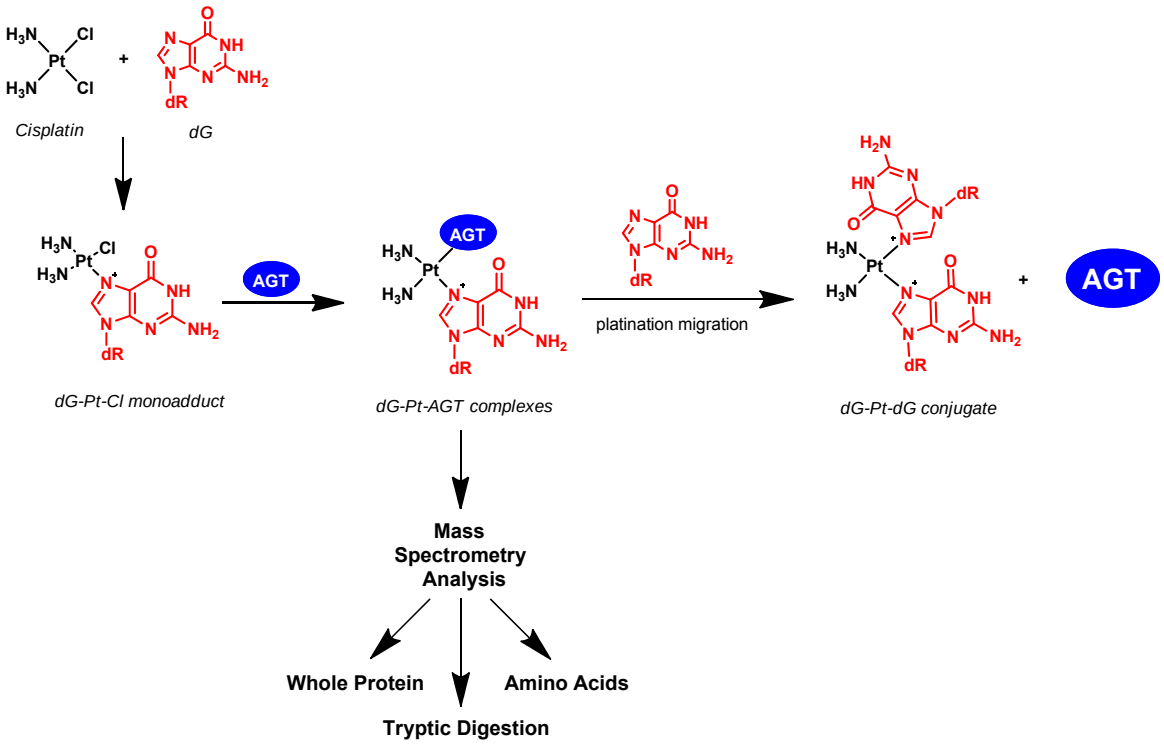
Recent studies have revealed that AGT is readily cross-linked to DNA in the presence of various *bis*-electrophiles, probably due to its high affinity for DNA and the remarkable nucleophilicity of AGT Cys^{145}, which is activated to a thiolate anion *via* a hydrogen bonding network in the protein active site *(116)*. As a result, the cytotoxicity and mutagenicity of *bis*-electrophiles such as 1,2-dibromoethane, dibromomethane, and 1,2,3,4-diepoxybutane (DEB), are enhanced in bacteria expressing human AGT protein due to the formation of toxic AGT-DNA cross-links *(117;118)*.

Based on the known ability of AGT to form toxic DPC lesions *(118;119)*, this protein was selected for our *in vitro* studies of DNA-protein cross-linking by cisplatin. Gel shift assay and electrospray mass spectrometry were used to demonstrate that cisplatin forms covalent cross-links between AGT protein and the N7 position of dG. Furthermore, peptide sequencing by tandem mass spectrometry has led to identification of six amino acid residues within AGT involved in cross-linking (Scheme 2.1).

2.2 Materials and Methods

Chemicals and Reagents - Cisplatin, dG, and human recombinant glyceraldehyde 3-phosphate dehydrogenase (GAPDH) were purchased from Sigma-Aldrich (St. Louis, MO), Boc-L-Lys-OH was obtained from Fluka (Buchs, Switzerland), and mass spectrometry-grade Trypsin Gold were purchased from Promega (Madison, WI). Adenosine 5'-[γ-^{32}P]-triphosphate was obtained from Perkin-Elmer (Boston, MA). T4 polynucleotide kinase and proteinase K were purchased from New England Biolabs (Beverly, MA). Synthetic DNA oligodeoxynucleotides were prepared at

Scheme 2.1 Mass spectrometry-based approach employed to characterize AGT-DNA cross-links of cisplatin and platination migration from AGT protein to dG to release intact AGT and form G-G cisplatin cross-links.



the University of Minnesota Microchemical Facility (Minneapolis, MN). Recombinant *C*-terminal histidine-tagged hAGT was produced as described previously *(120)*. *Cis*-1,1-diammine-2-chloro-2-(2'-deoxyguanosine-7-yl)-platinum (II) (dG-Pt-Cl) monoadducts and *cis*-1,1-diammine-2,2-*bis*-(2'-deoxyguanosine-7-yl)platinum (II) (dG-Pt-dG) conjugates were produced according to previous reports *(121)*.

1,1-Cis-diammine-2-(5-amino-5-carboxypentyl)amino-2-(2'-deoxyguanosine-7-yl)-platinum(II) (dG-Pt-Lys) **-** Cisplatin (10 mg, 33.33 μmol) was dissolved in 1 mL of 10 mM TRIS-HCl buffer (pH 7.2). $AgNO_3$ (11.32 mg, 66.66 μmol) was added, and the solution was kept in the dark with stirring for 4 h at room temperature. Following centrifugation, to the filtrate was added 2'-deoxyguanosine (dG) (9 mg, 33.33 μmol) and Boc-L-Lysine (8.2 mg, 33.33 μmol), and the resulting mixture was incubated for 48 h at 37 °C. The precipitate was isolated by filtration, and the supernatant was separated by semi-preparative HPLC on a Supelcosil LC-18-DB column (25 cm x 10 mm, 5 μm) eluted with a linear gradient of acetonitrile (B) in 15 mM ammonium acetate, pH 4.9 (A). The solvent composition was changed from 0 to 24% B in 24 min and further to 60% in 6 min. Under these conditions, Boc-protected dG-Pt-Lys eluted at 20.4 min. ESI^+-MS/MS (dG-Pt-Lys-Boc): *m/z* 741.3 $[M]^+ \rightarrow$ *m/z* 724.2 $[M - NH_3]^+$, 624.5 $[M - NH_3 - Boc]^+$. Following HPLC purification, the Boc protective group was removed by incubation in 10% TFA (0.5 mL) at room temperature for 30 min. The deprotected dG-Pt-Lys was purified using the same HPLC method. Under these conditions, dG-Pt-Lys eluted at 8.6 min. UV: λ_{max} 260nm, λ_{min} 280 nm (pH 4.9); ESI^+-MS/MS (dG-Pt-Lys): *m/z* 641.2 $[M]^+ \rightarrow$ *m/z* 624.2 $[M–NH_3]^+$, *m/z* 508.1 $[M–NH_3–$ deoxyribose+H$]^+$, and *m/z* 357.1 $[M–NH_3–$deoxyguanosine$]^+$.

Denaturing PAGE of Cisplatin-Induced DNA-Protein Cross-links **-** DNA 18-mer, 5′-GGA GCT GGT GGC GTA GGC-3′ (200 pmol), was 5′-end-labeled with ^{32}P in the presence of [γ-^{32}P]ATP and T4 polynucleotide kinase by standard methods *(122)*, purified by 12% denaturing PAGE, and

desalted by size exclusion. The ^{32}P-labeled duplex (0.93 nmol) was incubated with O^6-alkylguanine DNA alkyltransferase (AGT) (2.0 μg) or human recombinant glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (2.0 μg) in the presence of 1–100 mol equiv of cisplatin (1, 5, 10, and 20 nmol, respectively) for 3 h at 37 °C. The reaction mixtures were separated by 12% SDS-PAGE, and the radiolabeled products were visualized using a Packard Cyclone Phosphoimager (Packard BioScience, Meridan, CT).

Reaction of Synthetic dG-Pt-Cl monoadduct with recombinant AGT Protein - Human recombinant AGT protein (50 μg, 2.3 nmol) was incubated with 10 equiv of synthetic dG-Pt-Cl (23 nmol) in 10 mM Tris-HCl buffer (pH 7.4) for 4 h at 37^{o}C. Any unreacted dG-Pt-Cl was removed by size exclusion chromatography using Micro Bio-Spin 6 columns (Bio-Rad, Hercules, CA), in which the buffer was exchanged to 0.1% TFA in H_2O following manufacturer's instructions. The cross-linked protein was isolated by HPLC using an Agilent 1100 HPLC system equipped with a DAD UV detector. Agilent Zorbax 300 SB-C3 column (2.1 x 150 mm, 5 μm) was eluted with 0.1% TFA in water (A) and 0.1% TFA in 4:1 acetonitrile/water (B) at a flow rate of 0.4 mL/min under 10^{o}C. The solvent composition began at 0% B and was linearly changed to 15% B over 15 min and further to 50% from 15 to 25 min. The column was washed at 50% B for 15 min and re-equilibrated to 0% B for 2 min. Under these conditions, the modified AGT proteins eluted as a single peak at ~ 30 min, and the unreacted AGT protein eluted as a single peak at ~ 32 min. HPLC peaks containing modified AGT protein was collected and dried under vacuum, followed by HPLC-ESI$^+$-MS analysis as described below.

Trans-platination Reactions of dG-Pt-AGT in the Presence of dG - HPLC-purified cisplatin AGT-dG cross-links (~25 μg) were incubated with 10 molar equivalent of dG (~1.1 nmol) in SDS buffer (1% SDS, 10 mM Tris-HCl, 1% glycerol, 50 μM EDTA, pH 8.5) at 70 ^{o}C for 10 min or 60 min. Control experiment was conducted in the absence of dG at 70 ^{o}C for 10 min. Proteins were

subjected to size exclusion as described above, followed by HPLC-ESI$^+$-MS analysis. To detect *cis*-1,1-diammine-2,2-*bis*-(2'-deoxyguanosine-7-yl)-platinum (II) (dG-Pt-dG) cross-links, the solution was passed through Amicon Ultra-0.5 mL Centrifugal Filters (10K MWCO, Millipore, Temecula, CA) to remove proteins prior to HPLC-ESI$^+$-MS/MS analysis as described below.

Tryptic Digestion of Platinated AGT - Control or platinated AGT protein (~ 50 μg) was digested with trypsin (2 μg) in 25 mM ammonium bicarbonate buffer (pH 7.9) for 8 h at 37 ^{o}C. Samples were dried, desalted by ZipTip C18 purification (ZipTip C18 Pipette Tips, Millipore, Temecula, CA), and finally reconstituted in 0.1% formic acid (25 μL) prior to MS analysis as described below.

Total Digestion of Platinated AGT to Amino Acids - Tryptic peptides (from ~50 μg of protein) were filtered through Microcon YM-10 membrane filters to remove trypsin. Proteinase K (20 μg) was added to the filtrate, and proteolysis proceeded at 37 ^{o}C for 24 h. Samples were dried and subjected to off-line HPLC separation using an Agilent Technologies HPLC system (1100 model) incorporating a diode array detector and a Supelcosil LC-18-DB (4.6 x 250 mm, 5 μm) column (Sigma-Aldrich, St. Louis, MO). The column was eluted at a flow rate of 1 mL/min using 15 mM ammonium acetate, pH 4.9 (A) and acetonitrile (B). The solvent composition was changed linearly from 0 to 24% B over 24 min and further to 60% B in 6 min. HPLC fractions containing dG-Pt-Lys (5-7 min) were collected, dried under vacuum, and reconstituted in 25 μL of 15 mM ammonium acetate buffer for HPLC-ESI$^+$-MS/MS analysis.

Mass Spectrometry - HPLC-ESI$^+$-MS analysis of modified AGT proteins was performed with an Agilent 1100 capillary HPLC-ion-trap MS system operated in ESI$^+$ mode (*m/z* 200-2000). For whole protein mass spectrometry, chromatography was achieved using an Agilent Zorbax Extend SB 300-C8 column (150 x 0.3 mm, 3.5 μm) eluted at a flow rate of 12 μL/min with a mobile

phase of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The solvent composition was held at 30% B for the first 5 min, followed by a linear increase to 80% B over 25 min, and further to 95% B in 5 min. Using these conditions, dG-Pt-Cl-modified AGT proteins (dG-Pt-AGT) eluted ~14.5 min. Deconvolution of the protein charge envelope was performed using the commericial deconvolution software.

AGT tryptic peptides were analyzed by HPLC-ESI$^+$-MS/MS with a Thermo Scientific LTQ Orbitrap Velos mass spectrometer in line with an Eksigent NanoLC-Ultra 2D HPLC system, a nanospray source, and Xcalibur 2.1.0 software for instrument control. Peptide mixtures (8 µL) were loaded on a Symmetry C18 trapping column (180 µm x 20 mm, Waters, Milford, MA) using 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B) at a flow composition of 95% A and 5% B at 5 µL/min for 3 minutes. Following trapping, the flow was reversed and decreased to 0.3 µL/min. The peptides were eluted off the trap column and onto a capillary column (75 µm ID, 10 cm packed bed, 15 µm orifice) created by hand packing a commercially purchased fused-silica emitter (New Objective, Woburn MA) with Zorbax SB-C18 5 µm separation media (Agilent, Santa Clara, CA). The gradient program started at 5% B, followed by a linear increase to 60% B over 60 min, and further to 95% B in 5 min. Liquid chromatography was carried out at an ambient temperature. Centroided MS-MS scans were acquired using an isolation width of 2.5 *m/z*, an activation time of 30 ms, an activation Q of 0.25, 35% normalized CID collision energy, and 1 microscan with a max ion time of 100 ms for each MS/MS scan. The mass spectrometer was calibrated prior to each analysis, and the spray voltage was adjusted to assure a stable spray. Typically, the tune parameters were as follows: spray voltage of 1.6 kV, a capillary temperature of 275 °C, and an S-lens RF Level of 50%. Peptide MS/MS spectra were collected using data-dependent scanning in which one full scan mass spectrum was followed by eight MS/MS spectra. Dynamic exclusion was enabled for 60 s and singly charged species were excluded.

Spectral data were analyzed using Thermo Proteome Discoverer 1.2 (ThermoScientific,

San Jose, CA) that linked raw data extraction, database searching, and probability scoring. The raw data were directly uploaded, without any format conversion, to search against human AGT protein FASTA database (http://www.uniprot.org/uniprot/P16455) combined with its reversed counterpart by using the SEQUEST algorithm *(123;124)*. Search parameters included trypsin specificity and up to 2 missed cleavage sites. The dG-Pt-AGT conjugates were expected to experience fragmentation at the ESI^+ source by a loss of one- or two amino groups (due to the fragile nature of coordination between platinum and $–NH_3$ ligands) or a loss of 2-deoxyribose, as well as the possible breakage of the glycosodic bond of dG. Furthermore, dG-Pt-Cl -induced platination at the *N*-donor residues (*N*-terminus, histidine, lysine, or arginine), *S*-donor residues (cysteine or methionine), and *O*-donor residues (threonine, tyrosine, aspartic acid, or glutamic acid) were specified as the following dynamic modifications to identify spectra of modified peptides: (A) cross-link to dG: +495.1063 Da ($dG+Pt+2NH_3$), +478.0795 Da ($dG+Pt+NH_3$, *a loss of* $-NH_3$), or +461.0533 Da ($dG+Pt$, *a loss of* $2\text{-}NH_3$); (B) cross-link to guanine: +379.0589 Da ($Gua+Pt+2NH_3$), +362.0324 Da ($Gua+Pt+NH_3$, *a loss of* $–NH_3$), or +345.0058 Da ($Gua+Pt$, *a loss of* $2\text{-}NH_3$). Stringent criteria were implemented for all of the identified peptides: mass tolerance ≤ 100 ppm, Xcorr ≥ 2.5, ΔCn ≥ 0.5, and RSp (preliminary score rank) ≤ 3. Peptides sequences with MS/MS spectra not obeying with these criteria were removed from the final target list. Platinated peptides were included in the final lists only if they exhibited good quality MS/MS spectra (at least 40% of the observed MS/MS ions should match the theoretical b^+ or y^+-type peptide fragment ions) and with both platinated and non-platinated fragment ions whose relative abundances were unambiguously higher than the baseline.

HPLC-ESI^+-MS/MS of dG-Pt-Lys conjugates was conducted with a Thermo-Finnigan TSQ Vantage mass spectrometer in line with an Eksigent MicroAS autosampler and nanoLC 2D HPLC pump, a heated ESI source, and an Xcalibur 1.4 software for instrument control. Chromatographic separation was accomplished using a Hypercarb column (100 mm x 0.5 mm, 3 μm, ThermoScientific, Waltharm, MA) eluted with a gradient of 15 mM ammonium acetate (A)

and 1:1 acetonitrile:water with 1% formic acid (B) at a flow rate of 13 μL/min. The gradient program began at 2% B, followed by a linear increase to 8% B in 10 min, further to 80% B in 18 min, and finally back to 2% B in 2 min. Using this gradient, dG-Pt-Lys eluted at ~14.5 min. ESI was achieved at a spray voltage of 3.2 kV and a capillary temperature of 200°C. CID was performed with Ar as a collision gas (1.0 mTorr) at a collision energy of 25V. The MS parameters were optimized for maximum response during infusion of a standard solution of dG-Pt-Lys and may vary slightly between different experiments. HPLC-ESI^+-MS/MS analyses were performed in the selected reaction monitoring (SRM) mode using the transition corresponding to major fragment ions observed upon CID fragmentation of dG-Pt-Lys in a triple quadrupole mass spectrometer (*m/z* 641.3 $[M]^+ \rightarrow$ 508.2 $[M–NH_3–deoxyribose+H]^+$, and 340.1 $[M–2NH_3–deoxyguanosine]^+$).

HPLC-ESI^+-MS/MS analysis of dG-Pt-dG cross-links was performed using an Agilent 1100 series capillary LC Ion Trap MS system operated in the ESI^+ mode. Auto MS^2 was used to isolate and fragment the $[M]^+$ ion of dG-Pt-dG (*m/z* 762.3). Chromatographic separation was achieved using a Zorbax SB-C18 column (150 mm × 0.5 mm, 5 μm) eluted at a flow rate of 15 μL/min. The mobile phase consisted of 15 mM ammonium acetate, pH 4.9 (A) and acetonitrile (B). The solvent composition was held with a linear gradient of 0-15% B over the course of 30 min.

2.3 Results

2.3.1 SDS-PAGE Analysis of Cisplatin-Induced DNA-Protein Cross-Links

To examine the ability of cisplatin to cross-link proteins to DNA, recombinant O^6-alkylguanine DNA alkyltransferase (AGT) or glyceraldehyde-3-phosphate dehydrogenase (GAPDH) protein was incubated with 5'-^{32}P-end labeled DNA duplexes in the presence of increasing amounts of cisplatin, followed by SDS-PAGE analysis. The results presented in Figure 2.1 reveal that a slow-migrating species corresponding to protein-DNA conjugates was observed

when duplex DNA was exposed to cisplatin in the presence of AGT or GAPDH. No such product was formed in control experiments in which either protein or oligonucleotide was omitted (Lanes 1-3, Figures 2.1A and 2.1B). Cross-linking of AGT to DNA by cisplatin displayed a pronounced concentration dependence, with 2-5% of recombinant protein being cross-linked to DNA (Lanes 4-8, Figure 2.1A). Additional low-mobility bands were observed on the top of the gel, probably corresponding to cross-linking of multiple proteins to DNA.

2.3.2 HPLC-ESI$^+$-MS Analysis of dG-Pt-Cl Monoadduct Induced AGT-dG Cross-Links: Whole Protein Results

Initial platination of DNA by cisplatin leads to the formation of DNA monoadducts, which retain one reactive Pt functionality (Scheme 1.3). These monoadducts can then react with nucleophilic sites within proteins, giving rise to DNA-protein cross-links. To gain an insight into the nature of cisplatin-induced AGT-DNA linkages, recombinant AGT protein was incubated with synthetic *cis*-1,1-diammine-2-chloro-2-(2'-deoxyguanosine-7-yl)-platinum (II) (dG-Pt-Cl) as a model of monoplatinated DNA (4 h at 37 ^{o}C). Following size exclusion chromatography to remove the bulk of unreacted dG-Pt-Cl, AGT-dG cross-links were separated from unmodified AGT by reverse phase HPLC (Scheme 2.1 and Figure 2.2). Three HPLC peaks were observed (Figure 2.2). HPLC-ESI$^+$-MS/MS analyses revealed that the first peak eluting at 4.5 min corresponded to unreacted dG-Pt-Cl, the second peak (30.1 min) contained AGT-dG cross-links, and the third peak (32.1 min) was due to unreacted AGT protein (Figures 2.2 and 2.3). ESI$^+$ spectrum of unreacted AGT contains multiple *m/z* signals corresponding to the various charge states of the protein (+17 - +37) (Figure 2.3A). Deconvolution of the mass spectrum results in a molecular weight of 21,878 Da, which matches the theoretical value of 21,877 Da (Figure 2.3A). A similar spectrum was observed for the peak eluting at 32.1 min (Figure 2.2), with the exception of a small additional signal at 22,355 Da corresponding to AGT protein containing a single cisplatin cross-link to dG ($AGT+Pt+dG+NH_3$, M = 22355 Da) (Figure 2.3B). ESI$^+$ analysis of the

Figure 2.1 Detection of cisplatin-induced DNA-protein cross-links by gel electrophoresis. 12% SDS-PAGE analysis of ^{32}P-endlabeled DNA duplexes (5'-GGA GCT GGT GGC GTA GGC-3' +strand) following incubation with (**A**) recombinant human AGT or (**B**) recombinant GAPDH protein in the presence of 1 (lane 4), 5 (lane 5), 10 (lane 6), 50 (lane 7) or 100 (lane 8) molar equivalents of cisplatin. Free duplex DNA (labeled "Free oligo") migrates to the bottom of the gel, whereas DNA-protein cross-links display substantially reduced mobility.

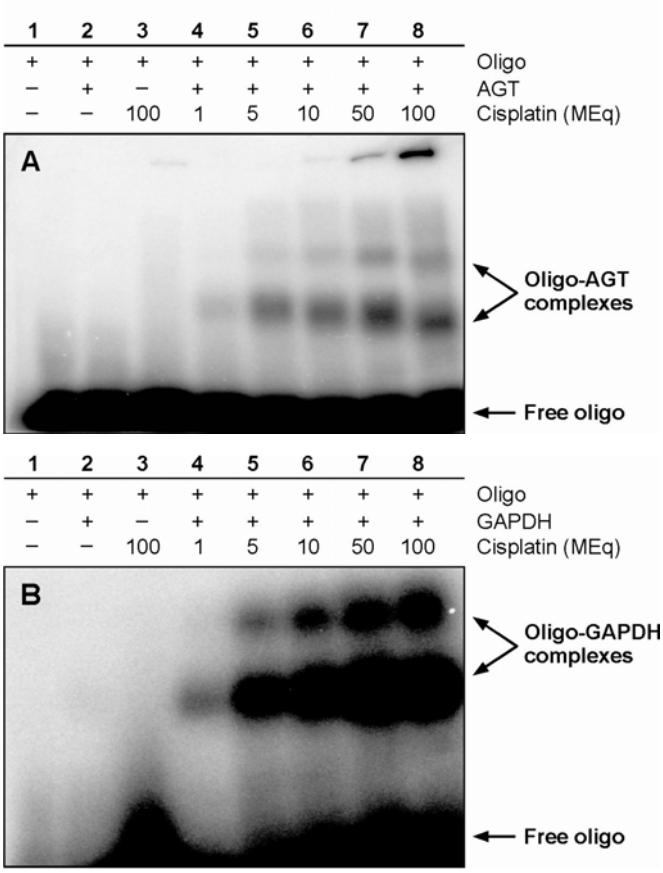
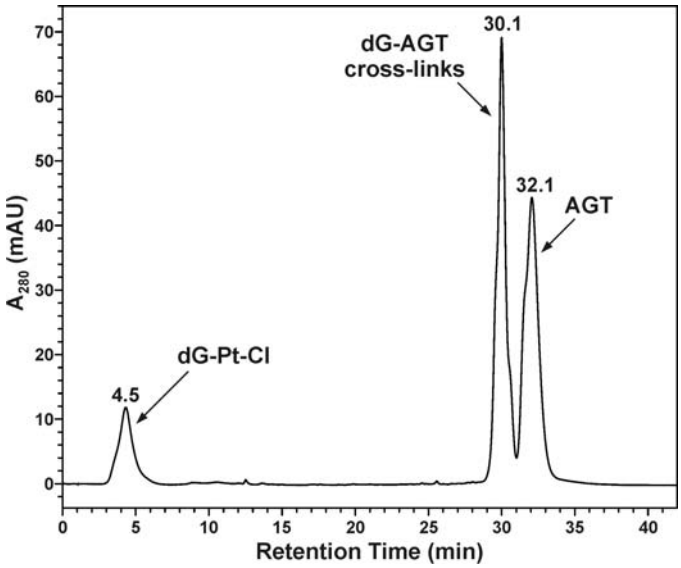


Figure 2.2. HPLC separation of reaction mixtures following incubation of recombinant AGT protein with dG-Pt-Cl as models for monoalkylated DNA to induce cross-linking. AGT-dG conjugates were identified by HPLC-ESI$^+$-MS as shown in Figure 2.3.



material eluting at 30.1 min (Figure 2) reveals the presence of AGT containing a single platinum cross-link to dG (*AGT+Pt+dG+2NH$_3$*, M = 22372 Da), two platinum cross-links to dG (*AGT+2Pt+2dG+2NH$_3$*, *M* = 22832 Da), and a triple platinum cross-link to dG (*AGT+3Pt+3dG+4NH$_3$*, *M* = 23328 Da) (Figure 2.3C). Taken together, those results indicate that AGT protein can be covalently modified by dG-Pt-Cl to form cross-links at at least three different sites of the protein. Our ability to separate the structurally modified AGT protein from the unreacted AGT has facilitated the further investigation of the specific sites within AGT protein involved in cross-linking to dG (see below).

2.3.3 Peptide Mapping by HPLC-ESI$^+$-MS/MS

Our ESI$^+$ MS results for dG-Pt-Cl-treated AGT protein detected AGT species containing one-, two-, or three- platinum cross-linked to dG (Figure 2.3C), suggesting the presence of at least three distinct cross-linking sites within the protein. Further insight into the identities of AGT amino acid residues responsible for reaction with dG-Pt-Cl was provided by HPLC-ESI$^+$-MS/MS analysis of tryptic digests of the platinated protein as described below.

Proteolytic digestion of native AGT protein with trypsin provides a good protein sequence coverage (82%, Table 2.1). Only one tryptic peptide representing amino acids 37-96 is not detected due to its size. To locate platinated sites within the protein, dG-Pt-Cl modified AGT (from HPLC peak at 30.1 min at Figure 2.2, see ESI MS spectrum in Figure 2.3C) was subjected to tryptic digestion, followed by nanoHPLC-nanospray ESI-MS/MS on an Orbitrap Velos mass spectrometer. Spectral data were analyzed using Thermo Proteome Discoverer 1.2 to search against human AGT protein FASTA database. Because the cross-linked moiety (*dG-Pt-2NH$_3$*) is likely to experience fragmentation in the ESI$^+$ source by a loss of one or two amino groups or a loss of 2-deoxyribose, multiple *m/z* values were specified for each target peptide: (A) cross-link to dG: +495.1063 Da (*dG+Pt+2NH$_3$*), +478.0795 Da (*dG+Pt+NH$_3$, a loss of -NH$_3$*), or +461.0533 Da (*dG+Pt, a loss of 2-NH$_3$*); (B) cross-link to guanine: +379.0589 Da

Figure 2.3. HPLC-ESI$^+$-MS and deconvoluted spectra (inset) of (**A**) unreacted AGT protein, (**B**) dG-Pt-Cl monoadduct-treated AGT protein that had been separated by HPLC with retention time 32.1 min in Figure 2.2, and (**C**) dG-Pt-Cl monoadduct-treated AGT protein that had been separated by HPLC with retention time 30.1 min in Figure 2.2.

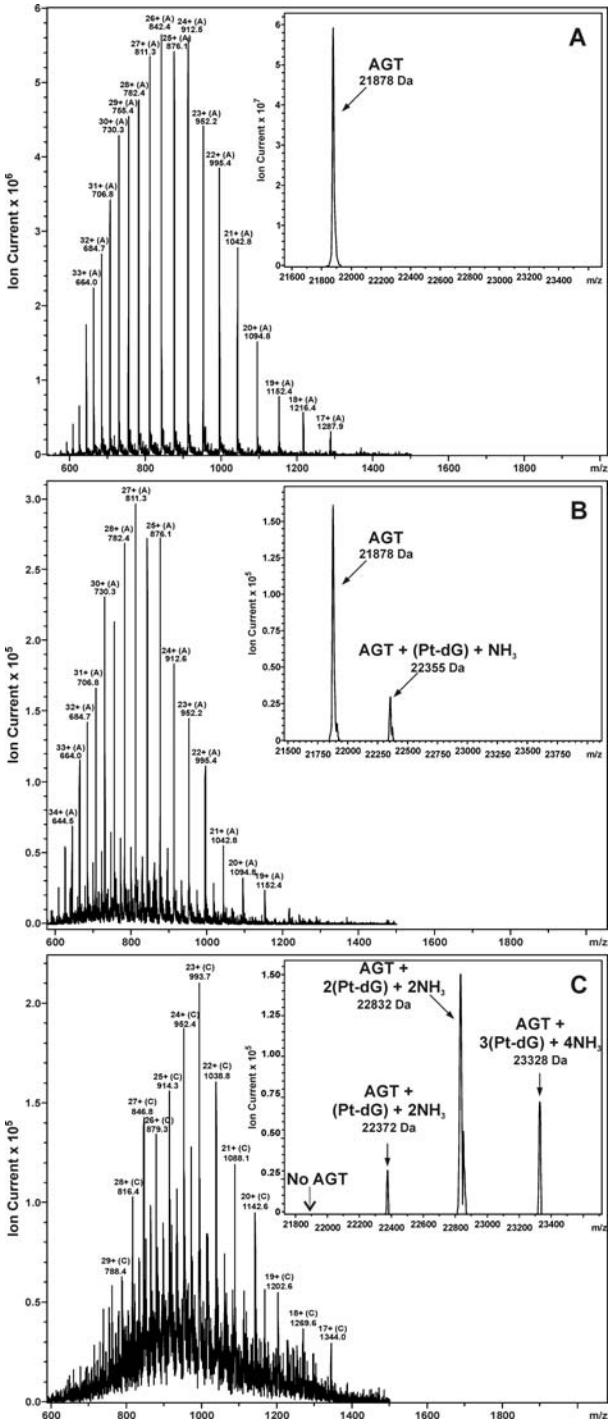


Table 2.1. HPLC-ESI^{+}-MS/MS analysis of AGT tryptic peptides (unmodified protein).

Position	Peptide sequence	$[M + H]^{+}_{calculated}$	$[M + 2H]^{2+}_{calculated}$	$[M + 3H]^{3+}_{calculated}$	Observed ions
1-8	MDKDCEMK	998.38	499.6939	333.4652	500.2
9-18	RTTLDSPLGK	1,086.60	543.8039	362.8718667	363.21, 544.31
10-18	TTLDSPLGK	930.5	465.7539	310.8385333	466.26
19-32	LELSGCEQGLHEIK	1,554.77	777.8889	518.9285333	519.27, 778.39
33-36	LLGK	429.28	215.1439	143.7652	215.43
37-96	GTSAADAVEVPAPAAVLGGPEPLMQCTAW LNAYFHQPEAIEEFPVPALHHPVFQQESFTR	6,468.13	3234.5689	2156.7152	ND*
97-101	QVLWK	673.40	337.2039	225.1385333	337.21
102-107	LLKVVK	699.51	350.2589	233.8418667	350.26
108-125	FGEVISYQQLAALAGNPK	1,905.00	953.0039	635.6718667	636.01, 935.51
126-135	AARAVGGAMR	959.52	480.2639	320.5118667	320.51, 480.26
129-135	AVGGAMR	661.35	331.1789	221.1218667	331.18
136-147	GNPVPILIPCHR	1,314.72	657.8639	438.9118667	658.37
148-165	VVCSSGAVGNYSGGLAVK	1,666.84	833.9239	556.2852	834.43
166-175	EWLLAHEGHR	1,246.61	623.8089	416.2085333	416.55, 624.32
176-193	LGKPGLGGSSGLAGAWLK	1,667.93	834.4689	556.6485333	566.99
194-207	GAGATSGSHHHHHH	1,428.60	714.8039	476.8718667	715.34

*Not detected.

(*Gua+Pt+2NH₃*), +362.0324 Da (*Gua+Pt+NH₃, a loss of –NH₃*), or +345.0058 Da (*Gua+Pt, a loss of 2-NH₃*). Furthermore, dG-Pt-Cl monoadduct-induced platination at the *N*-donor residues (*N*-terminus, histidine, lysine, or arginine), *S*-donor residues (cysteine or methionine), and *O*-donor residues (threonine, tyrosine, aspartic acid, or glutamic acid) were specified as potential dynamic modifications. The following stringent criteria were further implemented for all of the identified peptides: mass tolerance ≤ 100 ppm, Xcorr ≥ 2.5, ΔCn ≥ 0.5, and RSp (preliminary score rank) ≤ 3. Peptides sequences with MS/MS spectra not obeying with these criteria were removed from the final target list.

HPLC-ESI-MS/MS analyses detected a prominent, triply charged peptide at *m/z* 801.12 ($[M+3H]^{3+}$) corresponding to AGT residues F^{108}GEVISYQQLAALAGNPK^{125} containing a cisplatin-dG cross-link (calculated *M* = 2401.28, Δ*M* = 495.11 Da (*dG+Pt+2NH₃*), Table 2.2 and Figure 2.4A). When this peptide was subjected to collision-induced dissociation (CID), the resulting MS/MS spectrum was consistent with the presence of *dG- Pt-2NH₃* adduct at Glu^{110} (Table 2.2 and Figure 2.4A). While the masses of y_2, y_3, y_5-y_{10} and y_{12} fragments were in agreement with theoretical values for unmodified peptide, the mass of the y_{17} fragment was increased by 495.11 Da (observed *M* = 2255. 05 Da, calculated *M* = 1759.95 Da for the unmodified y_{17} peptide fragment). The b_3-b_7 and b_9-b_{15} fragments also contained the *dG-Pt-2NH₃* moiety as indicated by the 495.11 Da mass increase, suggesting the adduct resides at either the 2nd or the 3rd residue of the peptide (G^{109} or E^{110}). Since Gly is a non-nucleophilic residue, these results are suggestive of Glu^{110} participation in AGT-dG cross-linking by cisplatin.

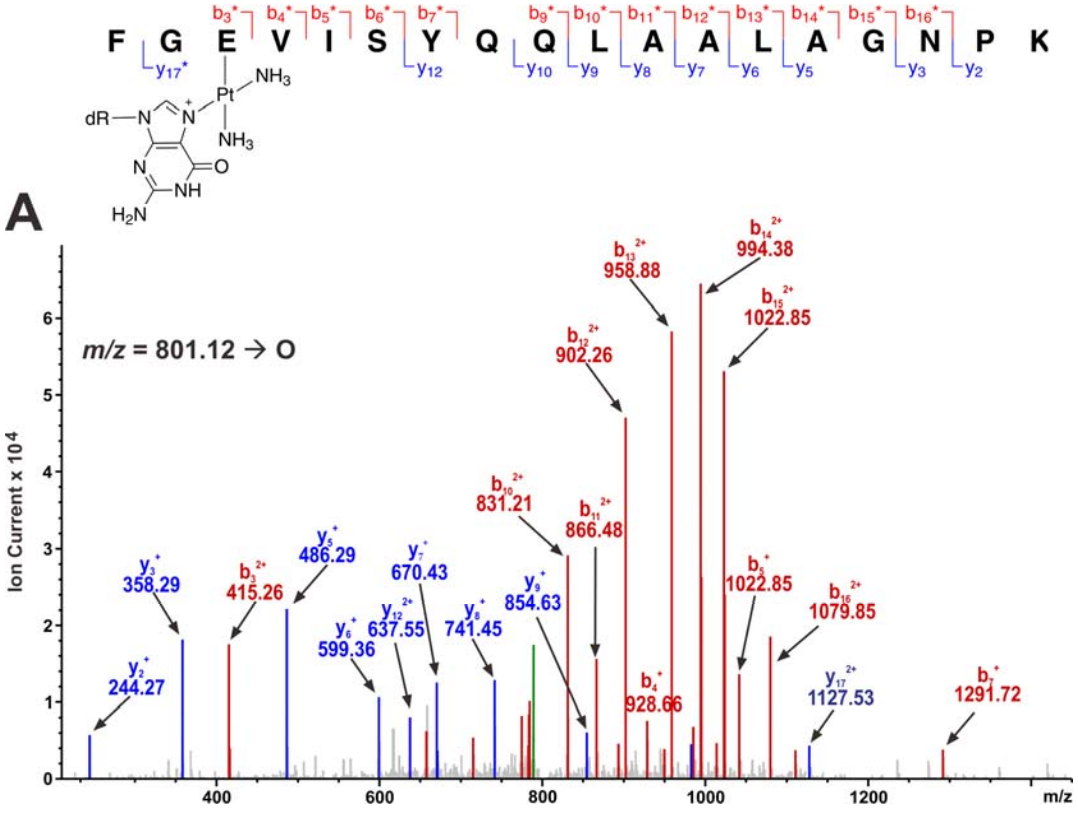
The second cross-linking site was located upon detection of the peptide F^{108}GEVISYQQ LAALAGNPKAAR^{128} containing a single platinum-dG adduct (*m/z* 850.12 $[M+3H]^{3+}$; calculated *M* = 2547.35 Da, Δ*M* = 344.00 Da (*Gua+Pt*)) (Table 2.2 and Figure 2.4B). The cross-linking site was mapped to Lys^{125} based on the MS/MS fragmentation patterns, especially the diagnostic y_3, y_4 and b_{18} ions at *m/z* 317.18, 395.15 and 1116.51, respectively (Figure 2.4B). The y_3 ion mass

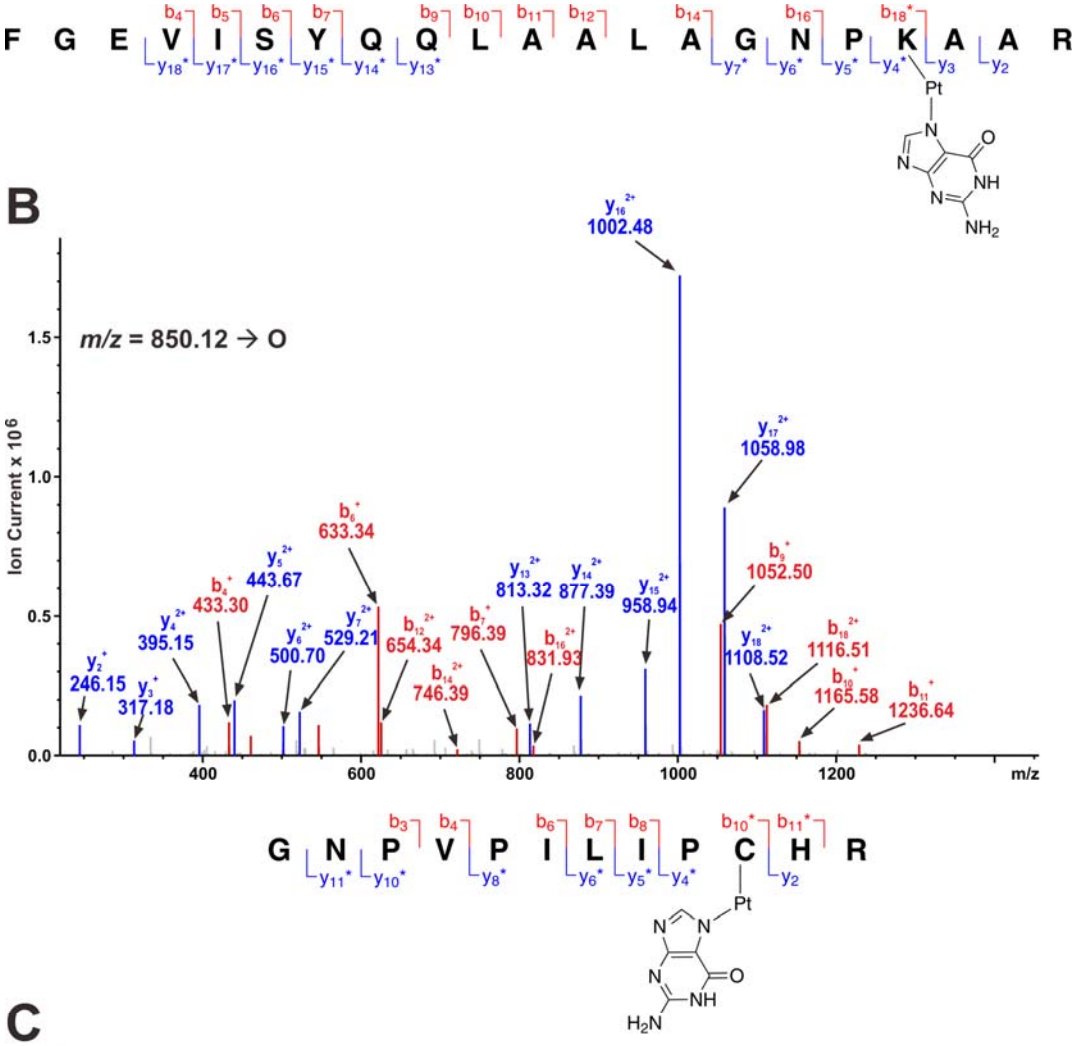
Table 2.2. Platinated peptide sequences of AGT after the treatment with dG-Pt-Cl monoadduct followed by HPLC-ESI$^+$-MS/MS analysis of the resulting tryptic digests. The assigned platinum binding site is given as bold italic type within peptide sequence.

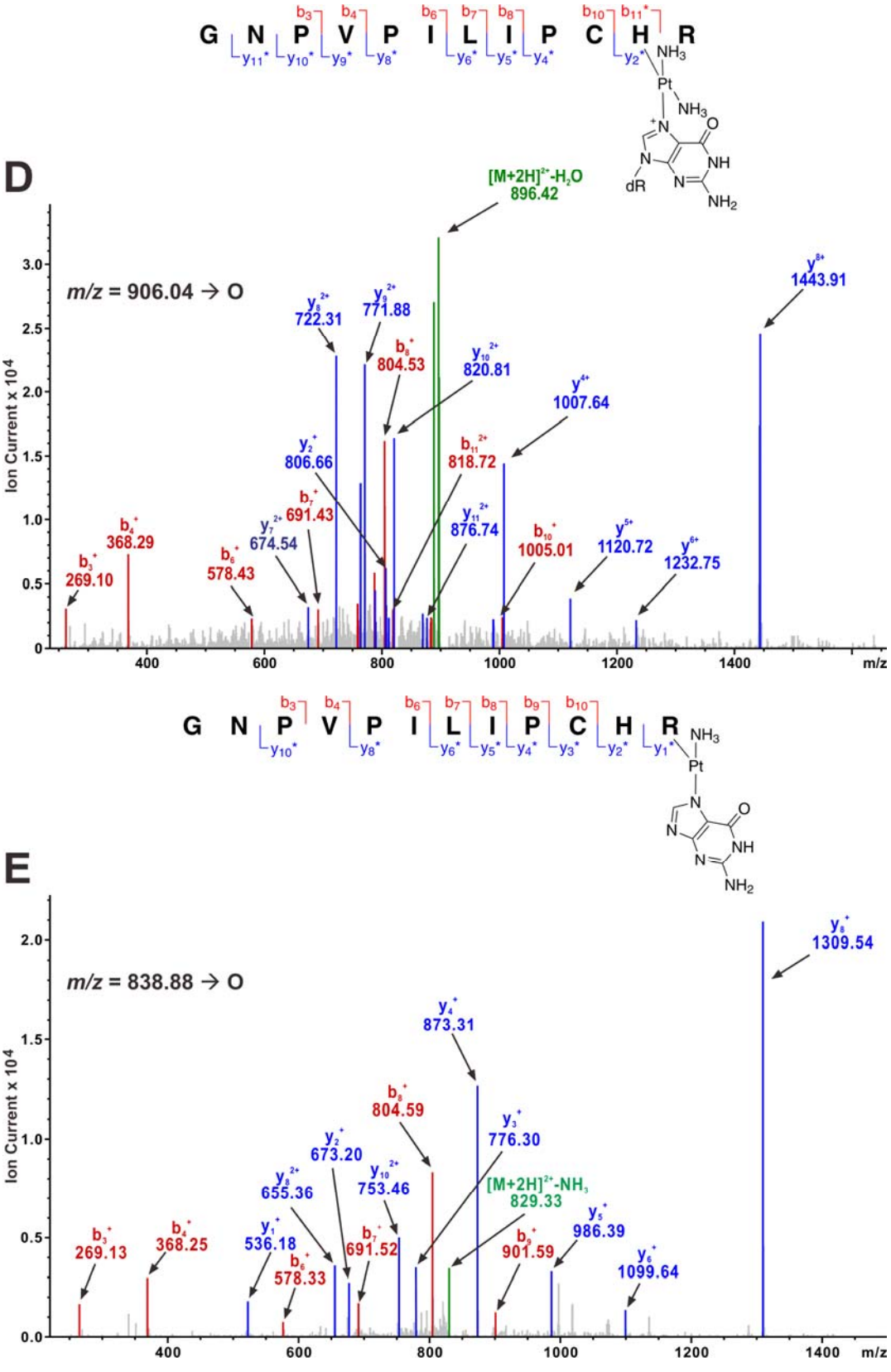
Peptide sequence	Pt adduct	Pt adduct mass (Da)	Charge	SEQUEST parameters		Ions*	MS/MS spectrum
				Xcorr	ΔCn		
F^{108}G***E***VISYQQLAALAGNPK^{125}	+(Pt-dG-2NH$_3$)	495.11	+3	4.53	1.0	34/68	Figure 2.4A
F^{108}GEVISYQQLAALAGNP***K***AAR^{128}	+(Pt-Gua)	344.00	+3	3.39	1.0	28/60	Figure 2.4B
G^{136}NPVPILIP***C***HR^{147}	+(Pt-Gua)	344.00	+2	2.54	1.0	22/44	Figure 2.4C
G^{136}NPVPILIPC***H***R^{147}	+(Pt-dG-2NH$_3$)	495.11	+2	3.07	1.0	20/50	Figure 2.4D
G^{136}NPVPILIPCH***R***^{147}	+(Pt-Gua-NH$_3$)	361.03	+2	3.05	1.0	16/35	Figure 2.4E
V^{148}V***C***SSGAVGNYSGGLAVK^{165}	+(Pt-Gua-2NH$_3$)	378.05	+3	4.95	1.0	30/68	Figure 2.4F

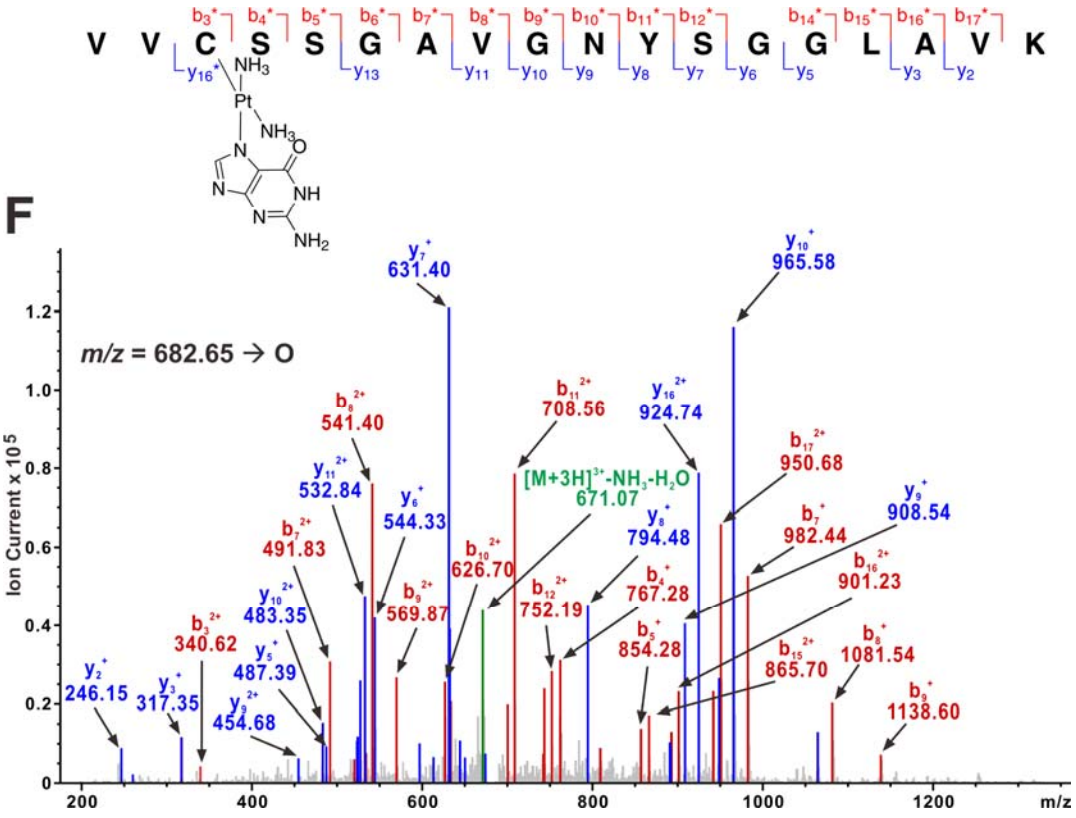
*Ratio of the assigned *b* and *y* ions to the total number of all possible fragment ions.

Figure 2.4. HPLC-ESI$^+$-MS/MS analysis of tryptic peptides derived from cisplatin-induced AGT-dG conjugates. Fragment ions containing cross-linked platinum adduct are indicated by "*". (**A**). MS/MS spectrum of hAGT tryptic peptide F^{108}GEVISYQQLAALAGNPK^{125} containing a dG-Pt-2NH$_3$ adduct on Glu^{110}. (**B**). F^{108}GEVISYQQLAALAGNPKAAR^{128} containing a cisplatin-guanine adduct on Lys^{125}. (**C**) G^{136}NPVPILIPCHR^{147} containing a cisplatin-guanine adduct on Cys^{145}. (**D**) G^{136}NPVPILIPCHR^{147} containing a dG-Pt-2NH$_3$ adduct on His^{146}. (**E**) G^{136}NPVPILIPCHR^{147} containing a Gua-Pt-NH$_3$ adduct on Arg^{147}. (**F**) V^{148}VCSSGAVGNYSGGLAVK^{165} containing a Gua-Pt-2NH$_3$ adduct on Cys^{150}.




F G E V I S Y Q Q L A A L A G N P K A A R
Pt
B
m/z = 850.12 → O
Ion Current x 10^6
m/z
y16 2+ 1002.48
y17 2+ 1058.98
b6+ 633.34
y5 2+ 443.67
b4+ 433.30
y4 2+ 395.15
y7 2+ 529.21
y6 2+ 500.70
b12 2+ 654.34
b14 2+ 746.39
b7+ 796.39
y13 2+ 813.32
y14 2+ 877.39
b16 2+ 831.93
y15 2+ 958.94
b9+ 1052.50
y18 2+ 1108.52
b18 2+ 1116.51
b10+ 1165.58
b11+ 1236.64
y2+ 246.15
y3+ 317.18
G N P V P I L I P C H R
Pt
C
m/z = 830.49 → O
Ion Current x 10^4
m/z
y8 2+ 646.58
[M+2H]2+-NH3-H2O 812.57
y10 2+ 744.70
y4+ 855.54
y8+ 1291.75
b10 2+ 674.54
b11 2+ 743.82
b4+ 368.27
y2+ 312.32
b7+ 691.48
y11 2+ 802.45
y5+ 968.61
y6+ 1081.64
b3+ 269.20
b6+ 578.34
b10+ 1347.76


D
G N P V P I L I P C H R
[M+2H]2+-H2O
896.42
m/z = 906.04 → O
Ion Current x 10^4
m/z
E
G N P V P I L I P C H R
m/z = 838.88 → O
Ion Current x 10^4
m/z


F
V V C S S G A V G N Y S G G L A V K
m/z = 682.65 → O
Ion Current x 10^5
m/z
y7+ 631.40
y10+ 965.58
b8 2+ 541.40
b11 2+ 708.56
y16 2+ 924.74
b17 2+ 950.68
y11 2+ 532.84
[M+3H]3+-NH3-H2O 671.07
y9+ 908.54
y6+ 544.33
y8+ 794.48
b7+ 982.44
b7 2+ 491.83
b10 2+ 626.70
b16 2+ 901.23
y10 2+ 483.35
b12 2+ 752.19
b4+ 767.28
b9 2+ 569.87
b3 2+ 340.62
b5+ 854.28
b8+ 1081.54
y5+ 487.39
b15 2+ 865.70
y2+ 246.15
y3+ 317.35
y8 2+ 454.68
b9+ 1138.60

matched the theoretical value for unmodified peptide, while the y_4 and b_{18} both experienced a +344.06 Da mass shift corresponding to guanine-platinum adduct, which suggested that the cross-linking took place at Lys^{125}.

Importantly, significant cross-linking was observed at three active site residues directly involved in AGT catalysis: Cys^{145} (Figure 2.4C), His^{146} (Figure 2.4D), and Arg^{147} (Figure 2.4E). While Cys^{145} serves as alkyl acceptor during AGT repair reaction, His^{146} acts as a base to deprotonate Cys^{145}. Arg^{147}, together with Pro^{144} and Leu168, provides a hydrophobic slot for His^{146}. In this hydrophobic environment, His^{146} accepts a hydrogen bond from water and donates one to the negatively charged Glu^{172} carboxylate, which pairs with Arg^{146} within a salt bridge *(125)*. HPLC-ESI^+-MS/MS analyses revealed doubly charged ions corresponding to three different guanine-cisplatin cross-link-containg G^{136}NPVPILIPCHR^{147} peptides (*m/z* 830.49 $[M+2H]^{2+}$, calculated M = 1658.91 Da, ΔM = 344.00 Da (*Gua+Pt*); *m/z* 906.04 $[M+2H]^{2+}$, calculated M = 1810.07 Da, ΔM = 495.11 Da ($dG+Pt+2NH_3$); and *m/z* 838.88 $[M+2H]^{2+}$, calculated M = 1675.74 Da, ΔM = 361.03 Da ($Gua+Pt+NH_3$), Table 2.2). MS^2 sequencing of the first doubly charged ion (*m/z* 830.49 $[M+2H]^{2+}$) yielded characteristic product ions corresponding to y_2, y_4, and b_{10}, which were instrumental in the assignment of the modification site within this peptide (Figure 2.4C). Among them, b_{10} and y_4 both experienced a ΔM = 344.02 Da (*Gua+Pt*) mass shift, indicating that guanine-platinum cross-linking took place at Pro^{144} or Cys^{145}. Since Pro^{144} is not likely to be modified by platinum, Cys^{145} has been assigned as the reactive binding site for cisplatin. The next adjacent residue His^{146} was similarly identified as a reactive binding site for cross-linking from the diagnostic b_{10}, b_{11}, and y_2 ions (Figure 2.4D). CID of the doubly charged ion of the peptide (*m/z* 838.88 $[M+2H]^{2+}$, calculated M = 1675.74 Da, ΔM = 361.03 Da ($Gua+Pt+NH_3$)) produced an MS/MS spectrum containing *b*- (b_3, b_4, and b_6-b_{10}) and *y*-series ions (y_1-y_6, y_8, and y_{10}), consistent with the presence of a guanine-platinum adduct at Arg^{147}.

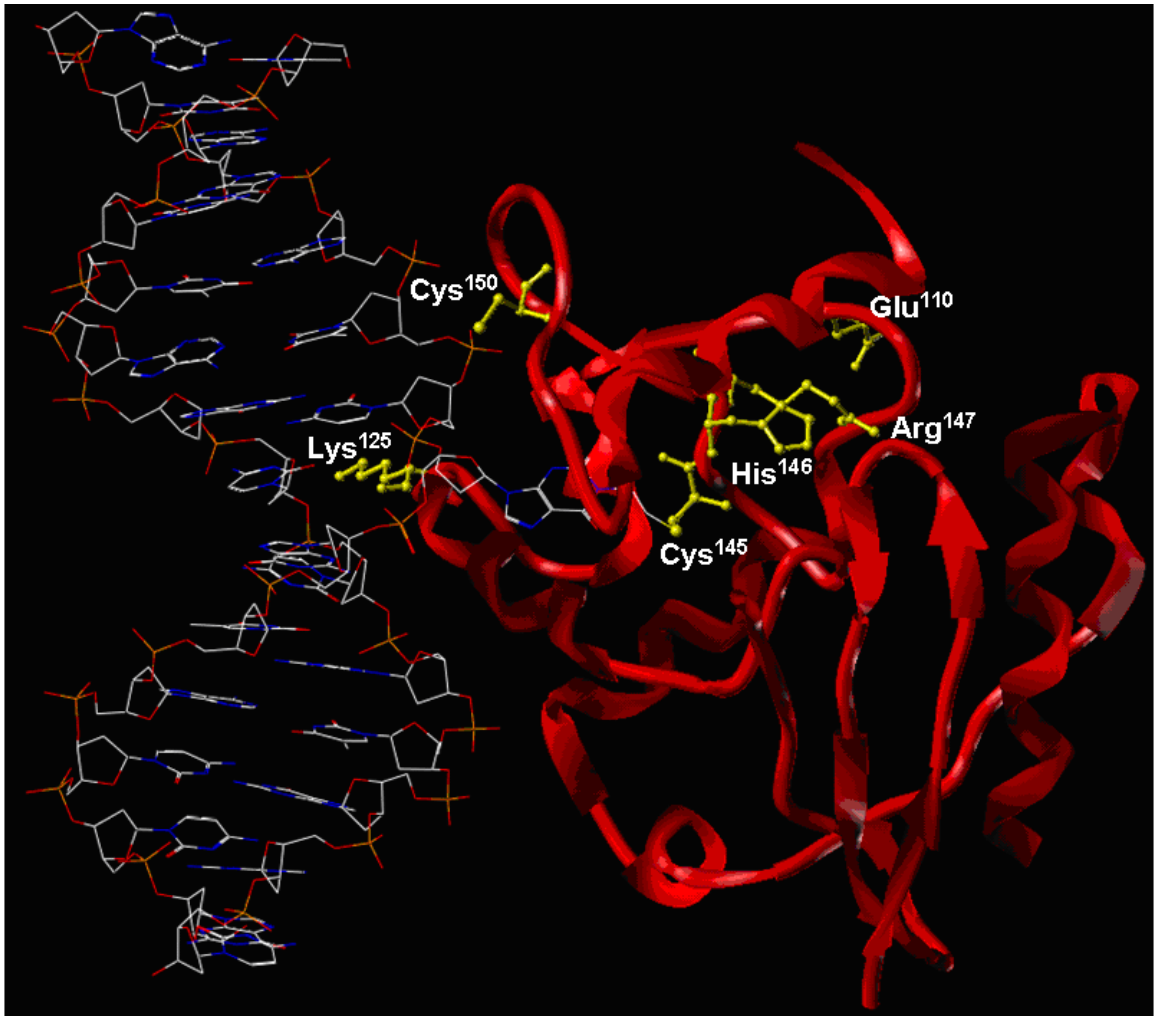
Figure 2.4F depicts the CID-MS^2 spectrum of the ions at *m/z* 682.65 ($[M+3H]^{3+}$), corresponding to the platinated peptide $V^{148}VCSSGAVGNYSGGLAVK^{165}$ containing a $Pt(NH_3)_2$-guanine adduct. The observed *b*- and *y*-series ions y_2-y_3, y_5-y_{11}, and y_{13} ions are in agreement with the fragmentation of unmodified peptide, whereas *m/z* 340.62 (b_3^{2+}) and 924.74 (y_{16}^{2+}) both experience +378.10 Da ($Gua+Pt+2NH_3$) mass increase (Figure 2.4F), making it possible to assign the modification site to Cys^{150} of the AGT protein.

Taken together, our nanoHPLC-nanospray MS/MS results for tryptic digests of platinated AGT reveal a total of six dG-Pt-Cl binding sites: Glu^{110}, Lys^{125}, Cys^{145}, His^{146}, Arg^{147}, and Cys^{150} (Table 2.2 and Figure 2.4). A crystal structure of human AGT protein bound to DNA is shown in Figure 2.5. As is apparent from the crystal structures, Cys^{145}, His^{146}, and Arg^{147} are located directly in the AGT active site pocket ($I^{143}PCHRV^{148}$). The side chain of Glu110 is in a close proximity to the protein active site. Lys^{125} is located at the DNA-binding domain of the protein, which interacts with double stranded DNA *via* helix-turn-helix (HTH) motif. Mutations at Lys^{125} (i.e. Lys125Ala) have been previously shown to significantly disrupt the interaction of wild type AGT with DNA, indicating Lys^{125} is essential for DNA binding *(115;125)*.

2.3.4 Platination Migration from DNA-Pt-Protein to DNA-Pt-DNA

One interesting property of platinum-induced adducts is that the Pt-S and Pt-N coordination bonds are potentially reversible, making it possible to observe "platination migration" from one nucleophilic site within biomolecules to another *(126-129)*. For example, Reediji and colleagues employed synthetic *S*-guanosyl-L-homocystine (sgh) as a model compound to examine intramolecular migration of platinum from cysteine thiol to the N7 of guanosine *(126)*. More recently, Sadler's group investigated the intermolecular displacement of Pt-S bound in $Pt(dien)^{2+}$ model compound by N7-guanosine 5'-monophosphate *(127)*. In 2000, Reediji *et al* further demonstrated that the sulfur atom of the platinum-thioether adduct can be

Figure 2.5 Crystal structure of human AGT protein bound to DNA (PDB 1T39) showing the sites of cisplatin-mediated DNA-protein cross-linking.



substituted by guanine nucleobase within oligonucleotides *(128)*.

To investigate the possibility that AGT-DNA cross-links may undergo platination migration to form guanine-guanine cross-link and to release the intact AGT protein, HPLC-purified dG-AGT conjugates (dG-Pt-AGT) were incubated with excess dG at 70 ^{o}C for varying periods of time, followed by capillary HPLC-ESI^{+}-MS analysis of the products (Scheme 2.1). HPLC-ESI^{+}-MS analysis of modified proteins confirmed that AGT was completely modified by dG-Pt-Cl, with no unreacted AGT protein present (Figure 2.3C). Deconvoluted ESI^{+} MS spectra reveal AGT protein bearing one (*AGT+Pt+dG+NH$_3$*, *M* = 22355 Da), two (*AGT+2Pt+2dG+2NH$_3$*, *M* = 22832 Da) or three-platinum adducts cross-linked to dG (*AGT+3Pt+ 3dG+4NH$_3$*, *M* = 23328 Da) (Figure 2.3C). In contrast, HPLC-ESI^{+}-MS analysis of dG-Pt-AGT cross-links incubated in the presence of free dG revealed the presence of free AGT protein (*M* = 21879 Da) (Figure 2.6A). When the incubation time was extended to 60 min, the unreacted AGT species (*M* = 21877 Da) became the dominant species (Figure 2.6B). In contrast no unmodified AGT species was observed when dG-Pt-AGT complexes were heated in the absence of dG (Figure 2.6C).

Further experimental evidence for platination migration was obtained from the concomitant formation of dG-Pt-dG conjugates. HPLC-ESI^{+}-MS/MS detected a prominent peak co-eluting with the authentic standard of *cis*-1,1-diammine-*bis*-(2'-deoxyguanosine-7-yl)-platinum (II) (dG-Pt-dG, *m/z* 762.3) in samples treated with dG for 60 min at 70 ^{o}C (Figure 2.7) but not in control samples incubated in the absence of dG (results not shown). MS/MS fragmentation of *m/z* 762.3 yielded the product ions corresponding to the loss of ammonia (*m/z* 745.2 [M-NH$_3$]$^{+}$), the removal of 2'-deoxyribose and ammonia (*m/z* 629.1 [M-NH$_3$-dR+H]$^{+}$), and the loss of one dG (*m/z* 495.1 [M-dG]$^{+}$). This compound had the same MS/MS fragmentation pattern and HPLC retention time as a synthetically prepared standard of dG-Pt-dG (Figure 2.7).

Figure 2.6 Platination migration from AGT to dG. (**A**) Deconvoluted ESI^+-MS spectra of dG-Pt-AGT that had been incubated with dG for 10 min at 70 ^oC. (**B**) Spectra of dG-Pt-AGT that had been incubated with dG for 60 min at 70 ^oC. (**C**) Spectra of dG-Pt-AGT complexes that had been incubated in the absence of dG for 10 min at 70 ^oC.

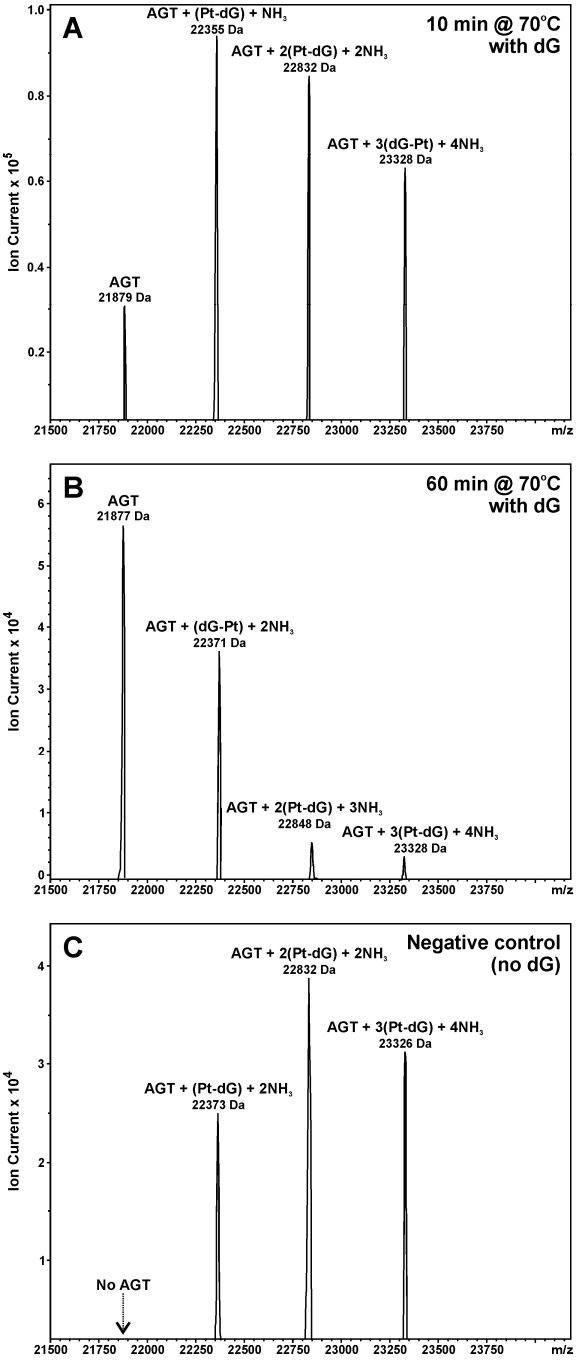
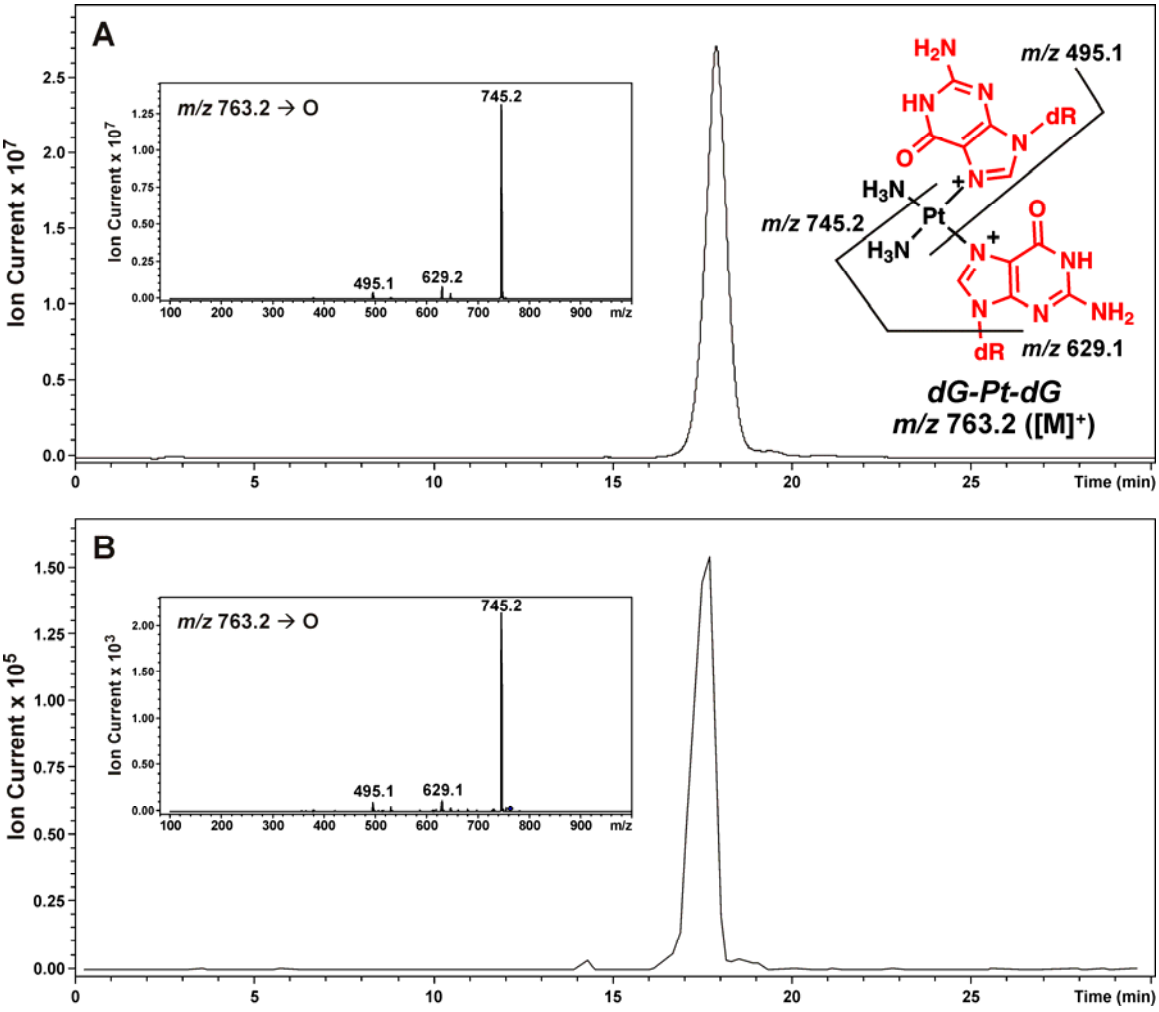


Figure 2.7 HPLC-ESI$^+$-MS/MS analysis of dG-Pt-dG conjugates produced as a result of platination migration from dG-Pt-AGT cross-links to dG. Extracted ion chromatogram of dG-Pt-dG (*m/z* 763.2 $[M]^+$). Inset: MS/MS fragmentation. (**A**) Synthetic dG-Pt-dG; (**B**) Sample treated with dG for 60 min at 70°C.



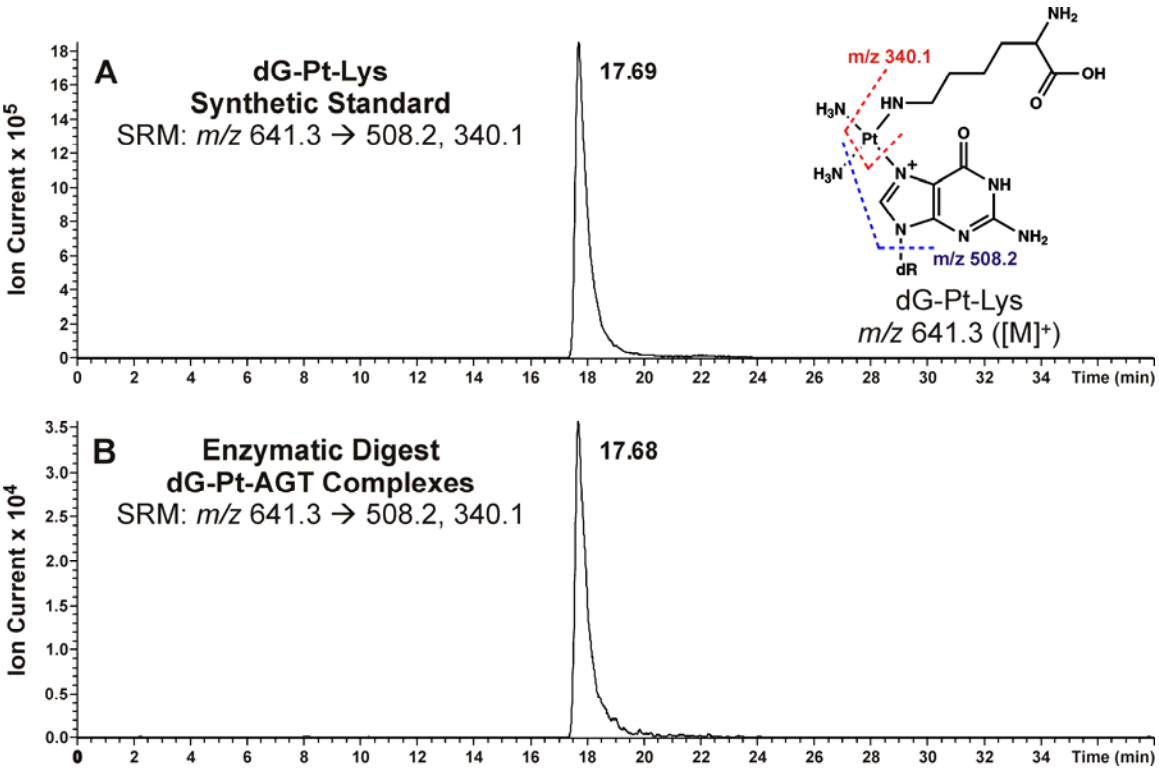
In contrast, no dG-Pt-dG cross-link peak was detected from the control sample in which no dG was added (results not shown). Taken together, our results indicate that cisplatin-induced DNA-protein cross-links can be transformed into DNA-DNA cross-links upon heating, releasing intact proteins. As described in *Chapters III* and *IV* of this Thesis, we took advantage of this reaction to release proteins from DNA following DPC isolation, simplifying their identification by mass spectrometry-based proteomics.

2.3.5 Capillary HPLC-ESI$^+$-MS/MS Analysis of dG-Pt-Lys Conjugates in Total Protein Digests

To confirm that cisplatin-induced AGT-DNA cross-linking involves covalent modification of proteins and DNA, HPLC-ESI$^+$-MS/MS analysis of the total enzymatic digests of platinated protein was performed. Synthetically prepared 1,1-*cis*-diammine-2-(5-amino-5-carboxypentyl)amino-2-(2'-deoxyguanosine-7-yl)-platinum (II) (dG-Pt-Lys) was used as an authentic standard. Following cross-linking reaction, AGT protein was subjected to complete digestion in the presence of trypsin and proteinease K, followed by HPLC-ESI$^+$-MS/MS analysis of the resulting amino acid mixtures in parallel with authentic dG-Pt-Lys.

Representative extracted ion chromatograms for capillary HPLC-ESI$^+$-MS/MS analysis of dG-Pt-Lys conjugates in digests of cisplatin AGT-dG cross-links are shown in Figure 8. HPLC-ESI$^+$-MS/MS of the digest mixtures detected a prominent peak (Figure 2.8B) co-eluting with the authentic standard of dG-Pt-Lys (Figure 2.8A). These data confirm that cisplatin-induced DNA-protein cross-linking can take place between the N7 position of guanine in DNA and the ε-amino group of lysines in AGT protein. Analogous studies involving cisplatin- induced dG conjugates to cysteine, histidine, arginine, and glutamatic acid are currently underway in our laboratory. These experiments are complicated by the requirement for the preparation of synthetic

Figure 2.8 HPLC-ESI$^+$-MS/MS analysis of dG-Pt-Lys conjugates in total proteolytic digests of AGT treated with dG-Pt-Cl to generate cross-links. Following HPLC purification of dG-Pt-AGT complexes, they were subjected to enzymatic digestion to release amino acid-nucleobase conjugates. (**A**) Synthetic dG-Pt-Lys; (**B**) Enzymatic digests of dG-Pt-AGT complexes.



dG-Pt-Cys, dG-Pt-His, dG-Pt-Arg, and dG-Pt-Glu conjugates. We have encountered some difficulties with chemical stability of dG-Pt-Cys conjugates, which decomposed during the Boc deprotection step in the presence of TFA (results not shown). Alternative experimental strategies for preparation of these standards are now being considered.

2.4 Discussion

Studies described above provide a direct evidence for the ability of cisplatin to form DNA-protein cross-links involving AGT protein. Gel-shift assay and mass spectrometry methods revealed that multiple sites of AGT that can participate in cross-linking (Scheme 2.1, Figures 2.1-2.3). Subsequent HPLC-ESI$^+$-MS/MS sequencing of tryptic peptides originating from AGT treated with dG-Pt-Cl monoadduct demonstrated that the cross-linking can occur at six sites within this protein, including Glu^{110}, Lys^{125}, Cys^{145}, His^{146}, Arg^{147}, and Cys^{150} (Figure 2.4). Finally, the exact chemical structure of the amino acid-nucleoside conjugates was established as 1,1-*cis*-diammine-2-(5-amino-5-carboxypentyl)amino-2-(2'-deoxyguanosine-7-yl)- platinum (II) (dG-Pt-Lys) based on the HPLC-ESI$^+$-MS/MS analysis of amino acid-nucleoside conjugates in total protein digests in comparison with the corresponding authentic standard (Figure 2.8).

Our experiments were conducted with dG-Pt-Cl as a model of monoalkylated DNA because platination of N7 position of guanine accounts for over 98% among all the DNA-damage induced by cisplatin *(109)*. However, similar specificity is expected for AGT-DNA cross-linking in cells, resulting in N7-guanine-AGT cross-links. Experiments are now in progress in our laboratory involving DNA oligonucleotides and calf thymus DNA to confirm these initial findings.

Our laboratory previously employed mass spectrometry-based methods to characterize AGT-DNA cross-linking by antitumor nitrogen mustards *(54)* and diepoxybutane (DEB) *(55)*.

For both *bis*-electrophiles, cross-linking took place specifically at the two active site cysteine residues within the protein: Cys^{145} and Cys^{150} *(54;55)*. While cisplatin also targets Cys^{145} and Cys^{150} (Figures 2.4C and 2.4F), additional sites within AGT are involved in cross-linking, including the non-thiol side chains of Glu^{110}, Lys^{125}, His^{146}, and Arg^{147}. This is indicative of a higher reactivity of cisplatin towards proteins and a different cross-linking chemistry characteristic for platinum compounds. Indeed, our earlier experiments with free amino acids (not shown) revealed that basic amino acids such as lysine and arginine were more reactive towards cisplatin than was cysteine (results not shown).

Previous studies investigated the reactions of model proteins such as cytochrome C *(130)*, myoglobin *(131)*, insulin *(132;133)*, ubiquitin *(134)*, and human serum proteins *(135)* with platinum-based drugs. These researchers identified multiple Pt binding sites at *N*-donor residues (His and Lys), *S*-donor residues (Cys and Met), and *O*-donor residues (Thr, Tyr, Asp, and Glu). This is truly consistent with our results for AGT-dG cross-linking (Table 2.2 and Figure 2.4).

Crystal structures of human AGT protein reveal two distinct domains: the *N*-terminal domain (residues 1-85) and the *C*-terminal domain (residues 86-207) *(115)*. The *N*-terminal domain is composed of three β-sheets separated by two α-helixes, whereas the *C*-terminal domain contains two β-sheets and five α-helixes *(115)* (Figure 2.5). The *C*-terminal domain is involved in DNA-binding *via* the helix-turn-helix (HTH) motif and is responsible for DNA repair function. AGT active site contains the conserved active site cysteine motif ($I^{143}PCHRV^{148}$), and the O^6-alkylguanine-binding channel, that are critical for nucleotide flipping and alkyl transfer reaction *(125;136)*.

The observed specificity of cisplatin-mediated cross-linking involving Glu^{110}, Lys^{125}, Cys^{145}, His^{146}, Arg^{147}, and Cys^{150} of the AGT protein may be a result of their proximity to the AGT active site. All six residues are found in the *C*-terminal domain of AGT (Figure 2.5). Furthermore, Cys^{145} is a highly nucleophilic residue that acts as alkyl acceptor in the AGT repair

reaction *(125)*. His^{146}, Arg^{147}, and Cys^{150} are other nucleophilic active site residues that are located in the immediate proximity to DNA in the AGT-DNA complex (Figure 2.5) *(125)*. While the side chain of Glu^{110} is close to the reactive site pocket (I^{143}PCHRV^{148}) in space, Lys^{125} is directly located in the DNA-binding domain (Figure 2.5). In contrast, the other three cysteines within AGT, Cys^{5}, Cys^{24} and Cys^{62} are not identified as the primary cross-linking sites, confirming our hypothesis that cross-linking could potentially take place at residues in close proximity DNA or directly involved in DNA-binding.

Our results reveal a transient nature of dG-AGT cross-links, which can be readily converted to dG-dG cross-links upon heating (Figures 2.6 and 2.7). Historically, there has been a long-term debate over the competition of purine bases, protein side chains, and glutathione for coordination sites of platinum *(137)*. The Hard Soft Acid Base (HSAB) theory predicts that the preferential binding ligands for platinum are *S*-containing biomolecules such as methionine and cysteine residues of proteins or glutathione *(138)*. This should leave little opportunity for platinum to coordinate with *N*-donor ligands in DNA, since numerous *S*-donor biomolecules present in cytosol (*e.g.* glutathione) should capture the drug before it enters the nucleus. Indeed, *in vitro* studies have demonstrated kinetically favored affinity of cisplatin toward S over N *(127)*. However, it is generally accepted that the ultimate biological target of cisplatin is DNA due to the ability of this drug to form DNA intra- and interstrand cross-links between N7 of the purine bases *(111)*.

To determine how cisplatin can induce DNA-DNA cross-links despite its potential reactions with glutathione, peptides, and proteins in the cytosol before reaching the nucleus, Reedijk *(126)* and Sadler *(127)* conducted intra- and intermolecular competition experiments with *S*-ligands and nucleobases. These studies detected platination migration from *S*-ligands to the N7 of guanine. Deubel et al *(137)* employed density functional theory to calculate Pt-L bond energy in $[Pt(NH_3)_3L]^{2+}$ complexes, where L represents peptide side chain, sulfur-containing protein

agents, and guanine bases of DNA. These results revealed that platinum prefers *N*-ligands over *S*-ligands in terms of orbital interactions, electrostatics, and intramolecular hydrogen binding *(137)*.

Taken together, previous studies indicate that there is a balance between Pt-S and Pt-N binding, where Pt-S binding is kinetically favored but Pt-N binding is more thermodynamically stable *(129)*. As a result, platinum initially bound to proteins may further react with DNA to yield Pt-DNA complexes and to release intact proteins. Indeed, our results presented above (Figures 2.6 and 2.7) provide evidence that AGT-DNA cross-links can rearrange to DNA-DNA cross-links, with N7-G winning the competition for the coordination sites of cisplatin upon heating of dG-AGT complexes. However, such platination migration from AGT to DNA requires high temperature and may not be relevant under physiologically *(128)*. Therefore, cisplatin induced DPCs may be stable enough to persist in cells long enough to disrupt several cellular processes, such as DNA replication and/or transcription.

In conclusion, our study demonstrates that cisplatin is capable of sequentially platinating nucleophilic sites within dG and AGT protein to form covalent cross-links between the N7 position of guanine and several reactive sites within AGT, including Glu^{110}, Lys^{125}, Cys^{145}, His^{146}, Arg^{147}, and Cys^{150}. To our knowledge, this is the first report of specific, structurally defined DNA-protein cross-linking involving cisplatin. These results are important because cisplatin-mediated DPC formation is likely to contribute to both on-target and off-target toxicity of this drug and potentially may explain its selective activity against certain types of tumors. Inspired by these results, experiments are currently underway to further elucidate the mechanisms of cisplatin-induced DPCs formation and to develop specific methodology that could monitor the formation and repair of cisplatin-induced DPCs in mammalian cells. These experiments will help us understand the contribution of DPCs to its cytotoxicity and mutagenicity *in vivo*.

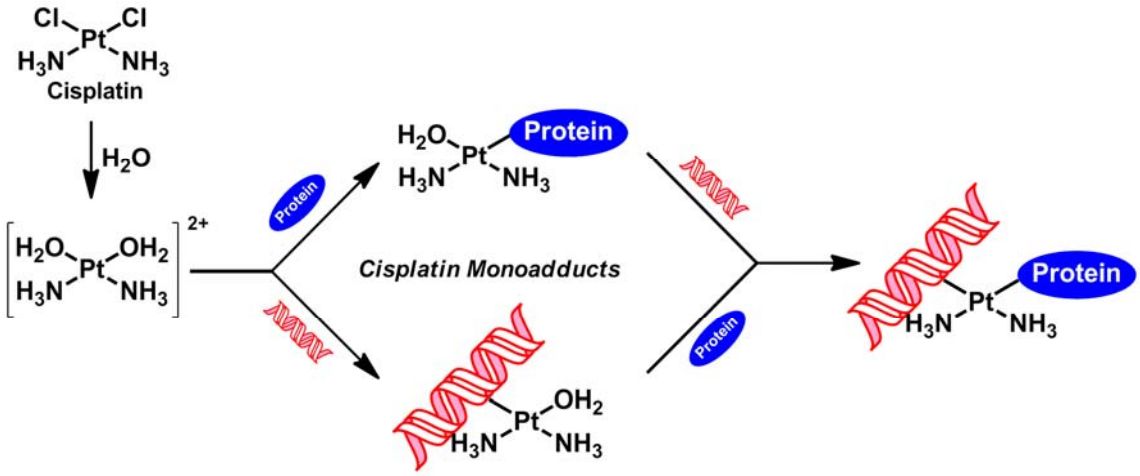
III. DNA-protein Cross-linking by 1,1,2,2-*Cis*-diamminedichloro platinum (II) (Cisplatin) in Nuclear Protein Extracts from Human Cervical Carcinoma (Hela) Cells

3.1 Introduction

Cis-diamminedichloroplatinum (II) (*Cis*-DDP, cisplatin) is an efficient antitumor agent widely used in the treatment of solid tumors of the brain, head and neck, ovary, testicles, and bladder *(68)*. It is generally accepted that the biological activity of cisplatin is due to its ability to form DNA adducts, e.g. 1,2-intrastrand DNA-DNA cross-links between neighboring purine bases (1,2-GG or 1,2-AG intrastrand, 90% of total lesions), 1,3-intrastrand adducts between purine bases separated by another base (1,3-GG or 1,3-AG, 5-10% of total lesions), interstrand DNA-DNA crosslinks (2%), and monofunctional adducts (3% of total lesions) *(69;70)*. However, cisplatin is also capable of forming ternary DNA-platinum-protein cross-links (DPCs) by sequentially platinating nucleophilic sites within DNA and proteins (Scheme 3.1) *(32;71;72)*. Zwelling *et al*. first employed alkaline elution methodology to detect DPC formation in L1210 mouse leukemia cells treated with cisplatin *(32)*. This observation was subsequently confirmed by other groups that utilized similar biophysical methods for detecting cisplatin-induced DPCs *(71)*. So far, cisplatin has been reported to form DPCs involving HMG proteins, histones, and cytokeratins *(73)*. The resulting DNA-protein cross-links have been proposed to play an important role in cytotoxicity of cisplatin *(74)*. However, little is known about the identities of other cellular proteins that participate in cross-linking to DNA in the presence of cisplatin. Furthermore, the covalent structures of cisplatin-mediated DPCs and their cellular abundance have not been determined.

Bulky DPC are likely to cause steric hindrance and to block DNA-protein interactions,

Scheme 3.1 Cisplatin activation and the formation of DNA-protein cross-links.

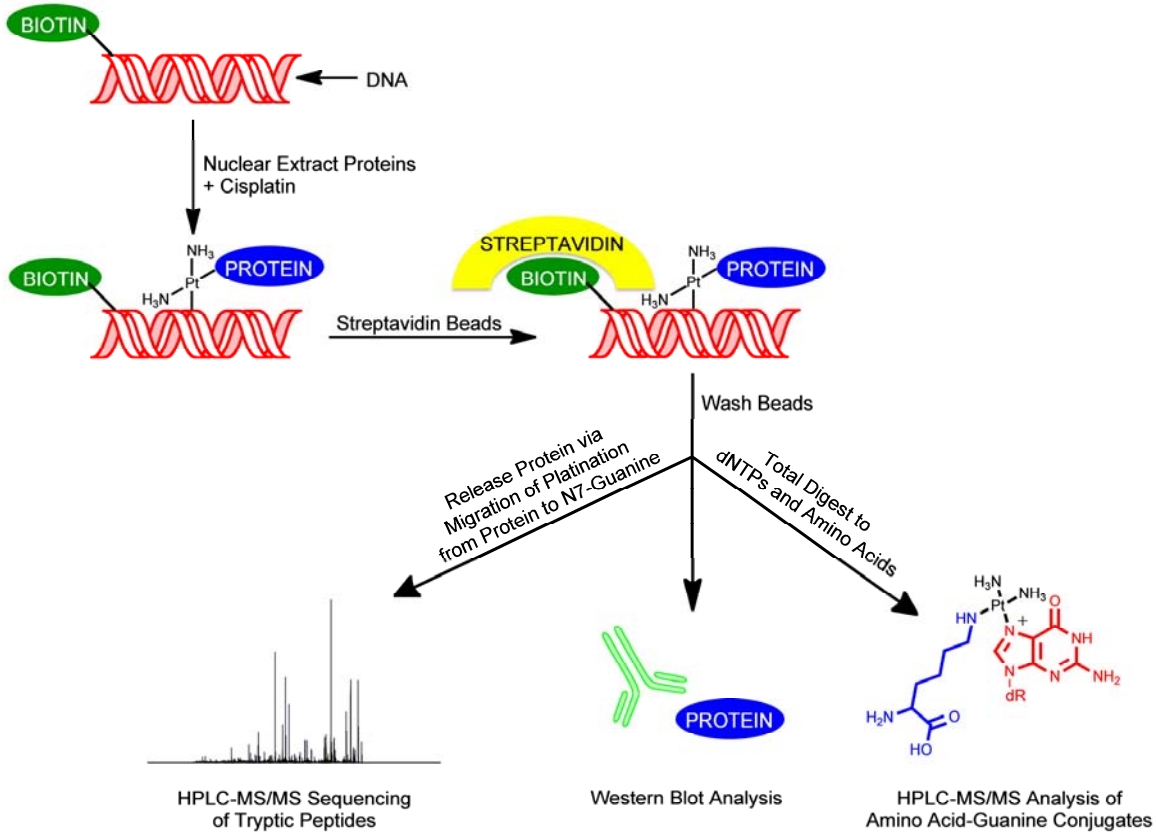


interfering with basic cellular functions such as DNA replication, transcription, repair, recombination, and chromatin remodeling, potentially contributing to the biological effects of platinum compounds *(1;38)*. The purpose of the present study was to conduct a comprehensive investigation of cisplatin-mediated DNA-protein cross-linking using an affinity capture methodolody coupled with mass spectrometry-based proteomics and immunological detection (Scheme 3.2) *(35;36)*. These results are significant because the identification of specific nuclear proteins that become cross-linked to DNA in the presence of cisplatin will help provide a better understanding of their role in cytotoxicity of platinum compounds, potentially facilitating the development of tumor-specific platinum therapies.

3.2 Materials and Methods

Chemicals and Reagents - Cisplatin, phenylmethanesulfonyl fluoride (PMSF), pepstatin, leupeptin, aprotinin, dithiothreitol (DTT), DNase I, alkaline phosphatase and iodoacetamide were purchased from Sigma-Aldrich (St. Louis, MO). Mass spectrometry grade Trypsin Gold was obtained from Promega (Madison, WI). Proteinase K was purchased from Worthington Biochemical Corp. (Lakewood, NJ). Primary polyclonal antibodies to human actin, DNA-(apurinic or apyrimidinic site) lyase (Ref-1), GAPDH, eukaryotic translation elongation factor 1 alpha 1 (EF-1α1), nucleoin (C-23), X-ray repair cross complementing protein 1 (XRCC-1), and poly(ADP-ribose) polymerase (PARP) were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). The primary polyclonal antibody to the ATP-dependent DNA helicase subunit 2 (Ku) was obtained from Lab Vision/NeoMarkers (Fremont, CA). The primary monoclonal antibody to O^6-alkylguanine DNA alkyltransferase (AGT) was purchased from Millipore (Temecula, CA). Alkaline phosphatase-conjugated anti-mouse and anti-rabbit IgG secondary antibodies were obtained from Sigma-Aldrich (St. Louis, MO). Synthetic DNA oligodeoxynucleotides were

Scheme 3.2 Experimental scheme for biotin capture enrichment of DNA-protein cross-links from nuclear protein extracts incubated with cisplatin in the presence of double-stranded DNA.



prepared by standard solid phase synthesis at the University of Minnesota Biomedical Genomics Center (Minneapolis, MN). Recombinant GAPDH protein was obtained from Sigma-Aldrich (St. Louis, MO). Recombinant AGT protein was a generous gift from Professor Anthony Pegg (Pennsylvania State University), and *cis*-1,1-diammine-2-(5-amino-5-carboxypentyl)-amino-2-(2'-deoxyguanosine-7-yl)-platinum(II) (dG-Pt-Lys) was prepared as described previously (*see Chapter II*).

Preparation of Nuclear Protein Extracts from Hela Cells - Human cervical carcinoma (Hela) cells were maintained as exponentially growing monolayer cultures in Dulbecco's Modified Eagle's Medium supplemented with 9% fetal bovine serum, in a humidifier incubator at 37°C with 5% CO_2. For preparation of nuclear extracts, ~10^8 cells were harvested, washed with ice cold phosphate-buffered saline, and suspended in a hypotonic buffer (10 mM Tris-HCl – pH 7.4/10 mM $MgCl_2$/10 mM KCl/1 mM DTT) containing 1 mM PMSF *(139)*. Following 5 min incubation on ice, cells were broken by 20 strokes in a Dounce homogenizer and centrifuged at 2000*g* for 10 min. The sedimented nuclei were re-suspended in hypotonic buffer containing 350 mM NaCl and a protease inhibitor cocktail (1 μg/mL pepstatin; 0.5 μg/mL leupeptin; 0.75 μg/mL aprotinin; 1 mM PMSF) and kept on ice for 1 h. The resulting nuclear lysate was centrifuged at 160,000*g* at 4 °C for 30 min, and the nuclear proteins were isolated in the clear supernatant. The extracts were dialyzed for 2 h at 4°C against 10 mM Tris-HCl – pH 7.4/10 mM KCl/10 mM $MgCl_2$ using Slide-A-Lyzer dialysis cassettes with a 3.5 kDa molecular weight cut-off (Pierce Biotechnology, Rockford, IL). Protein concentrations were determined using the Bio-Rad Protein Assay (Bio-Rad, Hercules, CA).

DNA-Protein Cross-linking and Biotin Capture - 5'-Biotinylated double-stranded oligodeoxynucleotides (5'-GGA GCT GGT GGC GTA GGC-3' (+) strand, 3.12 nmol) were

incubated with nuclear protein extracts from HeLa cells (500 μg total protein) in the absence or in the presence of cisplatin (1-50 μM, 1 mL total volume) at 37°C for 3 h. Biotinylated DNA, along any with DNA-protein cross-links, was captured on streptavidin Sepharose High Performance beads (600 μL slurry, GE Healthcare, Piscataway, NJ), overnight at 4°C, with rotation. To remove non-covalently bound proteins, the beads were washed twice with 1% SDS (1 h wash followed by another 30 min, with rotation, at room temperature), twice with 4 M urea containing 1 M NaCl (30 min followed by another 15 min, with rotation, at room temperature), and twice with phosphate-buffered saline (no incubation). Following each washing step, the beads were centrifuged at 2,000*g* for 1 min, and the supernatant was discarded. Proteins covalently cross-linked to DNA were released from the beads by adding 110 μL NuPage 4X LDS Sample Buffer (Invitrogen, Carlsbad, CA) and heating to 90°C for 15 min. As described in Chapter 2, these conditions result in trans-platination reaction that releases intact proteins and forms G-G cisplatin-DNA cross-links (*see Chapter II*). The proteins were subsequently analyzed by HPLC-ESI^+-MS/MS and Western blotting as described below.

SDS-PAGE of Protein-Oligonucleotides Conjugates - The proteins were separated by SDS-PAGE using 12% Tris-HCl Ready Gels (Bio-Rad, Hercules, CA). The gels were stained with Bio-Rad Silver Stain Kit (Bio-rad, Hercules, CA) or with the SimplyBlue SafeStain (Invitrogen, Carlsbad, CA) according to the manufacturers' instructions.

Protein Identification by Mass Spectrometry - For proteomic analyses, the cross-linked proteins were separated using 12% Tris-HCl Ready Gels (Bio-Rad, Hercules, CA) and stained with SimplyBlue SafeStain (Invitrogen, Carlsbad, CA). Gel lanes were cut into slices, washed with 100 mM ammonium bicarbonate, and subjected to treatment with dithiothreitol and

iodoacetamide as previously described *(140)*. Gel pieces were dehydrated with acetonitrile, dried under vacuum, and reconstituted in 25 mM ammonium bicarbonate buffer containing mass spectrometry grade trypsin (10 μg). The samples were digested overnight at 37°C. Tryptic peptides were extracted with 1% aqueous formic acid/60% acetonitrile, evaporated to dryness, and re-suspended in 0.1% aqueous formic acid for mass spectrometric analysis.

Tryptic peptides were analyzed by HPLC-ESI$^+$-MS/MS with a ThermoScientific LTQ OrbitrapXL mass spectrometer in line with an Eksigent MicroAS autosampler and nanoLC 1D HPLC pump, a nanospray source, and an Xcalibur 1.4 software for instrument control. Peptides were resolved with a 100 μm x 13 cm fused silica capillary column (Poly micro Technologies, LLC, Phoenix, CA) manually packed with 5 μm, 300 Å Jupiter C18 packing (Phenomenex, Torrence, CA). The column was eluted at a flow rate of 0.3 μL/min with a gradient of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The solvent composition was initially set at 5% B, followed by a linear increase to 60% B over 60 min, and further to 95% B in 5 min. Liquid chromatography was carried out at an ambient temperature. Centroided MS-MS scans were acquired using an isolation width of 2 *m/z*, an activation time of 30 ms, an activation Q of 0.250, 30% normal collision energy, and 1 microscan with a max ion time of 100 ms for each MS/MS scan. The mass spectrometer was tuned prior to each analysis using synthetic peptide TpepK (AVAGKAGAR), so that some parameters may have varied slightly from experiment to experiment. Typical tune parameters were as follows: spray voltage of 1.75 kV, a capillary temperature of 160°C, a capillary voltage of 50 V, and tube lens voltage of 120 V. MS/MS spectra of the peptides were collected using data-dependent scanning in which one full scan mass spectrum was followed by eight MS/MS spectra. Dynamic exclusion was enabled for 60 s and singly charged species were excluded from MS/MS.

Spectral data were analyzed using an in-house developed software pipeline "TINT" that linked raw data extraction, database searching, and probability scoring. Raw data were extracted

and converted to the mzXML format using ReadW. Spectra that contained fewer than 6 peaks or had less than 20 measured total ion current (TIC) were excluded. Data were searched using the SEQUEST v.27 algorithm *(123;124)* on a high speed, multiprocessor Linux cluster in the Minnesota Super Computing Institute at the University of Minnesota using the human subset consisting of the NCBI derived human protein database v200806 combined with its reversed counterpart along with common protein contaminants totaling 70,711 entries. Search parameters included trypsin specificity and up to 2 missed cleavage sites. Cysteine carboxamidomethylation (+57.0215 Da) was set as a fixed modification, and methionine oxidation (+15.9949 Da) was set as a variable modification. Precursor mass tolerance was set to 1.25 *m/z* within the calculated average mass, and fragment ion mass tolerance was set to 0.5 *m/z* of their monoisotopic mass. Identified peptides were filtered using scaffold 2 software (Proteome Software, INC., Portland, OR) *(141)*, to a target false discovery rate (FDR) of 5%. The FDR was calculated with the following expression FDR = (2R)/(R+F)*100, where R is the number of passing reversed peptide identifications and F is the number of passing forward (normal orientation) peptide identifications. The second round of filtering removed proteins supported by less than four distinct peptide identifications in the analyses. Indistinguishable proteins were recognized and grouped. Parsimony rules were applied to generate a minimal list of proteins that explained all of the peptides that passed our entry criteria *(142)*. Furthermore, manual statistical analyses were performed to ensure that the levels of proteins captured from treated samples were significantly higher than those in untreated controls. Specifically, a protein was included in the list only if the spectra of its peptides were exclusively found in the treated samples or the amount of total peptide spectra counts from treated samples was at least four times higher than those from untreated controls. Finally, to be considered a positive identification, all proteins were required to have a minimum of four unique peptide spectra and 10 total spectral counts.

Western Blot Analysis of Identified Proteins - Nuclear protein extracts from HeLa cells were incubated with 5'-biotinylated oligonucleotide duplexes in the presence of 0-50 μM cisplatin as described above, and the resulting DPCs were captured on streptavidin beads. Cross-linked proteins were released from DNA by heating, separated by 12% gel SDS-PAGE, and transferred to Trans-Blot nitrocellulose membranes (Bio-Rad, Hercules, CA). Membranes were immediately blocked in Tris-buffered saline (TBS) containing 5% (w/v) bovine serum albumin. Following 1–2 h incubation with the primary antibody at room temperature, the blots were washed three times with TBS buffer and incubated overnight at 4 °C with the corresponding alkaline phosphatase-conjugated secondary antibody. Following three additional washes with TBS buffer, the blots were developed using SIGMA Fast BCIP/NBT (Sigma, St. Louis, MO). The developed blots were scanned as image files, and the optical densities of the bands were then quantified using ImageJ software available free of charge from the NIH website (www.ncbi.nlm.nih.gov). The extent of DNA-protein cross-linking was estimated by comparing the band intensities in cross-linked samples to that of a known amount of nuclear proteins extract loaded on a separate lane of the gel.

HPLC-ESI$^+$-MS/MS Analysis of dG-Pt-Lys - Nuclear protein extracts from Hela cells (1.3 mg total protein) were incubated with 8 nmol of double-stranded 5'-bionylated oligodeoxynucleotides (5'-GGA GCT GGT GGC GTA GGC-3' and the complementary strand) in the presence of 50 μM cisplatin, and the resulting DPCs were captured on streptavidin beads. Following incubation with 70% acetonitrile/5% aqueous formic acid to release DPCs (overnight incubation at 4°C, 30 min incubation at room temperature, and 30 min incubation at 90°C), the beads were centrifuged for 1 min at 2,000*g*, and the supernatants containing DPCs were decanted and pooled. DNA was digested with DNase I (20 U) and alkaline phosphatase (60 U) overnight

at 37ºC *(143)* to produce protein-nucleoside conjugates. Samples were dried under vacuum, reconstituted in 25 mM ammonium bicarbonate, and digested to peptides with trypsin (20 μg, 0.86 nmol, 37°C overnight). To achieve complete hydrolysis to amino acids, tryptic peptides were dried under vacuum, reconstituted in water, and digested with proteinase K (20 μg, 0.70 nmol) for 48 h at room temperature. The digest mixtures were subjected to off-line HPLC separation using the same HPLC method employed to isolate the synthetic dG-Pt-Lys standard as described above. HPLC fractions containing dG-Pt-Lys (8-9 min) were dried under vacuum and re-suspended in 15 mM ammonium acetate, pH 5.0 (25 μL) for HPLC-ESI^+-MS/MS analysis (injection volume, 8 μL).

HPLC-ESI^+-MS/MS of dG-Pt-Lys conjugates was conducted with an Agilent 1100 capillary HPLC system interfaced to a Thermo-Finnigan TSQ Vantage mass spectrometer. Chromatographic separation was accomplished using a Hypercarb column (100 mm x 0.5 mm, 5 μm, ThermoScientific, Waltharm, MA) eluted with a gradient of 15 mM ammonium acetate (A) and 1:1 acetonitrile:water with 5% formic acid (B) at a flow rate of 13 μL/min. The gradient program began at 2% B, followed by a linear increase to 8% B in 10 min, further to 11% B in 17 min, and finally to 30% B in 2 min. Using this gradient, dG-Pt-Lys eluted at ~10 min. ESI was achieved at a spray voltage of 3.2 kV and a capillary temperature of 250°C. CID was performed with Ar as a collision gas (1.0 mTorr) at a collision energy of 25V. The MS parameters were optimized for maximum response during infusion of a standard solution of dG-Pt-Lys and may vary slightly between experiments. HPLC-ESI^+-MS/MS analyses were performed in the selected reaction monitoring (SRM) mode using transitions corresponding to major fragment ions observed upon CID fragmentation of dG-Pt-Lys in a triple quadrupole mass spectrometer (*m/z* 641.2 $[M]^+ \rightarrow$ 624.2 $[M–NH_3]^+$, 508.1 $[M–NH_3–\text{deoxyribose+H}]^+$, and 357.1 $[M–NH_3–\text{deoxyguanosine}]^+$).

3.3 Results

3.3.1 Strategy for Affinity Purification of Proteins Cross-linked to DNA by Cisplatin

Cisplatin-induced DPCs were enriched by the affinity capture-based methodology recently developed in our laboratory *(35;36)*. In brief, nuclear extract proteins from human cervical carcinoma (HeLa) cells were incubated with synthetic DNA duplexes containing a 5'-biotin tag in the presence of increasing concentrations of cisplatin (Scheme 3.2). Following biotin capture of the proteins covalently cross-linked to the biotinylated DNA, streptavidin beads were subjected to stringent washing with a series of buffers (e.g. 1% SDS, 4 M urea, 1 M NaCl, and PBS buffer) to remove non-covalently bound proteins. Finally, the cross-linked proteins were eluted by heating the beads in SDS-containing gel loading buffer via platination migration (*see Chapter II*).

Our previous methodology was developed for depurinating DPC lesions induced by nitrogen mustards *(36)* and 1,2,3,4-diepoxybutane (DEB) *(35)* that are readily released from the DNA backbone in the form of protein-guanine conjugates, simplifying their selective removal from streptavidin beads for MS analysis. Platination of the N7 position of dG, unlike N7 alkylation, does not weaken the glycosidic bond, so the corresponding protein-guanine conjugates cannot be released from DNA by heating *(144;145)*. However, our previous studies involving recombinant proteins revealed that the protein components of cisplatin-induced DPCs can be released upon heating at 90 ^{o}C *via* trans-platination reaction that produces guanine-guanine cross-links (*see Chapter II*). The resulting proteins were separated by SDS-PAGE and subjected to in-gel tryptic digestion, followed by mass-spectrometry based proteomics.

3.3.2 Mass Spectrometric Identification of Proteins Participating in Cross-linking

Silver staining of SDS-PAGE separated proteins revealed a concentration-dependent increase in protein signal in samples treated with increasing concentrations of cisplatin, indicating the formation of DPCs (Figure 3.1A). The cross-linked proteins span a wide range of molecular weights (25 – 250 KDa). A similar concentration-dependence was obtained when using SimplyBlue staining to visualize SDS-PAGE separated proteins (Figures 3.2A and 3.2B). A small amount of protein was also observed in control samples prepared in the absence of cisplatin (lane 3 in Figure 3.1A and Figure 3.2A), probably due to the endogenous DNA-protein cross-linking and/or the presence of non-covalently bound proteins with strong affinity for DNA, which are not removed during the washing step. However, protein background in untreated samples was minimal and did not interfere with proteomic analyses. Densitometric analyses (Figure 3.1B) revealed that as much as 7.9% (2.5% in Figure 3.2B) of total protein is cross-linked to DNA following treatment with 50 μM concentration of cisplatin. By comparison, 10-fold higher concentrations were required to obtain similar levels of cross-linking with an antitumor nitrogen mustard, mechlorethamine *(36)*, and 10,000-fold concentrations were required for DPC formation by 1,2,3,4-diepoxybutane *(35)*, indicating that cisplatin is a very effective DPC-inducing agent.

In order to identify the proteins which participate in cisplatin-induced DPC formation, nuclear protein extracts from HeLa cells (in triplicate) were incubated with biotinylated DNA in the presence of cisplatin (10 μM). Control samples (N = 3) were incubated with buffer only in the absence of drug. Following biotin capture enrichment and heating to induce trans-platination (Scheme 3.2), the released proteins were separated by SDS-PAGE (Figure 3.3). The gel lanes containing proteins were cut into five portions encompassing the entire molecular weight range and excised. Proteins were subjected to in-gel tryptic digestion *(140)*, and the resulting tryptic peptides were extracted and analyzed by HPLC-ESI^+-MS/MS analysis to identify the proteins participating in cross-linking. Protein identifications were based on the MS/MS spectra of tryptic

Figure 3.1 Concentration-dependent formation of DPCs in nuclear protein extracts prepared from HeLa human cervical carcinoma cells following exposure to cisplatin. (A) Nuclear protein extracts from HeLa cells (500 μg) and 5'-biotinylated double-stranded oligodeoxynucleotides (3.12 nmol) were incubated in the presence of 0–50 μM Cisplatin. The resulting DPCs were captured on streptavidin beads, and the proteins were resolved on 12% SDS-PAGE. Gels were stained with SilverQuest SilverStain to visualize the cross-linked proteins. (B) Densitometric analysis of protein bands in the 25 – 250 kDa molecular weight range was used to estimate the extent of total protein cross-linking to DNA in the presence of cisplatin. Known amounts of nuclear protein extract were analyzed as a control to estimate the cross-linking efficiency.

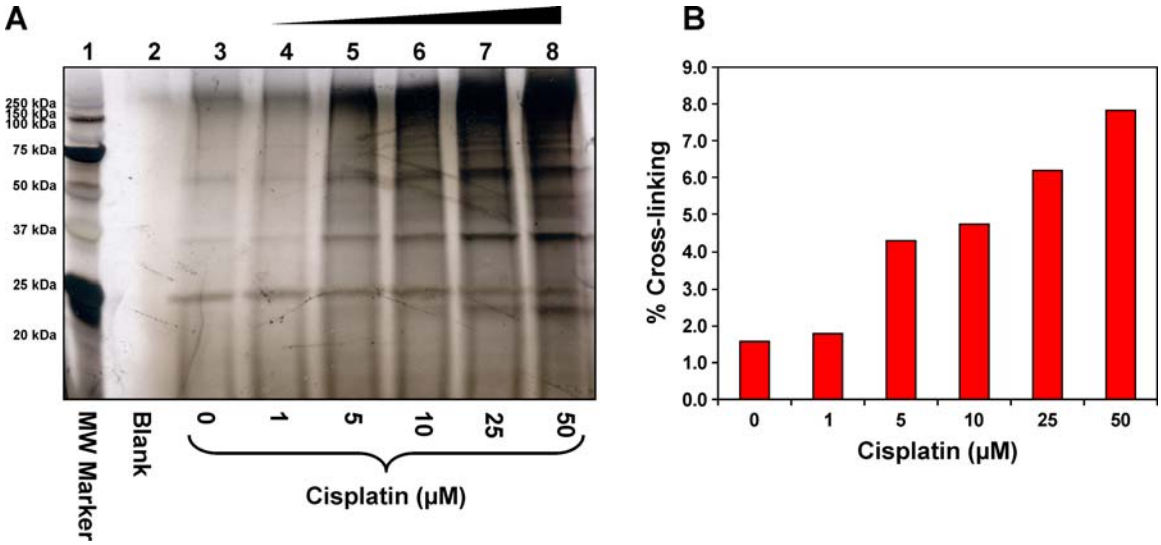


Figure 3.2 Concentration-dependent formation of DPCs in nuclear protein extracts prepared from human cervical carcinoma (HeLa) cells following exposure to cisplatin. (A) Nuclear protein extracts from HeLa cells (500 μg) and 5'-biotinylated double-stranded oligodeoxynucleotides (3.12 nmol) were incubated in the presence of 0–50 μM cisplatin. The resulting DPCs were subjected to biotin capture enrichment released from DNA by heating, and resolved by 12% SDS-PAGE. Gels were stained with SimplyBlue SafeStain to visualize the cross-linked proteins. (B) Densitometric analysis of protein bands in the 25 – 250 kDa molecular weight range to estimate the extent of total protein cross-linking to DNA in the presence of cisplatin. Band intensity was compared to staining of a known amount of nuclear protein extract analyzed as a control to estimate the cross-linking efficiency.

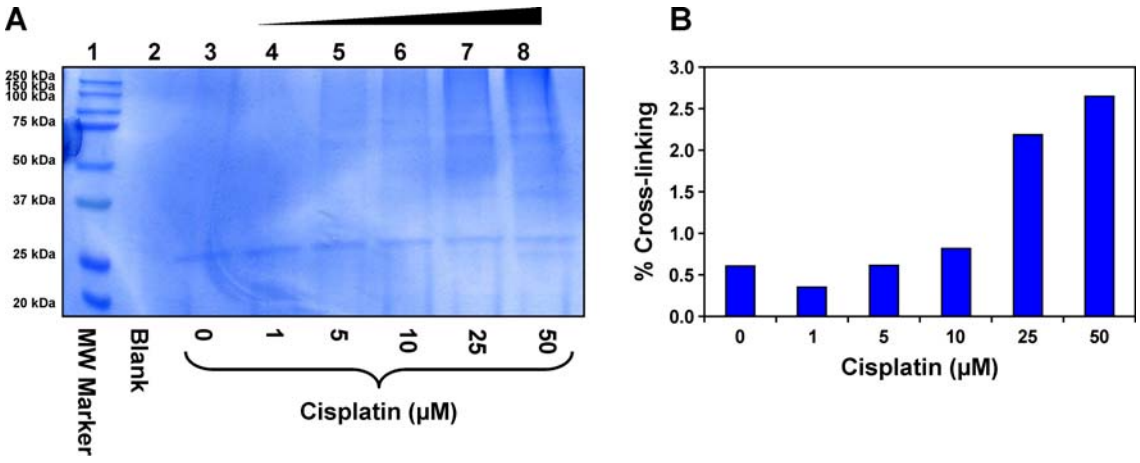
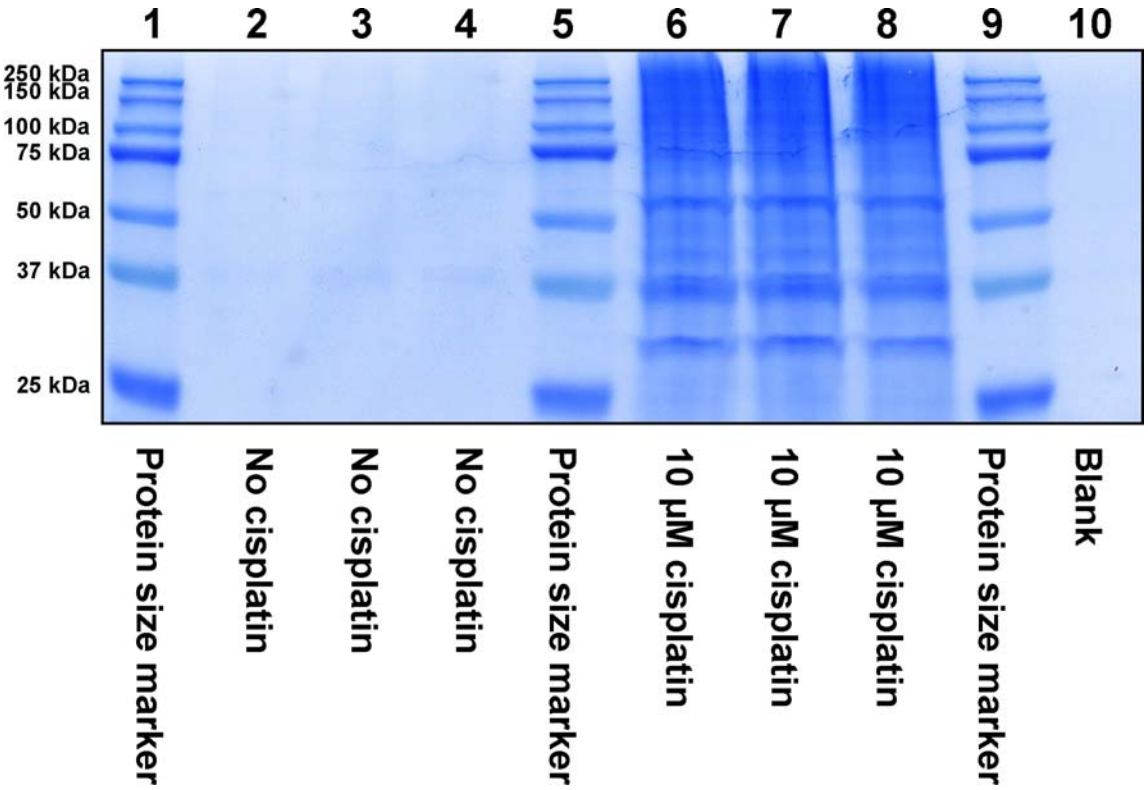


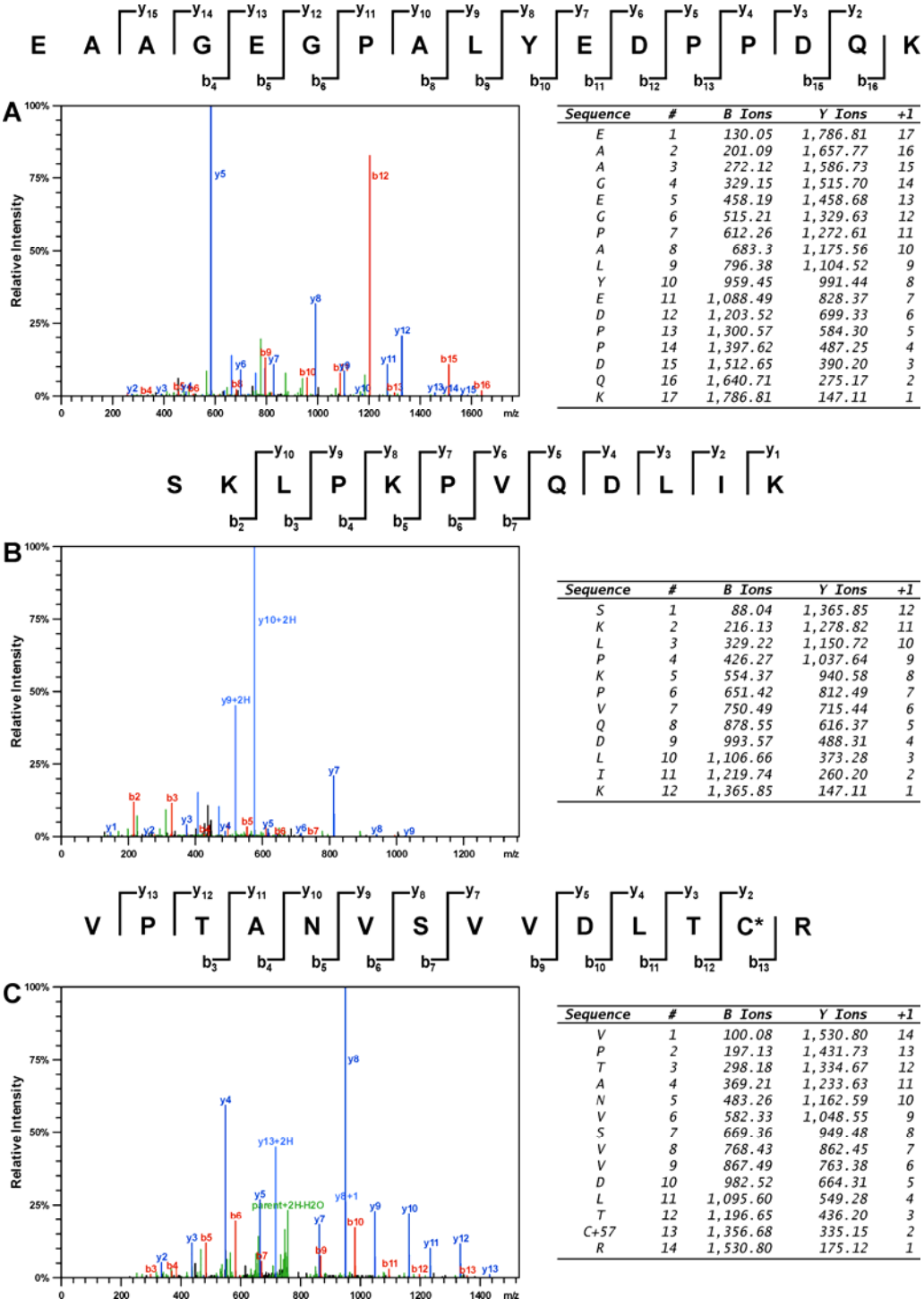
Figure 3.3 SDS-PAGE analysis of cisplatin-induced DPCs in nuclear extracts prepared from Hela cells. Nuclear proteins extracts (500 μg) and 5'-biotinylated double-stranded oligodeoxynucleotides (3.12 nmol) were incubated in triplicate with 0 (lanes 2-4) or 10 μM cisplatin (lanes 6-8). Following biotin capture enrichment and release *via* trans-platination, cross-linked proteins were resolved by 12% SDS-PAGE and visualized with SimplyBlue SafeStain. Each lane was cut into 5 sections and proteins present in each piece were subjected to in-gel tryptic digestion prior to HPLC-ESI$^+$-MS/MS analysis.



peptides, yielding characteristic *b*- and *y*-series fragment ions that were used to determine amino acid sequence (see examples in Figure 3.4). Spectral data were subjected to parsimony analysis, resulting in identification of a total of 131 HeLa nuclear proteins participating in cross-linking to DNA in the presence of cisplatin (Table 3.1).

Proteins identified by mass spectrometry-based proteomics were classified according to their cellular function, cellular distribution, participation in biological processes, and molecular functions using the GO database available via the European Bioinformatics Insititute (http://www.ebi.ac.uk/QuickGo) (Figure 3.5). Of the proteins listed in Table 3.1, 52 (39.2%) are known nuclear proteins (Figure 3.5A), with cellular roles in transcriptional regulation (*e.g.* poly-(ADP-ribose) polymerase (PARP)), heterogeneous nuclear ribonucleoproteins (*e.g.* GAPDH, Ref-1, actin) *(146-150)*, chromatin remodeling (*e.g.* actin) *(150)*, DNA replication (*e.g.* DNA polymerase δ) *(151)*, and DNA repair (*e.g.* Ref-1, AGT, PARP, flap endonuclease-1 (Fen-1)) *(114;149;152-154)*. An additional 44 proteins (33.08%) are classified as cytoplasmic, and 10 (7.52%) are classified as membrane-bound proteins (Figure 3.5A). The latter are likely to have additional cellular roles requiring nuclear localization (which are not annotated in the GO database), explaining their participation in cross-linking to DNA. We also observed proteins involved in cell motility/signaling/ architecture (*e.g.* actin), protein folding (*e.g.* T complex proteins) *(155)*, and biosynthesis (*e.g.* histidyl tRNA synthetase) *(156)*. Twenty-five proteins identified in biotin capture fractions of cisplatin-induced DPCs, including actin, GAPDH, EF-1α1, PARP, Ku, and Fen-1, have been previously shown to form cross-links to DNA in the presence of other *bis*-electrophiles, *e.g.* mechlorethamine and DEB *(35;36)*. Additionally, several of the identified proteins are closely related to proteins that form DPCs in the presence of mechlorethamine and DEB (*e.g.* different chains of tubulin). Altogether, approximately 19% of the 131 proteins identified in the present work also formed DPCs in the presence of mechlorethamine and DEB. The remaining proteins were not common among those three lists,

Figure 3.4 Examples of HPLC-ESI^{+}-MS/MS data used for the identification Ref-1 (A), PARP (B), and GAPDH protein (C) present in affinity-captured DPCs. *Cysteine carboxamidomethylation (+57).



Sequence	#	B Ions	Y Ions	+1
E	1	130.05	1,786.81	17
A	2	201.09	1,657.77	16
A	3	272.12	1,586.73	15
G	4	329.15	1,515.70	14
E	5	458.19	1,458.68	13
G	6	515.21	1,329.63	12
P	7	612.26	1,272.61	11
A	8	683.3	1,175.56	10
L	9	796.38	1,104.52	9
Y	10	959.45	991.44	8
E	11	1,088.49	828.37	7
D	12	1,203.52	699.33	6
P	13	1,300.57	584.30	5
P	14	1,397.62	487.25	4
D	15	1,512.65	390.20	3
Q	16	1,640.71	275.17	2
K	17	1,786.81	147.11	1

Sequence	#	B Ions	Y Ions	+1
S	1	88.04	1,365.85	12
K	2	216.13	1,278.82	11
L	3	329.22	1,150.72	10
P	4	426.27	1,037.64	9
K	5	554.37	940.58	8
P	6	651.42	812.49	7
V	7	750.49	715.44	6
Q	8	878.55	616.37	5
D	9	993.57	488.31	4
L	10	1,106.66	373.28	3
I	11	1,219.74	260.20	2
K	12	1,365.85	147.11	1

Sequence	#	B Ions	Y Ions	+1
V	1	100.08	1,530.80	14
P	2	197.13	1,431.73	13
T	3	298.18	1,334.67	12
A	4	369.21	1,233.63	11
N	5	483.26	1,162.59	10
V	6	582.33	1,048.55	9
S	7	669.36	949.48	8
V	8	768.43	862.45	7
V	9	867.49	763.38	6
D	10	982.52	664.31	5
L	11	1,095.60	549.28	4
T	12	1,196.65	436.20	3
C+57	13	1,356.68	335.15	2
R	14	1,530.80	175.12	1

Table 3.1 List of proteins that form covalent cross-links to chromosomal DNA in the presence of 10 μM 1,1,2,2-*cis*-diamminedichloroplatinum(II) (cisplatin) (10 μM for 3 h) in cell-free nuclear protein extracts from Hela cells.

Swiss-Port ID	*Protein*	*% Coverage*	*Peptide Sequences*	*Total Spectra*	*Primary Celluar Function*	*Protein MW (Da)*	*No. of Cys.*	*No. of Lys.*
P07355	annexin A2	27	10	17	*Cell Signalling/Motility/Architecture*	38,473	4	32
P60709	actin, cytoplasmic 1	33	10	79		41,737	6	19
Q05682	caldesmon	6.6	5	10		93,250	1	104
P68032	actin, α cardiac muscle 1	31	18	144		41,785	5	19
P08107	heat shock 70 kDa protein 1A/1B	42	30	111		69,921	5	50
P07900	heat shock protein HSP 90-α	7.7	5	7		84,529	7	80
P11142	heat shock cognate 71 kDa protein	41	23	65		70,767	4	44
Q00341	vigilin	21	25	40		141,324	11	109
P46940	ras GTPase-activating-like protein IQGAP1	3.7	5	5		189,120	11	134
P26038	moesin	6.9	4	4		67,689	2	62
Q13492	phosphatidylinositol-binding clathrin assembly protein	8.4	4	8		70,546	3	40
Q13283	ras-GTPase-activating protein-binding protein 1	27	8	20		52,033	1	19
P18754	regulator of chromosome condensation	27	9	19		44,969	8	24
P68363	tubulin α-1B chian	28	11	30		50,152	12	19
Q13509	tubulin β-3 chain	11	4	13		50,433	8	16
P00966	argininosuccinate synthetase	16	5	5	*Celluar Homeostasis/Cell Cycle*	46,530	5	33
O43684	mitotic checkpoint protein BUB3	25	6	23		37,155	8	17
P31327	carbamoyl-phosphate synthetase 1	4.2	4	6		160,549	20	100
P06493	cell division protein kinase 1	25	7	10		34,095	1	24
P78371	T-complex protein 1 subunit β	15	6	11		57,357	6	37
P49368	T-complex protein 1 subunit γ	15	7	10		60,534	10	36
P50991	T-complex protein 1 subunit δ	17	7	14		57,793	9	39
Q99832	T-complex protein 1 subunit η	6.1	4	5		59,367	9	41

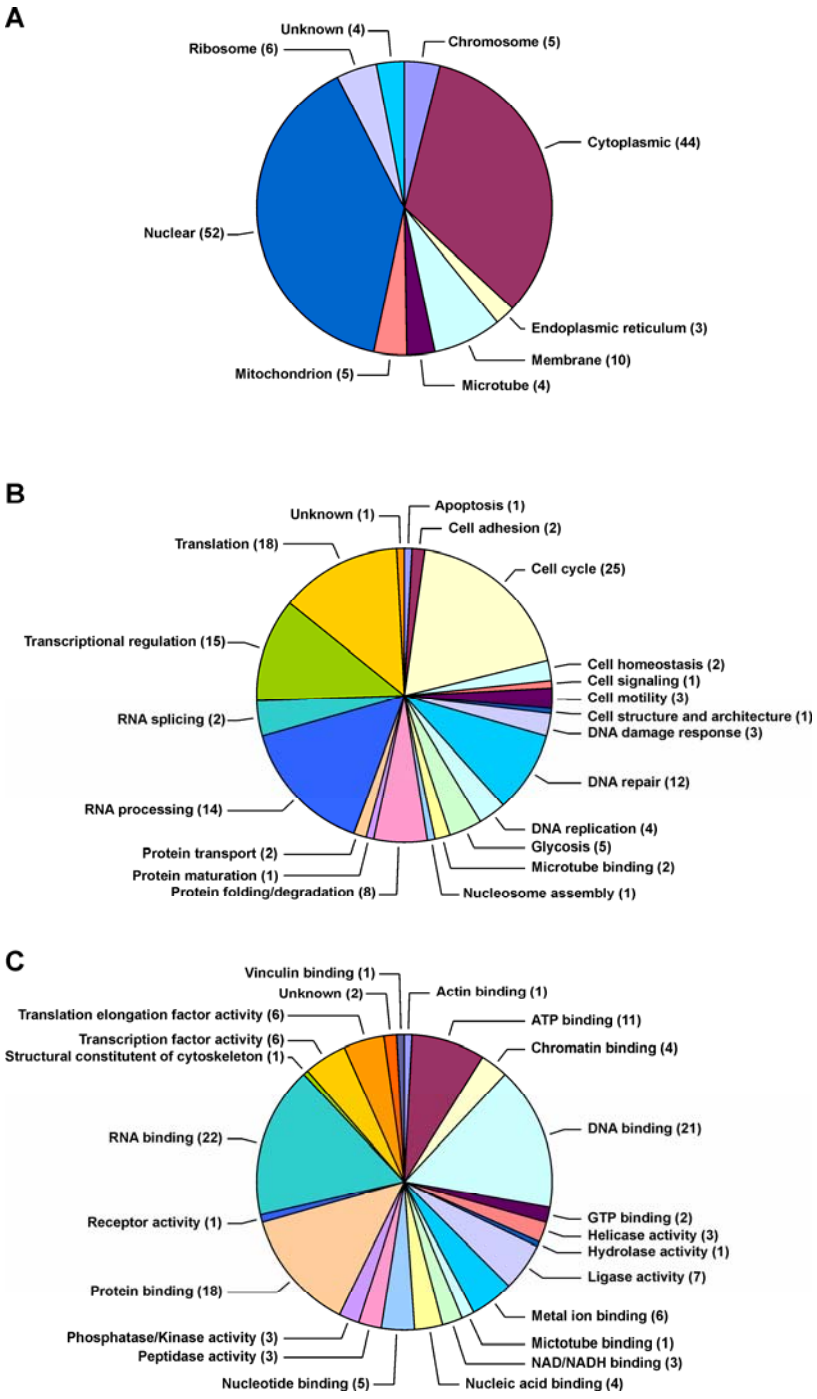
P21291	cysteine and glycine-rich protein 1	28	4	9	20,436	15	23
Q96EP5	DAZ-associated protein 1	16	4	9	43,383	4	18
Q9H4M9	EH-domain containing protein 1	9.6	6	6	60,627	1	44
P06733	α-enolase	19	6	13	47,038	6	38
P13929	β-enolase	15	4	5	46,801	6	37
P11413	glucose-6-phosphate 1-dehydrogenase	9.7	5	7	59,126	8	29
P04406	glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	38	8	53	35,922	3	26
P12081	histidyl-tRNA synthetase	10	4	8	57,279	10	43
P51610	host cell factor 1	3.3	5	12	208,732	35	73
Q9Y3I0	UPF0027 protein C22orf28	17	8	12	55,210	9	32
Q9H0C8	integrin-linked kinase-associated protein phosphatase 2C	12	4	6	42,907	10	29
O75874	isocitrate dehydrogenase 1 [NADP]	19	6	6	46,659	5	37
P33176	Kinesin-1 heavy chain	7.2	6	10	109,554	13	98
P00338	L-lactate dehydrogenase A chain	22	7	8	36,558	5	28
P31153	S-adenosylmethionine synthase isoform type-2	12	4	4	43,661	6	27
P11586	C-1-tetrahydrofolate synthase	5.6	5	9	101,428	12	62
O43175	D-3-phosphoglycerate dehydrogenase	17	7	24	56,519	13	26
P00558	phosphoglycerate kinase 1	23	7	13	44,483	7	42
P62195	26S protease regulatory subunit 8	22	7	7	45,495	3	33
P30101	protein disulfide isomerase A3	20	9	22	54,265	7	48
P14618	pyruvate kinase isozymes M1/M2	59	40	142	57,806	10	37
P62826	GTP-binding nuclear protein Ran	23	5	29	24,292	3	18
Q9P258	protein RCC2	52	28	86	56,085	13	33
P50454	serpin H1	11	4	4	44,523	1	34
Q01105	protein SET	27	6	11	33,489	0	23
Q9NTJ3	structural maintenance of chromosomes protein 4	5.1	6	6	147,182	12	148
Q9Y657	Spindling-1	20	4	4	29,601	2	21
O95347	structural maintenance of chromosomes protein 2	4.3	4	7	135,656	9	146
P27816	microtubule-associated protein 4	9.4	7	24	120,874	9	113

Q9NXG2	THUMP domain-containing protein 1	16	6	12		39,184	7	36
P02786	transferrin receptor protein 1	12	7	8		84,871	8	50
Q99536	synaptic vesicle membrane protein VAT-1homolog	24	7	14		41,920	4	20
P27695	DNA-(apurinic or apyrimidinic site) lyase (Ref-1)	58	15	76	*DNA Damage Response/DNA Repair*	35,423	7	29
P13010	X-ray repair cross-complementing protein 5 (Ku 80/86 or XRCC-5)	20	12	30		82,573	10	63
Q92466	DNA damage-binding protein 2	23	8	12		47,864	11	26
P09429	high-mobility group protein B1 (HMG B1)	40	8	13		24,763	3	43
P43246	DNA mismatch repair protein 2	5	5	8		104,743	13	70
P29372	DNA-3-methyladenine glycosylase	46	9	17		32,869	7	12
P09874	poly (ADP-ribose) polymerase1 (PARP-1)	25	22	82		112,953	71	126
P46063	ATP-dependent DNA helicase Q1	7.9	4	10		73,326	18	60
P23246	splicing factor, proline- and glutamine- rich	25	15	22		76,149	2	33
P12004	proliferating cell nuclear antigen	23	4	7		28,769	6	16
P04818	thymidylate synthetase	23	6	6		35,585	5	15
P12956	X-ray repair cross-complementing protein 6 (Ku 70 or XRCC-6)	26	19	56		69,712	5	59
Q16666	interferon, γ-inducible protein 16	12	6	14		88,256	8	89
P39748	flap endonuclease 1 (Fen-1)	38	12	39		42,593	6	38
P28340	DNA polymerase δ catalytic subunit	3.8	4	6		123,631	26	46
Q92879	CUGBP Elav-like family member1	8.2	4	7	*RNA Processing/mRNA Splicing*	52,063	7	24
Q9Y2L1	exosome complex exonuclease RRP44	8.7	6	9		109,003	19	70
Q9UQ80	proliferation-associated protein2G4	37	15	35		43,656	6	42
P60842	eukaryotic initiation factor 4A-I	24	11	24		46,154	4	20
P38919	eukaryotic initiation factor 4A-III	14	6	7		46,740	5	23
Q14240	eukaryotic initiation factor 4A-II	19	8	22		46,402	4	22
P35637	RNA-binding protein FUS	14	8	30		53,426	4	14
P09651	heterogeneous nuclear ribonucleoprotein A1	41	10	25		38,715	2	18
Q99729	heterogeneous nuclear ribonucleoprotein A/B	24	6	20		36,225	2	29
P22626	heterogeneous nuclear ribonucleoprotein A2/B1	25	10	16		37,430	1	19
Q14103	heterogeneous nuclear ribonucleoprotein D	33	9	28		38,434	3	32

P61978	heterogeneous nuclear ribonucleoprotein K	19	7	26		50,976	5	22
P14866	heterogeneous nuclear ribonucleoprotein L	22	6	18		64,132	11	29
P52272	heterogeneous nuclear ribonucleoprotein M	41	22	83		77,384	5	40
O60506	heterogeneous nuclear ribonucleoprotein Q	17	12	23		69,471	1	26
Q0VGD6	heterogeneous nuclear ribonucleoprotein R	9.3	5	13		67,861	7	42
Q12905	interleukin enhancer-binding factor 2	22	8	12		43,062	4	17
Q12906	interleukin enhancer-binding factor 3	24	19	48		95,338	7	65
Q08J23	tRNA (cytosine-5-)-methyltransferase NSUN2	11	5	12		86,471	14	57
Q15233	non-POU domain-containing octamer-binding protein	35	13	35		54,232	2	27
P11940	polyadenylate-binding protein 1	8.3	4	7		70,671	4	41
Q15365	poly(rC) binding protein 1	21	5	11		37,498	9	14
Q68Y55	poly(rC)-binding protein 2	24	4	15		34,917	6	17
P26599	polypyrimidine tract-binding protein 1	31	14	51		57,221	3	33
Q8NC51	plasminogen activator inhibitor 1 RNA-binding protein	15	5	21		44,834	1	35
Q15637	splicing factor 1	16	8	22		68,199	4	27
Q15459	splicing factor 3a, subunit 1	14	7	22		88,886	1	59
Q15428	splicing factor 3a, subunit 2	14	4	5		49,256	2	22
P61221	ATP-binding cassette, sub-family E, member 1	8.7	5	6		67,314	17	53
Q8NE71	ATP-binding cassette, sub-family F, member 1	6.9	5	10		95,926	8	101
O43776	asparaginyl-tRNA synthetase	11	6	9	*Transcriptional Regulation/Translation*	62,943	15	35
P82979	SAP domain-containing ribonucleoprotein	25	4	5		23,540	1	26
Q92499	ATP-dependent RNA helicase DDX1	13	8	15		82,432	17	62
O00148	ATP-dependent RNA helicase DDX39	13	7	15		48,998	9	28
Q59F66	DEAD box polypeptide 17 isoform 3	12	5	12		81,067	12	31
P68104	elongation factor 1α1 (EF-1α1)	33	13	99		50,141	6	47
P29692	elongation factor 1δ	9.3	4	4		30,991	2	20
P13639	elongation factor 2	17	14	25		95,207	17	65
P26641	elongation factor 1γ	15	8	29		49,988	6	32
Q9BY44	eukaryotic translation initiation factor 2A	29	12	18		64,859	10	55

Q04637	eukaryotic translation initiation factor 4γ	8.4	9	15		175,535	13	96
Q96I24	far upstream element-binding protein 3	11	5	11		61,509	6	24
Q96AE4	far upstream element-binding protein 1	27	15	39		67,429	3	28
P78347	general transcription factor II-I	15	12	26		112,285	9	85
Q8N1G4	leucine rich repeat-containing protein 47	11	5	12		63,342	10	34
P49736	DNA replication licensing factor MCM2	8.3	7	13		101,765	12	43
P25205	DNA replication licensing factor MCM3	13	8	12		90,850	11	61
P33991	DNA replication licensing factor MCM4	9.8	9	12		96,558	10	45
P33992	DNA replication licensing factor MCM5	7.2	4	9		82,286	14	47
O75475	PC4 and SFRS1-interacting protein	11	6	22		60,103	2	84
Q9Y285	phenylalanyl-tRNA synthetase, α chain	19	10	11		57,432	1	28
Q9NSD9	phenylalanyl-tRNA synthetase, β chain	7.5	4	5		66,130	11	46
Q9NTZ6	RNA-binding protein 12	7.1	6	10		97,395	4	39
P49591	seryl-tRNA synthetase	17	7	7		58,646	9	45
Q7KZF4	staphylococcal nuclease domain-containing protein 1	8.9	6	12		101,867	12	58
Q8TAQ2	SWI/SNF complex subunit SMARCC2	5.3	4	4		132,879	7	92
O14776	transcription elongation regulator 1	13	13	27		123,901	4	98
Q13263	transcription intermediary factor 1β	31	19	72		88,418	31	42
P54577	tyrosyl-tRNA synthetase	31	14	23		59,012	6	61
P26640	valyl-tRNA synthetase	6.5	8	12		140,345	22	57
P49005	DNA polymerase δ subunit 2	13	4	13	*Unknown*	51,289	11	17

Figure 3.5 GO annotations for the cellular distributions (A), biological processes (B), and molecular functions (C) of proteins that form DPCs in the presence of cisplatin. The numbers of proteins in each category is labeled on the charts.



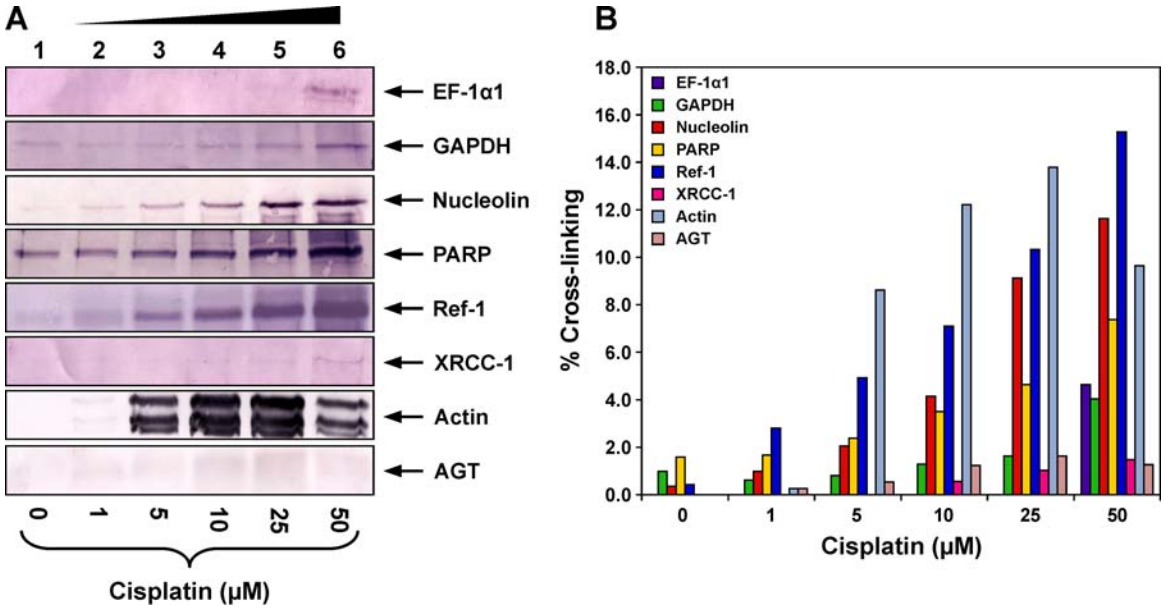
which is reasonable due to the distinct mechanisms of cross-link formation by the three *bis*-electrophiles.

The proteins identified in our proteomics screen were additionally categorized according to the molecular weight region of the gel in which they were present prior to tryptic digestion (Figure 3.3). The majority of proteins detected fell into the predicted molecular weight region. However, some proteins were also found in higher molecular weight fractions, suggesting that cisplatin can form ternary DNA-protein-protein or protein-DNA-protein complexes. Further studies using physiologically relevant concentrations of cisplatin are required to establish whether such complexes are biologically relevant.

3.3.3 Western Blot Analysis to Confirm Protein Identities

The identities of a subset of proteins which were detected by mass spectrometry-based proteomics were confirmed by western blot analysis using commercially available antibodies. Nuclear protein extracts derived from HeLa cells were incubated with biotinylated DNA in the presence of 0– 50 μM cisplatin to induce DPCs. Following biotin capture (Scheme 3.2), protein release *via* trans-platination, and SDS-PAGE separation, the proteins were transferred to nitrocellulose membranes and subjected to Western blot analysis using commercial antibodies against actin, AGT, GAPDH, PARP, Ref-1 and EF-1α1 (Figure 3.6). These target proteins were selected based on either their identification from the proteomics screen (actin, PARP, GAPDH, Ref-1, EF-1α1) or due to their previously demonstrated ability to form DPCs in the presence of other *bis*-electrophiles, such as antitumor nitrogen mustards and DEB (AGT, XRCC-1, and nucleolin) *(35;36)*. Densitometric analysis of protein signals observed in the Western blots of biotin captured mixtures and of the total nuclear extracts derived from the same number of cells was used to estimate the efficiency of DPC formation for specific proteins in the presence of cisplatin (Figure 3.6B).

Figure 3.6 Western blot analysis of proteins participating in DPC formation in nuclear protein extracts from human cervical carcinoma HeLa cells. (A) Nuclear extract proteins were incubated with 0 (lane 1), 1 (lane 2), 5 (lane 3), 10 (lane 4), 25 (lane 5) or 50 μM cisplatin (lane 6) in the presence of 5'-biotinylated double-stranded oligodeoxynucleotides (5'-GGA GCT GGT CGT GGC CTA-3' (+) strand). Following biotin capture enrichment, removal of unbound proteins, and thermal release of proteins participating in cross-linking, they were resolved by SDS-PAGE and transferred to nitrocellulose membranes. Western blot analyses were performed using primary antibodies against EF-1α1, GAPDH, Nucloin, PARP, Ref-1, XRCC-1, actin, and AGT. (B) Densitometric analysis of Western blots to estimate the extent of protein cross-linking to DNA in the presence of cisplatin.



Western blot analysis confirmed the identities of five proteins detected in mass spectrometric analysis of cisplatin-induced DPCs: actin, GAPDH, PARP, Ref-1 and EF-1α1 (Figure 3.6A). Three additional proteins not detected by mass spectrometry: AGT, XRCC-1 and nucleolin, were found to form DPCs in the presence of cisplatin (Figure 3.6A). These three proteins were not detected in the mass spectrometry-based proteomics screen (Table 3.1), probably because of their low abundance in the nucleus *(157)*. In all cases, the intensity of antigen-specific staining of streptavidin-captured fractions progressively increased as cisplatin concentration was increased from 1 to 50 μM (Figure 3.6A, lanes 2-6). Densitometric analysis revealed that between 2 and 16% of specific proteins became cross-linked to DNA in the presence of 50 μM cisplatin (Figure 3.6B), with Ref-1 forming the largest number of DPC lesions (~ 16% of total protein), followed by GAPDH (4%), PARP (7%), EF-1α1 (5%), and nucleolin (12%). As discussed above, the overall DPC yield under these conditions is < 8% (Figure 3.1B), suggesting that EF-1α1, GAPDH, PARP, nucleolin and Ref-1 are specifically targeted for cross-linking to DNA in the presence of cisplatin as compared to other nuclear proteins.

3.3.4 HPLC-ESI$^+$-MS/MS Analysis of dG-Pt-Lys Conjugates

To identify the chemical structures of cisplatin-induced amino acid-nucleoside conjugates, DNA-protein cross-links eluted from streptavidin beads were subjected to DNase I/alkaline phosphatase-mediated DNA degradation to release protein-nucleoside conjugates, followed by digestion with trypsin and proteinase K to achieve complete proteolytic digestion of proteins to amino acids. The resulting amino acid-nucleoside conjugates were enriched by off-line HPLC and analyzed by capillary HPLC-ESI$^+$-MS/MS. Our earlier studies conducted with various amino acids (not shown) indicated that cisplatin-induced DNA-protein cross-linking took place primarily between the lysine amino side chain within proteins and the N7 position of guanine in DNA. Therefore, our detection efforts have focused on *cis*-1,1-diammine-2-(5-amino-5-

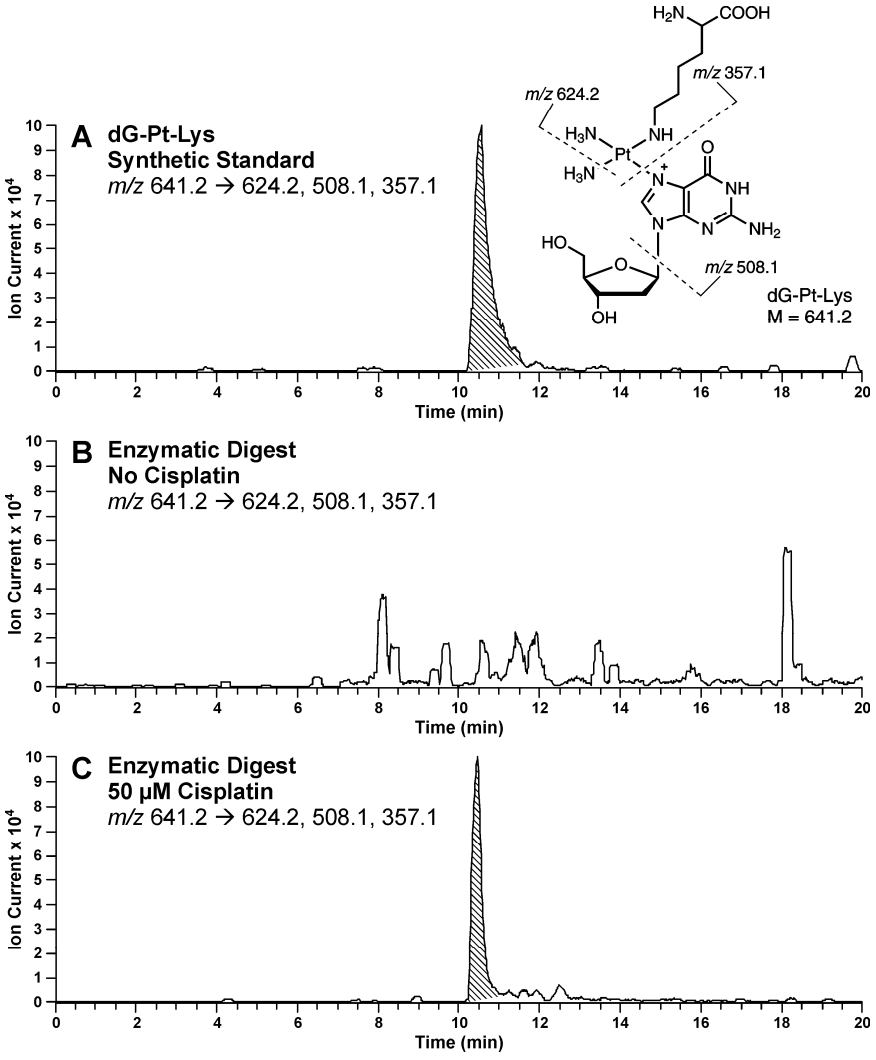
carboxypentyl)amino-2-(2'-deoxyguanosine-7-yl)-platinum(II) (dG-Pt-Lys) (Scheme 3.2). Authentic standard of dG-Pt-Lys was prepared by incubating 2'-deoxyguanosine with Boc-protected lysine in the presence of activated cisplatin, followed by deprotection and HPLC purification (*see Chapter II*).

HPLC-ESI$^+$-MS/MS analysis of dG-Pt-Lys in enzymatic digests was conducted in the selected reaction monitoring (SRM) mode by following three characteristic mass transitions (*m/z* 641.2 → *m/z* 624.2, 508.1, and 357.1) corresponding to a loss of NH_3, the cleavage of the glycosidic bond, and the loss of dG, respectively (Figure 3.7A). dG-Pt-Lys were detected in cisplatin-treated nuclear samples (Figure 3.7C), but not in the untreated control (Figure 3.7B). These results indicate that cisplatin can form covalent cross-links between the N7 position of guanine in duplex DNA and the side-chain ε-amino group of lysine in proteins. All of proteins identified in the present work contain multiple lysine residues (Table 3.1).

3.4 Discussion

Cisplatin and its analogues are important antitumor drugs capable of cross-linking cellular biomolecules. They bind to the N-7 positions of purines in DNA, producing mostly intrastrand-1,2-d(GpG) and 1,2-d(ApG) cross-links *(70;143)*. Interestingly, non-covalent binding of nuclear proteins to cisplatin-damaged DNA is thought to play an important role in toxicity of these compounds *(158)*. For example, binding of high mobility group (HMG) proteins to cisplatin-induced intrastrand-1,2-GG adducts and 1,2-AG adducts blocks their repair by nucleotide excision repair proteins and inhibits transcription, causing tumor-selective toxicity *(71-73;159)*. Recently, Lippard *et al.* employed photoreactive aryl azide group-containing PtBPn complexes and biotinylated DNA probes to affinity purify and identify 14 nuclear proteins that bind to cisplatin-DNA adducts (1,2-intrastrand DNA-DNA cross-links), including PARP-1,

Figure 3.7 HPLC-ESI$^+$-MS/MS analysis of 1,1-diammine-2-(5-amino-5-carboxypentyl)amino-2-(2'-deoxyguanosine-7-yl)-platinum(II) (dG-Pt-Lys) conjugates in total proteolytic digests of cisplatin-induced DPCs. Nuclear protein extracts from HeLa cells were exposed to cisplatin in the presence of biotinylated DNA duplexes. Following biotin capture enrichment, DPCs were subjected to enzymatic digestion to release amino acid-nucleoside conjugates. Shown here are HPLC-ESI$^+$-MS/MS traces of synthetic dG-Pt-Lys (A); enzymatic digests of HeLa nuclear protein extracts following incubation with DNA in the absence of cisplatin (negative control) (B); and enzymatic digests of HeLa nuclear protein extracts following incubation with 50 μM cisplatin (C).



HMG B1/B2/B3, DNA ligase III, XRCC-1, ATP-dependent DNA helicase (Ku) 70/86, and human DNA mismatch repair protein (hMutS) *(160-162)*. However, these and other proteins may also be trapped to cisplatin-DNA monoadducts, potentially leading to their irreversible cross-linking to DNA and the formation of covalent DNA-protein lesions.

Perhaps due to their inherent complexity, cisplatin-induced DNA-protein cross-links are poorly characterized, and their potential role in toxicity is not understood. Previous studies have shown that several DNA-binding proteins, including HMG proteins 1, 2, and E, and histone proteins, can become cross-linked to DNA in the presence of cisplatin *(71-73;159)*. However, these studies employed antibodies against specific proteins and therefore could not identify other potential targets of cisplatin-induced cross-linking or determine the structures of the resulting macromolecular lesions.

In the present work, we employed an affinity capture approach developed in our laboratory *(35;36)* to isolate human nuclear extract proteins forming DPCs to biotinylated DNA duplexes in the presence of cisplatin (Scheme 3.2). Using mass spectrometry-based proteomics, we identified 131 HeLa nuclear proteins which became trapped on DNA following exposure to cisplatin (Table 3.1). As shown in Table 3.1, those identified proteins were classified according to their primary cellular functions: (1) DNA damage response and repair (15, 11% of total proteins), (2) RNA splicing and processing (29, 22% of total proteins), (3) cell signaling, cell motility and cell architecture (15, 11% of total proteins), (4) cell cycle and homeostasis (41, 31% of total proteins), and (5) transcriptional regulation and translation (30, 23% of total proteins). HPLC-ESI^+-MS/MS analyses of total proteolytic digests revealed that cisplatin-mediated DPCs are formed by the irreversible binding of cisplatin to the N7 position of guanine in DNA and the amino acid side chain of lysine residues with nuclear proteins (Figure 3.7).

A number of factors can influence the propensity of specific proteins to form DPC lesions, including their abundance, cellular distribution, and their participation in specific biological

processes/molecular functions involving DNA. Many of the proteins found to form cisplatin-mediated DPCs are known to associate with DNA in the nucleus due to their involvement in DNA replication, DNA damage response, and DNA repair (Table 3.1). For example, apurinic-apyrimidinic endonuclease 1 (Ref-1) is involved in DNA damage signaling response, and has been previously shown to form cross-links to DNA in the presence of nitrogen mustards *(36)*, 1,2,3,4-diepoxybutane (DEB) *(35)*, and formaldehyde *(163)*. In contrast, other proteins identified in our screen, *e.g.* GAPDH, actin, and enolase (Table 3.1), are known to participate in cellular processes not involving DNA (*e.g.* glycolysis, cell motility). However, it is important to note that many those identified proteins may have additional cellular roles and may participate in multiple biological processes, which would explain their ability to associate with DNA and form DPCs. For example, enolase is primarily known for its role in cell glycolysis, but is also reported to participate in DNA binding and transcriptional regulation *(164)*.

The majority of the proteins participating in cisplatin-mediated cross-linking (Table 3.1) are unique to this drug and have not been detected in our previous studies of nitrogen mustard *(36)* and DEB *(35)* mediated cross-linking. Only 19% of the proteins identified in the present work (including PARP, HMG, Ku and hMutS) are also targeted for cross-linking to DNA in the presence of nitrogen mustards or DEB *(35;36)*, while the remaining proteins appear to be specific to cisplatin-mediated cross-linking. This may be a result of different mechanisms for DPC formation employed by these drugs. For example, mechlorethamine and DEB preferentially target proteins with reactive thiol groups, leading to cysteine-guanine cross-links *(54;55)*. In contrast, our studies with free amino acids (*see Chapter II*) indicate that cisplatin is more reactive towards basic amino acids such as lysine.

Our proteomics screen identified high mobility group proteins as one of the targets for cisplatin-mediated cross-linking to DNA, consistent with a previous study *(159)*. In contrast to an earlier study *(71;72)*, no DPC lesions involving histone proteins were detected, which could be a

result of their low concentrations in our nuclear protein extracts which were prepared according to Jessberger *et. al*. *(139)*, their inefficient cross-linking to DNA, or their inefficient association with DNA duplexes used in the present study. A close examination of our results shown in Figure 3.5 reveals that many DPC-forming proteins are classified as RNA-binding involved in RNA processing, RNA splicing, and transcription (Figure 3.5 and Table 3.1). However, since our experiments included biotinylated DNA, those RNA-binding proteins are hypothesized to have DNA-binding capabilities which can be triggered by cisplatin-induced DNA damages.

The sequence of 18-mer DNA duplexes employed in the present study is derived from the frequently mutated region of the *K-ras* protooncogene (condons 10-15, 5'-GGA GCT GGT GGC GTA GGC-3'). We previously employed this DNA sequence for biotin capture of DPCs induced by other cross-linking agents, *e.g.* mechlorethamine *(36)* and DEB *(35)*, to allow for a direct comparison between different electrophiles. While it is possible that the selection of DNA sequence affects the types of proteins captured, systematic evaluation of the effects of DNA sequence on DPC formation was beyond the scope of the present investigation.

The results obtained from present work suggest that, in addition to DNA-DNA cross-linking, the formation of DPC lesions may be responsible for the biological activity of cisplatin. If not repaired, bulky DPC lesions are expected to interfere with critical DNA functions and disrupt key DNA-protein interactions, resulting in cytotoxic and genotoxic effects *(1)*. Some types of DPC lesions, such as those induced by metals, have been shown to persist through several cycles of DNA replication *(1;37;38)*. While cisplatin-induced DNA-DNA cross-links are repaired in cells primarily through nucleotide excision repair (NER) *(165)* and mismatch repair *(166)*, it is reasonable to expect that different repair mechanisms are required for the removal of cisplatin-induced DPCs due to their bulky nature. It has been reported that depending on their structure and protein size, DPCs can be repaired by nucleotide excision repair (NER), homologous recombination (HR), and proteolytic degradation *(45;167;168)*. Nakano *et al.*

reported that NER was involved in the removal of cross-linked proteins of relatively low molecular weight (<14 KDa), whereas HR repairs cross-links involving larger proteins *(45;167)*. This might suggest that the majority of DPCs induced by cisplatin are expected to be repaired by HR, due to their significant size (Table 3.1). However, future studies involving site specific cisplatin lesions are needed to test this hypothesis.

In conclusion, our study demonstrates the ability of cisplatin to covalently trap a number of nuclear proteins on DNA strands. At clinically relevant concentrations, cisplatin is a potent DPCs-inducing agent capable of irreversibly crosslinking 131 Hela nuclear proteins to DNA, therefore, it is reasonable to predict that cisplatin also forms DPC lesions within cells of cancer patients clinically treated with this drug, potentially contributing to both on-target and off-target toxicity of platinum compounds.

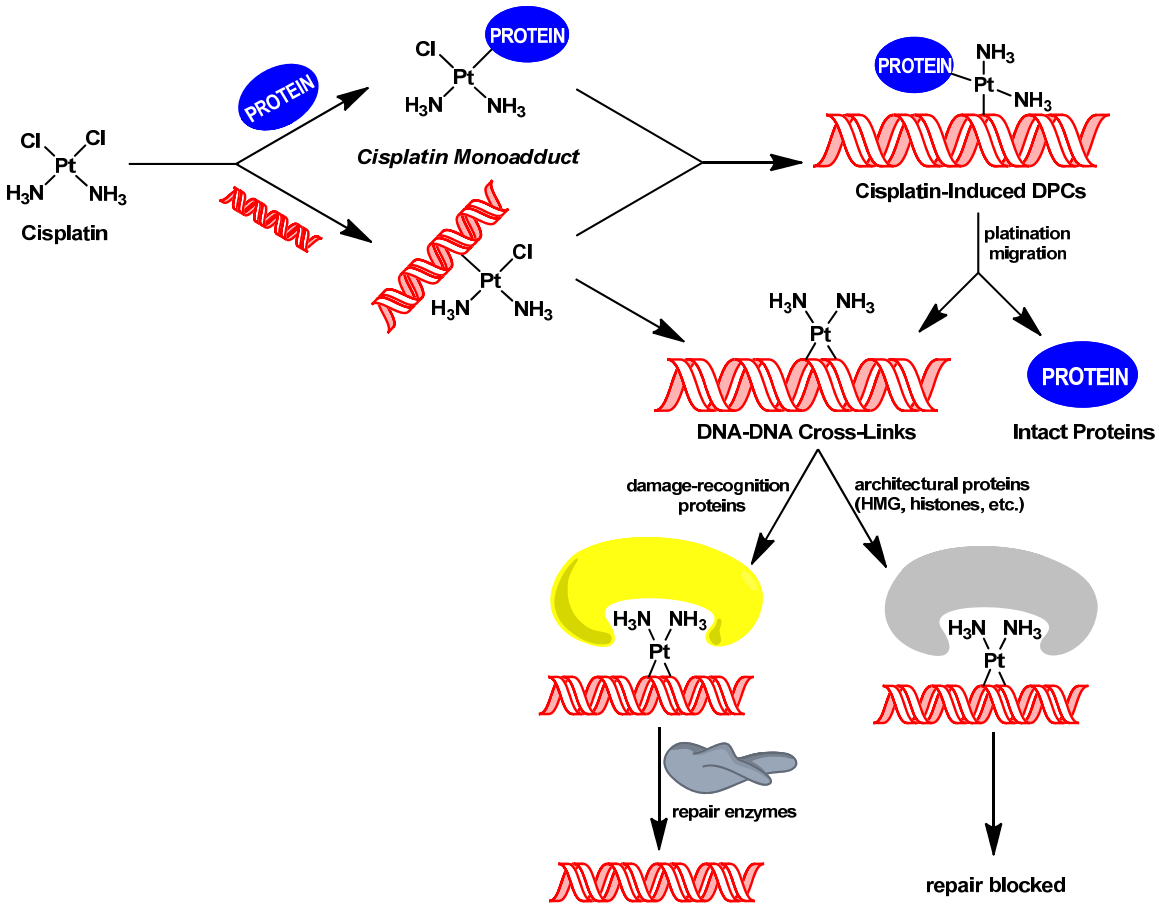
IV. Cisplatin-Induced DNA-Protein Cross-Linking in Human Fibrosarcoma (HT1080) Cells

4.1 Introduction

DNA-protein cross-links (DPCs) are bulky, helix-distorting lesions that are formed upon sequential alkylation of DNA and proteins by endogenous and exogenous *bis*-alkylating agents *(1)*. These macromolecular lesions can block DNA-protein interactions, interfering with basic cellular functions such as DNA replication, transcription, repair, recombination, and chromatin remodeling *(1;38)*. DPCs can be induced by a variety of cytotoxic, mutagenic, and carcinogenic agents including ionizing radiation *(24)*, some metals *(22)*, and common chemotherapeutic agents such as nitrogen mustards *(57;108)*, cisplatin and its analogs *(107)*, and chloroethyl nitrosoureas *(34)*. If left unrepaired, DPCs may persist through several cycles of replication, resulting in permanent DNA alterations and damaging cytotoxic and mutagenic effects *(1)*.

Platinum-based antitumor agents, *e.g.* 1,1,2,2-*cis*-diamminedichloroplatinum (II) (cisplatin) and its analog *cis*-diammine-1,1-cyclobutanedicarboxylate platinum(II) (carboplatin), are highly effective in clinical treatment of testicular and ovarian malignancies, bladder, cervical, head and neck, esophageal, and small cell lung cancer *(169;170)*. Upon entering cell nucleolus, cisplatin is spontaneously hydrolyzed *(171)*, yielding a highly reactive aquated species capable of platinating DNA to form a variety of nucleobase adducts *(109;111;112)*. The monofunctional DNA adducts formed initially can further react with neighboring bases to produce intrastrand and interstrand DNA-DNA cross-links. Alternatively, the monofunctional adducts can be trapped by nuclear proteins found in proximity to chromosomal DNA to form DPCs (Scheme 4.1) *(112)*. While cisplatin-induced DNA-DNA cross-links, including 1,2-d(GpG) intrastrand cross-links, 1,2-d(ApG) intrastrand cross-links, and 1,3-d(GpNpG) intrastrand cross-links *(109;111;112)*, are

Scheme 4.1 Formation of DNA-DNA cross-links and DPCs by cisplatin.



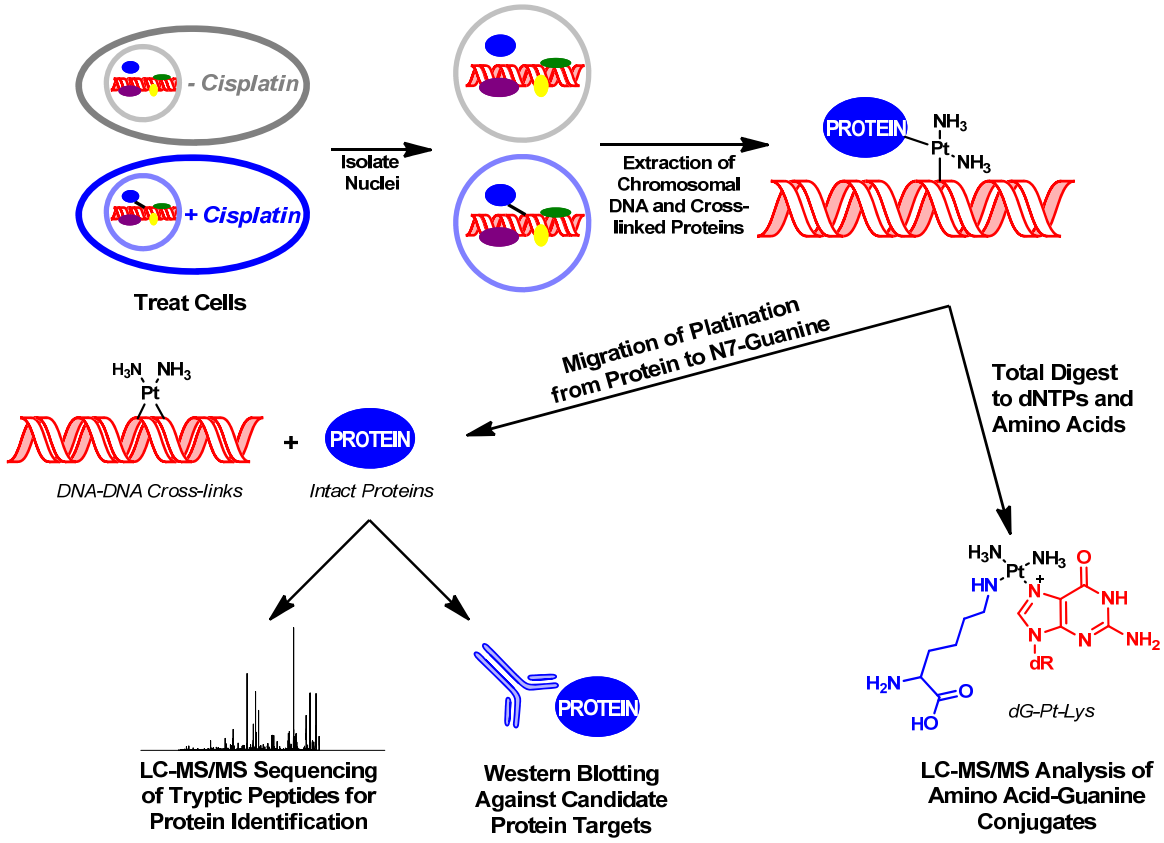
well characterized and are thought to play a prominent role in their antitumor effects *(172;173)*, little is known about the identities of the corresponding DPC lesions.

While the role of DNA-protein cross-linking in the biological effects of cisplatin is unknown, several studies using primarily biophysical methods have previously shown a relationship between DPC formation by cisplatin and cell killing *(32;71-74;174)*. It is hypothesized that bulky DPC lesions are capable of triggering cytotoxic mechanisms that lead to the death of cancer cells. However, the identities of cellular proteins participating in DPC formation by cisplatin DNA-protein lesions are not well characterized, and their role in therapeutic activity of this drug has not been established.

Previous studies have revealed that a number of nuclear proteins are capable of binding to cisplatin-damaged DNA (e.g. T4 endonuclease, DNA photolyase, human DNA mismatch repair protein (hMutS), and HMG proteins), providing a possible explanation to selective cytotoxicity of cisplatin towards certain types of tumors *(109;158;172)*. Recently, Lippard *et al.* employed photoreactive aryl azide group-containing PtBPn complexes and biotinylated DNA probes to affinity purify and identify 14 nuclear proteins that bind to cisplatin-DNA adducts, including PARP-1, HMG B1/B2/B3, DNA ligase III, XRCC-1/5/6, and hMutS *(160-162)*. While some of those described proteins that can specifically recognize cisplatin-damage DNA site may be directly or indirectly involved in DNA damage recognition, others are generally abundant nuclear or chromatin proteins that perform structural functions in a cell. Binding of those proteins to cisplatin-damaged DNA sites interferes with their repair by preventing the access of damage recognition proteins (Scheme 4.1) *(158)*. Of these proteins, the high-mobility group (HMG) family proteins *(159;175)* and the linker histones *(176)* are two representative classes that preferentially bind to cisplatin-modified DNA with high affinity, which has been postulated to mediate the antitumor properties of cisplatin (Scheme 4.1).

We previously employed an affinity capture-based approach to investigate cisplatin-mediated DPCs formation *in vitro*. 131 proteins from human cervical carcinoma (Hela) nuclear protein extracts were found to be cross-linked to DNA by cisplatin (*see Chapter III*). However, the *in vitro* study involved a short synthetic DNA duplex that was lacking normal chromosomal structure. The goal of the present work was to characterize cisplatin-mediated DPC formation in cultured human cells. We anticipated that the presence of intact chromatin structure and normal DNA-protein interactions in cells would lead to an increased efficiency of cisplatin-induced DPC formation. DPCs were isolated from treated cells using a modified phenol/chloroform extraction methodology developed in our laboratory (Scheme 4.2) *(177)*. Proteins were then released by heating due to the platination migration from proteins to DNA to form DNA-DNA cross-links (*see Chapter II*), resolved by gel electrophoresis, and identified by bottom-up mass spectrometry-based proteomics and western blotting. A total of 256 cross-linked proteins were identified, including a number of nuclear proteins that play an important role in DNA repair, transcriptional regulation, and chromatin remodeling, *e.g.* poly(ADP-ribose) polymerase 1 (PARP-1), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), x-ray cross-complementing (XRCC) proteins, high mobility group (HMG) proteins, and histone proteins. Finally, high performance liquid chromatography-electrospray ionization-tandem mass spectrometry (HPLC-ESI^+-MS/MS) analyses of total proteolytic digests revealed the presence of 1,1-*cis*-diammine-2-(5-amino-5-carboxypentyl)-amino-2-(2'-deoxyguanosine-7-yl)-platinum(II) (dG-Pt-Lys) conjugates in HT1080 cells treated with cytotoxic concentrations of cisplatin as a result of sequential platination of lysine residues within the proteins and guanine nucleobases in DNA by cisplatin. Taken together, these results indicate that cisplatin is capable of forming covalent DPCs *in vivo* involving a range of nuclear proteins. These findings are significant because the super-bulky DPC lesions are likely to contribute to the cytotoxic and mutagenic effects of platinum-based drugs.

Scheme 4.2 Strategy for the isolation and analysis of DPCs from cisplatin-treated mammalian cell cultures.



4.2 Materials and Methods

Chemicals and Reagents - 1,1,2,2-*Cis*-diamminedichloroplatinum(II) (cisplatin), leupeptin, pepstatin, aprotinin, phenylmethanesulfonyl fluoride (PMSF), dithiothreitol (DTT), iodoacetamide, chloroform, ribonuclease A, nuclease P1, phosphodiesterase I (PDE I), phosphodiesterase II (PDE II)and alkaline phosphatase were purchased from Sigma (St. Louis, MO). Mass spectrometry-grade Trypsin Gold was purchased from Promega (Madison, WI). Proteinase K was obtained from New England Biolabs (Beverly, MA). Primary polyclonal antibodies specific for glyceraldehyde 3-phosphate dehydrogenase (GAPDH), flap endonuclease 1 (Fen-1), nucleoin, actin, poly-(ADP-ribose) polymerase 1 (PARP-1), elongation factor 1α1 (EF-1α1), and DNA-(apurinic- or apyrimidinic-site) lyase (Ref-1) were obtained from Santa Cruz Biotechnology (Santa Cruz, CA). The monoclonal antibody specific for x-ray cross-complementing protein 1 (XRCC-1) was purchased from Lab Vision/NeoMarkers (Fremont, CA). The primary monoclonal antibody against AGT was purchased from Millipore (Temecula, CA). Alkaline phosphatase-conjugated anti-mouse and anti-rabbit secondary antibodies were purchased from Sigma (St. Louis, MO). *Cis*-1,1-diammine-2-(5-amino-5-carboxypentyl)-amino-2-(2'-deoxyguanosine-7-yl)-platinum(II) (dG-Pt-Lys) was prepared as described previously (*see Chapter II*).

Cell Culture - Human fibrosarcoma (HT1080) cells *(178)* were obtained from the American Type Cell Culture Collection. The cells were maintained as exponentially growing monolayer cultures in Dulbecco's modified Eagle's medium supplemented with 9% fetal bovine serum (FBS) maintained in a humidified incubator at 37ºC with 5% CO_2.

Assay for Cytotoxicity as a Result of Cisplatin Exposure - HT1080 cells were plated in Dulbeccos modified Eagle's medium containing 9% FBS at a density of 5 x 10^5 cells/dish and permitted to adhere overnight. On the following morning, cells (in triplicate) were treated with cisplatin (0, 5, 10, 50, or 100 μM) for 3 h at 37ºC. Following treatment, cells were placed in a drug-free media and allowed to recover for 18 h. The effect of cisplatin on cell survival was determined by direct cell counting in a haemocytometer. Cytotoxicity was expressed as the number of cells surviving cisplatin treatment relative to buffer-treated controls.

Isolation of Proteins Cross-Linked to Chromosomal DNA by Cisplatin - To analyze DPC formation in mammalian cells exposed to cisplatin, HT1080 cells were treated with increasing concentrations of cisplatin (0, 10, 50, 100, 250, or 500 μM) for 3 h at 37ºC. Following exposure, the cells were washed with ice cold phosphate-buffered saline (PBS) and re-suspended in PBS to a final density of ~2 x 10^6 cells/mL. To isolate nuclei, cells were lysed by adding an equal volume of 2X cell lysis buffer (20 mM Tris-HCl/10 mM $MgCl_2$/2% v/v Triton-X100/0.65 M sucrose), incubated on ice for 5 min, and centrifuged at 2,000 *g* for 10 min at 4 ºC. The nuclear pellet was re-suspended in a saline-EDTA solution (75 mM NaCl/24 mM EDTA/1% (w/v) SDS, pH 8.0) containing RNase A (10 μg/mL) and a protease inhibitor cocktail (1 mM PMSF; 1 μg/mL pepstatin; 0.5 μg/mL leupeptin; 1.5 μg/mL aprotinin) to a concentration of ~ 5 x 10^6 nuclei/mL and incubated for 2 h at 37ºC with gentle shaking. To isolate chromosomal DNA containing covalent DPCs, nuclear lysates were extracted with two volumes of Tris-buffer saturated phenol, and the resulting white emulsion was centrifuged for at 1,000 *g* for 15 min at room temperature. The aqueous layer and the interface material were subjected to a second extraction with two volumes of Tris buffer saturated phenol:chloroform (1:1). DNA was precipitated with cold ethanol. Samples were centrifuged at 4,000 *g* for 20 min at 4 ºC, and the resulting DNA pellet

was washed with cold 70% ethanol, air dried, and reconstituted in 1 mL MilliPore water. DNA concentrations were estimated by UV spectrophotometry. DNA amounts and its purity were determined by HPLC-UV analyses of enzymatic hydrolysates as described previously *(177)*.

Mass Spectrometric Identification of Cross-Linked Proteins - To identify cellular proteins that become covalently attached to chromosomal DNA in cisplatin-treated cells, HT1080 cells (~10^7 cells, in triplicate) were treated with 100 μM cisplain or solvent control for 3 h at 37 ºC. Chromosomal DNA containing covalently cross-linked proteins was isolated by modified phenol/chloroform extraction methodolody and quantified as described above. DNA (30 μg) was dissolved in 50 μL of 1X NuPAGE Sample Buffer (Invitrogen, Carlsbad, CA) and heated at 70 ^{o}C for 1 h to release the cross-linked proteins. Our earlier studies revealed that intact proteins can be released from cisplatin-induced DPCs by platination migration to nearby nucleobases within genomic DNA (*Chapter II* of this thesis). The resulting proteins were separated using 12% Tris-HCl Ready Gels (Bio-Rad, Hercules, CA) and stained with SimplyBlue SafeStain (Invitrogen, Carlsbad, CA). Gel lanes were divided into five sections encompassing the entire molecular weight range, and the gel sections were further diced into ~1 mm pieces. The proteins present within the gel pieces were subjected to in-gel tryptic digestion as described previously *(140)*. In brief, gel pieces were rinsed with 25 mM ammonium bicarbonate, and the protein thiols were subjected to reduction with DTT (300 mM) and alkylation with iodoacetamide. The gel pieces were then dehydrated by incubation with acetonitrile, dried under vacuum, and reconstituted in 25 mM ammonium bicarbonate buffer. Mass spectrometry grade trypsin (2-3 μg) (Promega, Madison, WI) was added, and the samples were digested overnight at 37 ºC. The resulting tryptic peptides were extracted with 60% acetonitrile containing 0.1% aqueous formic acid, evaporated to dryness, desalted by ZipTip C18 purification (ZipTip C18 Pipette Tips, Millipore, Temecula, CA), and finally reconstituted in 0.1% formic acid (25 μL) prior to MS analysis.

Tryptic peptides were analyzed by HPLC-ESI$^+$-MS/MS with a Thermo Scientific LTQ Orbitrap Velos mass spectrometer in line with an Eksigent NanoLC-Ultra 2D HPLC system, a nanospray source, and Xcalibur 2.1.0 software for instrument control. Peptide samples (8µL) were trapped on a 180 µm x 20mm Symmetry (Waters, Milford, MA) C18 column with 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B) at a flow composition of 95% A and 5% B at 5 µL/min for 3 minutes. The flow was reversed after trapping and decreased to 0.3 µL/min, and the peptides were eluted off the trap column and onto a capillary column (75 µm ID, 10 cm packed bed, 15 µm orifice) created by hand packing a commercially purchased fused-silica emitter (New Objective, Woburn MA) with Zorbax SB-C18 5 µm separation media (Agilent, Santa Clara, CA). The solvent composition was initially set at 5% B, followed by a linear increase to 60% B over 60 min, and further to 95% B in 5 min. Liquid chromatography was carried out at an ambient temperature. Centroided MS-MS scans were acquired using an isolation width of 2.5 *m/z*, an activation time of 30 ms, an activation Q of 0.25, 35% normalized CID collision energy, and 1 microscan with a max ion time of 100 ms for each MS/MS scan. The mass spectrometer was mass calibrated prior to each analysis, and the spray voltage was adjusted to assure a stable spray. Typical MS parameters were as follows: spray voltage of 1.6 kV, a capillary temperature of 275°C, and an S-lens RF Level of 50%. Peptide MS/MS spectra were collected using data-dependent scanning in which one full scan mass spectrum was followed by eight MS/MS spectra. Dynamic exclusion was enabled for 60 s and singly charged species were excluded.

Mass spectral data were analyzed using an in-house developed software pipeline, TINT, that linked raw data extraction, database searching, and probability scoring. Raw data were extracted and converted to the mzXML format using ReadW. Spectra that contained fewer than 6 peaks or had less than 20 measured total ion current (TIC) were excluded. Data were searched using the SEQUEST v.27 algorithm *(123;124)* on a high speed, multiprocessor Linux cluster in

the Minnesota Super Computing Institute at University of Minnesota using the human subset consisting of the NCBI derived human protein database v200806 combined with its reversed counterpart along with common protein contaminants totaling 70,711 entries. Search parameters included trypsin specificity and up to 5 missed cleavage sites. Cysteine carboxamidomethylation (+57.0215 Da) was set as a fixed modification, and methionine oxidation (+15.9949 Da) was set as a variable modification. Precursor mass tolerance was set to 1.25 *m/z* within the calculated average mass, and fragment ion mass tolerance was set to 0.5 *m/z* of their monoisotopic mass. Identified peptides were filtered using Scaffold 3 software (Proteome Software, INC., Portland, OR) *(141)*, to a target false discovery rate (FDR) of 5%. The FDR was calculated with the following expression: FDR = (2R)/(R+F)*100, where R is the number of passing reversed peptide identifications and F is the number of passing forward (normal orientation) peptide identifications. The second round of filtering removed proteins supported by less than four distinct peptide identifications in the analyses. Indistinguishable proteins were recognized and grouped. Parsimony rules were applied to generate a minimal list of proteins that explained all of the peptides that passed our entry criteria *(142)*. Furthermore, T-test analyses were performed to ensure that the levels of proteins captured from treated samples were significantly higher than those in untreated controls. All statistical analyses were conducted in Scafflod version 3.0. The significance level was set at 5%.

Western Blots Analysis of Identified Proteins - HT1080 cells (~10^7) were treated with cisplatin (0, 10, 50, 100, 250, or 500 μM) for 3 h at 37 ºC. Chromosomal DNA, along with any covalently bound proteins, was extracted and quantified as described above. Approximately 30 μg of DNA from each sample was dissolved in gel loading buffer and heated for1 h at 70 ºC to partially release the proteins *via* platination transfer. Proteins were separated by 12% SDS-PAGE and transferred to Trans-blot nitrocellulose membranes (Bio-Rad, Hercules, CA). Following blocking

in Tris-buffered saline (TBS) containing 5% (w/v) bovine serum albumin, the membranes were incubated with the primary antibody against target protein for 3 h at room temperature, rinsed with TBS buffer, and incubated overnight at 4 ºC with the corresponding alkaline phosphatase-conjugated secondary antibody. The blots were washed and developed with SIGMA Fast BCIP/NBT (Sigma, St. Louis, MO) according to the manufacturer's instructions. The developed blots were scanned as image files. ImageJ software (available free of charge from the NIH website, www.ncbi.nlm.nih.gov) was used to quantify the optical densities of the protein bands. The efficiency of DNA-protein cross-linking was approximated by comparing signal intensities of the protein which was co-purified with chromosomal DNA (corresponding to cross-linked protein) and the intensity of the corresponding protein band present in the whole cell protein lysate (representing total cellular proteins) *(177)*.

HPLC-ESI$^+$-MS/MS analysis of dG-Pt-Lys in cells exposed to cisplatin - HT1080 cells (~10^6) were treated with 100 μM cisplain or solvent control for 3 h at 37 ºC. Chromosomal DNA was isolated using the modified phenol/chloroform extraction procedure described above. DNA (50 μg) was digested with phosphodiesterase I (240 mU), phosphodiesterase II (240 mU), DNase I (120 mU) and alkaline phosphatase (6 U) overnight at 37ºC to produce protein-nucleoside conjugates. Samples were dried under vacuum, reconstituted in 25 mM ammonium bicarbonate, and digested to peptides with trypsin (2-3 μg, 37°C overnight). To achieve complete hydrolysis to amino acids, the resulting tryptic peptides were dried under vacuum, reconstituted in water, and digested with proteinase K (20 μg) for 48 h at room temperature. The digest mixtures were subjected to off-line HPLC separation using an Agilent Technologies HPLC system (1100 model) incorporating a diode array detector and a Supelcosil LC-18-DB (4.6 x 250 mm, 5 μm) column (Sigma-Aldrich, St. Louis, MO). The column was eluted at a flow rate of 1 mL/min using 15 mM ammonium acetate, pH 4.9 (A) and acetonitrile (B). The solvent composition was changed

linearly from 0 to 24% B over 24 min and further to 60% B in 6 min. HPLC fractions containing dG-Pt-Lys (5-7 min) were collected, dried under vacuum, and reconstituted in 15 mM ammonium acetate (25 μL) for HPLC-ESI$^+$-MS/MS analysis (injection volume, 8 μL).

HPLC-ESI$^+$-MS/MS of dG-Pt-Lys conjugates was conducted with a Thermo-Finnigan TSQ Vantage mass spectrometer in line with an Eksigent MicroAS autosampler and nanoLC 2D HPLC system, a heated ESI source, and an Xcalibur 2.1.0 software for instrument control. Chromatographic separation was accomplished using a Hypercarb HPLC column (100 mm x 0.5 mm, 5 μm, ThermoScientific, Waltharm, MA) eluted with a gradient of 15 mM ammonium acetate (A) and 1:1 acetonitrile:water with 1% formic acid (B) at a flow rate of 13 μL/min. The gradient program began at 2% B, followed by a linear increase to 10% B in 10 min and further to 80% B in 8 min. The column was washed with 80% B for 5 min, and the solvent composition was brought back to 2% B in 6 min. Using this gradient, dG-Pt-Lys eluted at ~17.3 min. ESI was achieved at a spray voltage of 3.2 kV and a capillary temperature of 200 °C. CID was performed with Ar as a collision gas (2.0 mTorr) at a collision energy of 25 V. MS parameters were optimized for maximum response during infusion of a standard solution of dG-Pt-Lys and may vary slightly between experiments. HPLC-ESI$^+$-MS/MS analyses were performed in the selected reaction monitoring (SRM) mode using the mass transition corresponding to the major fragmentation pathway of dG-Pt-Lys in a triple quadrupole mass spectrometer (*m/z* 641.3 $[M]^+$ → 508.2 $[M–NH_3–\text{deoxyribose}+H]^+$, and 340.1 $[M–2NH_3–\text{deoxyguanosine}]^+$).

4.3 Results

4.3.1 Cytotoxicity Experiments

To establish the effects of cisplatin treatment on cellular viability, HT1080 cells were treated with increasing concentrations of the drug (0, 5, 10, 50, or 100 μM) for 3 h. Following

overnight incubation in a drug free media, cells were counted, and cytotoxicity was measured as the percentage of cells surviving cisplatin treatment as compared to untreated controls. Treatment with cisplatin resulted in a significant decrease in cell numbers, with approximately 70% cell death observed following treatment with 100 μM cisplatin (Figure 4.1).

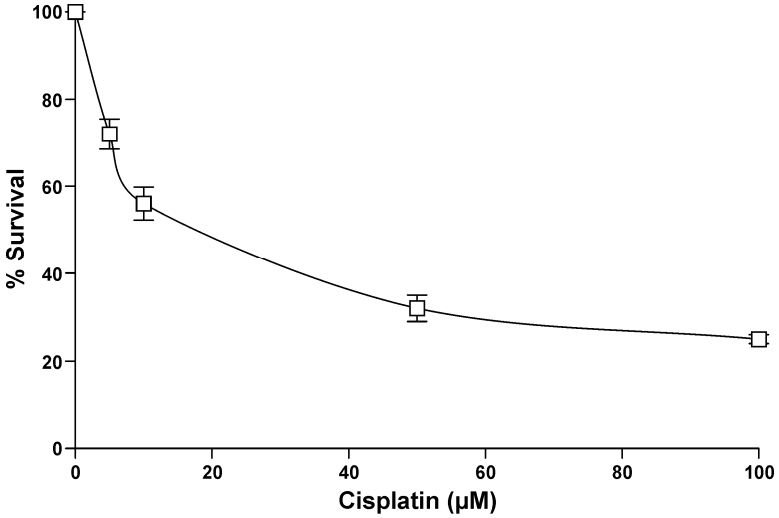
4.3.2 Identification of Cross-Linked Proteins by Mass Spectrometry-Based Proteomics

To determine the identities of the proteins participating in cisplatin-mediated DPC formation *in vivo*, HT1080 cells (~ 10^7 cells, in triplicate) were treated with 100 μM cisplatin, while control cells were incubated in standard media. Following DNA extraction by the modified phenol/chloroform extraction method developed in our laboratory (Scheme 4.2) *(177)*, ~ 30 μg of DNA from each sample was taken and heated in 1X NuPAGE loading buffer (1 h at 70 ºC). Our previous studies (*see Chapters II and III*) revealed that this treatment releases intact proteins *via* platinantion migration from protein to DNA (Scheme 4.2), simplifying protein analysis by mass spectrometry.

Proteins released from DNA were separated by SDS-PAGE (Figure 4.2). Distinct protein bands were observed for cisplatin-treated samples (lanes 6-8, Figure 4.2), while the untreated samples exhibited minimal protein signals (lanes 2-4, Figure 4.2). Protein bands within the molecular weight range of 15–250 kDa were excised from the gel and subjected to in-gel tryptic digestion *(140)*. The resulting peptides were extracted from the gel and subjected to HPLC-ESI^+-MS/MS analysis. Protein identification was based on the MS/MS spectra of tryptic peptides, yielding characteristic b- and y-series fragment ions that were used to determine their amino acid sequence (see examples in Figure 4.3).

Database searching and parsimony analysis of the MS/MS spectral data resulted in identification of 256 proteins that co-purified with chromosomal DNA from cisplatin-treated cells (Table 4.1). All protein identifications were supported by at least four unique peptides. Statistical

Figure 4.1 Cytotoxicity of cisplatin in HT1080 cells. Cells ($5x10^5$/dish, in triplicate) were treated with 0–100 μM cisplatin. Following 3 h treatment, the cell media was replaced with drug-free media, and the cells were incubated overnight before being harvested and counted. Cell survival was determined by comparing the number of cells surviving cisplatin treatment relative to untreated controls. Error bars represent the standard error from three experiments.



analyses were conducted to compare proteomics results for treated and untreated samples, and only proteins which exhibited significantly increased ion counts in treated samples ($p < 0.05$) were included in the list. As shown in Table 4.1, molecular weights of the identified proteins are generally consistent with their positions on the gel. A number of proteins were also present in a higher molecular weight fraction than expected, probably representing protein-protein cross-links released from ternary DNA-protein-protein conjugates. Of the identified proteins listed in Table S1, 126 (49.0%) are classified as nuclear proteins by the GO database available *via* the European Bioinformatics Institute (http://www.ebi.ac.uk/QuickGO) (Figure 4.4A). These include high mobility group (HMG) proteins, histone proteins, 40S ribosomal proteins, 60S ribosomal proteins, and elongation factors. This is not unexpected, considering those proteins are either localized in close proximity to DNA or are directly associated with DNA, increasing their chance of being covalently captured on DNA in the presence of cisplatin. An additional 46 proteins (17.9%) are classified as cytoplasmic, 46 (17.9%) as ribosomal proteins, and 7 (2.7%) as membrane-bound proteins (Figure 4.4A). Our observation of cross-linking of those proteins to DNA may be explained by additional roles that these proteins play in a cell.

DPC-forming proteins were further classified according to their GO annotations in regard to their molecular function and their participation in biological processes (Figures 4.4B and 4.4C). We found that a large portion of the identified proteins belong to the following three categories: DNA binding proteins (34, or 13.2%), RNA binding proteins (41, or 16.0%), and protein binding proteins (47, or 18.3%) (Table 4.1 and Figure 4.4B). By comparison, DNA and RNA-binding proteins comprise only 10.3% of the total nuclear proteome of human fibroblasts *(179)*. A total number of 52 identified proteins (20.2%) are known to play a role in RNA processing or splicing (Figure 4.4C), including arginine/serine-rich splicing factors, heterogeneous nuclear ribonucleoproteins, and ATP-dependent RNA helicases. An additional 10.5% of proteins are involved in transcriptional regulation (Figure 4.4C), including transcription activator BRG 1,

Figure 4.2 SDS-PAGE analysis of samples employed in the proteomics analysis of cisplatin-induced DPCs. HT1080 cells (~10^7) were treated 0 (lanes 2-4) or 100 μM cisplatin for 3 h (lanes 6-8). Following modified phenol/chloroform extraction of DNA in the presence of proteasome inhibitors and heating to release proteins, they were separated by 12% SDS-PAGE and visualized by staining with SimplyBlue SafeStatin. Proteins present in the 15 – 250 KDa molecular weight range were excised from the gel, subjected to in-gel tryptic digestion, and analyzed by HPLC-ESI^+-MS/MS.

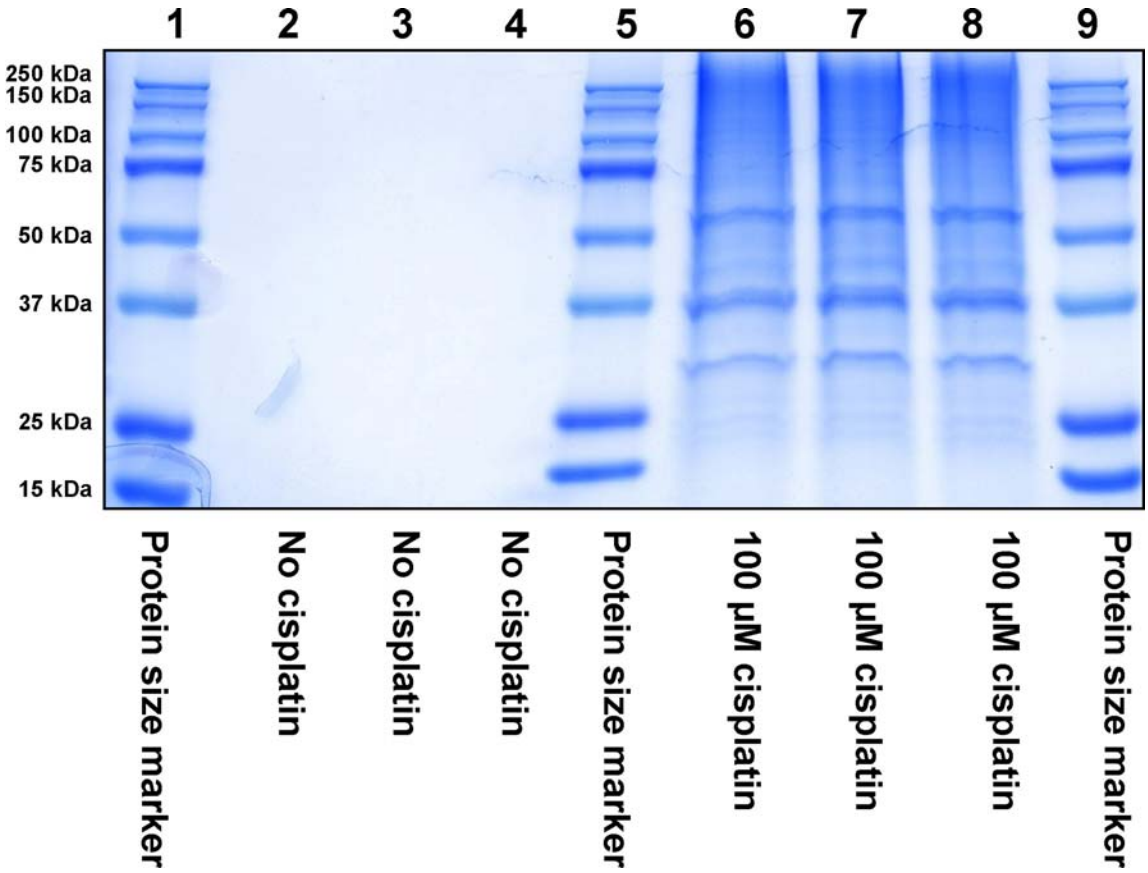


Table 4.1 Proteins that form cross-links to chromosomal DNA in human fibrosarcoma HT1080 cells treated with cisplatin (100 μM for 3 h).

Swiss-Port ID	*Identified Proteins*	*% Coverage*	*No. of Unique Peptides*	*No. of Assigned Spectra*	*Primary Cellular Function*	*Protein MW (Da)*	*Also Participate in DPC Formation by Mechlorethamine (177)*	*Also Form Cisplatin-induced DPCs in vitro (Nuclear Extracts) (see Chapter III)*
P62258	14-3-3 Protein epsilon	24	5	11	*Cell Signalling/ Motility/Architecture*	29175.0		
P62736	Actin, aortic smooth muscle	24	8	37		42010.1		
P60709	Actin, cytoplasmic 1 (β-actin)	35	7	97		41737.8		yes
Q01518	Adenylyl cyclase-associated protein 1	12	4	9		51855.5		
P12814	α-Actinin-1	13	10	19		103061.1		
O43707	α-Actinin-4	10	6	11		104857.2		
P04083	Annexin A1	37	9	22		38715.9		
P07355	Annexin A2	58	15	36		38606.1		yes
P23528	Cofilin-1	45	5	15		18503.2		
Q07065	Cytoskeleton-associated protein 4	9	4	7		66022.2		
Q16643	Drebrin	14	6	10		71428.6		
P14625	Endoplasmin	38	24	80		92471.7		
P15311	Ezrin	9	5	9		69414.7		
P21333	Filamin-A	3	4	6		280729.4		
Q99988	Growth/differentiation factor 15	23	6	12		34154.8		
P07900	Heat shock protein HSP 90-α	23	13	45		84663.2		yes
P08238	Heat shock protein HSP 90-β	21	12	75		83267.3		
O95373	Importin-7	7	5	15		119519.5		
P05787	Keratin, type II cytoskeletal 8	42	18	74		53706.2		
P02545	Prelamin-A/C	53	34	93		74140.7		
P20700	Lamin-B1	36	16	29		66409.6		
Q03252	Lamin-B2	22	13	46		67689.8		
Q15185	Prostaglandin E synthase 3	42	6	15		18697.9		
P61026	Ras-related protein Rab-10	28	5	9		22542.1		
Q13813	Spectrin α chain, brain	37	75	161		284542.7		
Q01082	Spectrin β chain, brain 1	27	46	91		274613.4		
Q9UH99	SUN domain-containing protein 2	10	4	9		80312.2		
P09493	Tropomyosin α-1 chain	14	4	4		32710.0		
Q71U36	Tubulin α-1A chain	41	14	53		50135.7		
P07437	Tubulin β chain	16	4	20		49670.6		

Q13885	Tubulin β-2A	36	11	91		49907.1		
P68371	Tubulin β-2C chain	20	6	18		49830.7		
Q9BUF5	Tubulin β-6 chain	18	4	6		49857.2		
P08670	Vimentin	65	28	990		53652.7	yes	
P31946	14-3-3 protein β/α	19	5	10	*Cellular Homeostasis/ Cell Cycle*	28083.1		
P62244	40S ribosomal protein S15a	29	4	9		14840.0		
P62847	40S ribosomal protein S24	29	4	8		15423.8		
P61247	40S ribosomal protein S3a	41	10	23		29945.3		
P62753	40S ribosomal protein S6	23	5	11		28681.7		
P08865	40S ribosomal protein SA	36	8	15		32854.1		
P10809	60 kDa heat shock protein, mitochondrial	23	8	19		61055.7		
P11021	78 kDa glucose-regulated protein OS=Homo	23	11	24		72334.7		
P84077	ADP-ribosylation factor 1	34	4	8		20697.6		
P18085	ADP-ribosylation factor 4	22	4	9		20511.6		
P15144	Aminopeptidase N	25	18	40		109542.4	yes	
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	3	4	7		151887.9		
P25705	ATP synthase subunit α, mitochondrial	14	7	10		59752.1		
P06576	ATP synthase subunit β, mitochondrial	10	4	7		56560.6		
O00571	ATP-dependent RNA helicase DDX3X	9	4	8		73245.8		
P27824	Calnexin	23	14	36		67570.2		
P27797	Calreticulin	52	18	54		48142.9		
Q9UQ88	Cell division protein kinase 11A	14	10	22		90976.0	yes	
Q68CQ4	Digestive organ expansion factor homolog	12	7	9		87057.4		
P62495	Eukaryotic peptide chain release factor subunit 1	13	4	7		49032.6		
Q99613	Eukaryotic translation initiation factor 3 subunit C	6	5	7		105347.2		
O60841	Eukaryotic translation initiation factor 5B	12	11	24		138831.5		
P02751	Fibronectin	8	10	21		262616.9		
P09382	Galectin-1	39	5	11		14715.8		
P14314	Glucosidase 2 subunit β	11	5	6		59425.8		
P04406	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	35	8	18		36053.4		yes
P62826	GTP-binding nuclear protein Ran	28	5	10		24423.1		yes
P11142	Heat shock cognate 71 kDa protein	13	5	16		70899.8		yes
Q1KMD3	Heterogeneous nuclear ribonucleoprotein U-like protein 2	20	13	24		85105.2		
Q9Y4L1	Hypoxia up-regulated protein 1	18	12	28		111336.8		
P05556	Integrin β-1	16	10	20		88415.1	yes	

P05783	Keratin, type I cytoskeletal 18	53	21	61		48059.0		
P35580	Myosin-10	8	11	25		229005.3		
P35579	Myosin-9	25	38	96		226537.5		
P07196	Neurofilament light polypeptide	53	25	69		61517.8	yes	
Q14978	Nucleolar and coiled-body phosphoprotein 1	8	5	10		73604.2	yes	
Q13823	Nucleolar GTP-binding protein 2	10	5	9		83656.0		
Q9Y2X3	Nucleolar protein 58	21	8	17		59580.2		
P19338	Nucleolin (C-23)	36	25	305		76615.9		
P06748	Nucleophosmin	47	10	52		32575.5	yes	
Q99733	Nucleosome assembly protein 1-like 4	21	6	11		42823.9		
P62937	Peptidyl-prolyl cis-trans isomerase A	28	5	9		18012.9		
Q13427	Peptidyl-prolyl cis-trans isomerase G	5	4	6		88619.0		
Q06830	Peroxiredoxin-1	59	9	18		22110.9		
P18669	Phosphoglycerate mutase 1	25	4	6		28804.8		
P13796	Plastin-2	15	6	10		70292.1		
O00622	Protein CYR61	21	8	15		42026.0		
P07237	Protein disulfide-isomerase	38	15	38		57118.1		
P13667	Protein disulfide-isomerase A4	13	9	17		72934.0		
Q58FF3	Putative endoplasmin-like protein	9	4	10		45859.7		
Q58FF8	Putative heat shock protein HSP 90-β-2	18	6	11		44350.2		
Q58FF7	Putative heat shock protein HSP 90-β-3	20	14	170		68326.5		
P14618	Pyruvate kinase isozymes M1/M2	13	5	12		57937.5		yes
P51149	Ras-related protein Rab-7a	33	6	10		23490.0		
Q14692	Ribosome biogenesis protein BMS1 homolog	8	6	8		145812.2		
Q14137	Ribosome biogenesis protein BOP1	23	12	20		83629.3		
Q9Y265	RuvB-like 1	16	6	10		50229.4		
P62136	Serine/threonine-protein phosphatase PP1-α catalytic subunit	25	6	9		37513.9		
Q9BXP5	Serrate RNA effector molecule homolog	16	11	16		100669.7		
Q9NQZ2	Something about silencing protein 10	19	10	35		54559.2		
P78371	T-complex protein 1 subunit β	13	4	7		57489.9		yes
P40227	T-complex protein 1 subunit ζ	17	4	7		58025.3		
P37802	Transgelin-2	23	4	7		22391.9		
P43307	Translocon-associated protein subunit α	20	4	9		32236.0		
Q9BV38	WD repeat-containing protein 18	16	4	4		47405.1		
Q86VM9	Zinc finger CCCH domain-containing protein 18	11	6	10		106379.8	yes	
P23396	40S ribosomal protein S3	35	7	14	*DNA Damage Response/DNA Repair*	26688.6		
Q9Y5B9	FACT complex subunit SPT16	40	41	113		119917.4		

Q08945	FACT complex subunit SSRP1	51	29	185		81077.6		
P09429	High mobility group protein B1 (HMG B1)	23	5	22		24894.7		yes
P26583	High mobility group protein B2 (HMG B2)	45	11	32		24034.6		
P17096	High mobility group protein HMG-I/HMG-Y	45	5	10		11676.2	yes	
P16403	Histone H1D	19	5	14		21365.8		
Q96QV6	Histone H2A type 1-A	40	5	18		14234.2		
Q96A08	Histone H2B type 1-A	28	4	9		14168.0		
P33778	Histone H2B type 1-B	36	8	31		13950.8		
P62805	Histone H4	54	6	79		11367.7		
P09874	Poly [ADP-ribose] polymerase 1 (PARP-1)	4	4	5		113087.8		yes
Q9NY61	Protein AATF	41	17	37		63135.0		
B2RPK0	Putative high mobility group protein B1-like 1	32	10	41		24238.8		
P23246	Splicing factor, proline- and glutamine-rich	16	7	11		76149.5		yes
Q9UIG0	Tyrosine-protein kinase BAZ1B	5	6	10		170907.0		
P13010	X-ray repair cross-complementing protein 5 (XRCC-5 or Ku80/86)	11	4	10		82707.1		yes
P12956	X-ray repair cross-complementing protein 6 (XRCC-6 or Ku70)	27	12	28		69846.4		yes
Q15029	116 kDa U5 small nuclear ribonucleoprotein component	13	9	12	*RNA Processing/ mRNA Splicing*	109438.1		
P62081	40S ribosomal protein S7	32	4	8		22127.5		
P62913	60S ribosomal protein L11	21	4	8		20253.2		
Q08211	ATP-dependent RNA helicase A	8	7	13		140961.5		
Q9NVP1	ATP-dependent RNA helicase DDX18	9	5	11		75409.7		
O00148	ATP-dependent RNA helicase DDX39	9	4	6		49129.8		yes
Q9H583	HEAT repeat-containing protein 1	3	5	8		242378.8		
Q32P51	Heterogeneous nuclear ribonucleoprotein A1-like 2	27	7	17		34225.4		
P51991	Heterogeneous nuclear ribonucleoprotein A3	25	8	21		39595.1		
O60812	Heterogeneous nuclear ribonucleoprotein C-like 1	29	13	279		32142.7		
P38159	Heterogeneous nuclear ribonucleoprotein G	22	9	19		42333.7		
P31943	Heterogeneous nuclear ribonucleoprotein H	21	6	15		49229.8		
P31942	Heterogeneous nuclear ribonucleoprotein H3	16	4	10		36927.6		
P61978	Heterogeneous nuclear ribonucleoprotein K	41	12	24		50978.5		yes
P14866	Heterogeneous nuclear ribonucleoprotein L	16	6	13		64132.8		yes
P52272	Heterogeneous nuclear ribonucleoprotein M	16	9	18		77517.3		yes
O60506	Heterogeneous nuclear ribonucleoprotein Q	15	9	15		69603.5		yes
Q00839	Heterogeneous nuclear ribonucleoprotein U	37	29	78		90585.2		
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	33	10	28		37430.3		yes

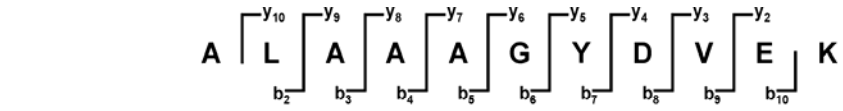
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	48	13	739		33607.5		
Q9UKD2	mRNA turnover protein 4 homolog	51	13	37		27561.3		
P78316	Nucleolar protein 14	7	5	7		97671.9	yes	
O00567	Nucleolar protein 56	24	10	18		66052.0		
Q9NR30	Nucleolar RNA helicase 2	14	9	13		87346.0		
P12270	Nucleoprotein TPR	16	30	52		267289.3	yes	
O00541	Pescadillo homolog	36	19	55		68004.9		
Q15365	Poly(rC)-binding protein 1	22	5	8		37498.2		yes
P26599	Polypyrimidine tract-binding protein 1	30	9	20		57222.5		yes
Q9HCG8	Pre-mRNA-splicing factor CWC22 homolog	4	4	11		105470.2		
Q92841	Probable ATP-dependent RNA helicase DDX17	9	5	8		72373.0		
P17844	Probable ATP-dependent RNA helicase DDX5	13	7	11		69149.7		
P46087	Putative ribosomal RNA methyltransferase NOP2	26	18	35		89303.6		
Q8IY81	Putative rRNA methyltransferase 3	34	19	40		96560.5		
O76021	Ribosomal L1 domain-containing protein 1	18	7	12		54974.7		
Q9NW13	RNA-binding protein 28	18	12	26		85739.1		
Q9UKM9	RNA-binding protein Raly	37	11	22		32463.9		
Q15287	RNA-binding protein with serine-rich domain 1	26	6	30		34209.6		
Q9Y3B9	RRP15-like protein	17	5	9		31484.5	yes	
Q8IYB3	Serine/arginine repetitive matrix protein 1	8	4	7		102337.5		
Q13435	Splicing factor 3B subunit 2	11	9	17		100229.4		
Q07955	Splicing factor, arginine/serine-rich 1	39	9	16		27745.1	yes	
O75494	Splicing factor, arginine/serine-rich 10	23	5	10		31301.7		
P84103	Splicing factor, arginine/serine-rich 3	29	5	11		19330.0		
Q08170	Splicing factor, arginine/serine-rich 4	14	7	15		56680.0		
Q13243	Splicing factor, arginine/serine-rich 5	19	5	12		31264.8		
Q16629	Splicing factor, arginine/serine-rich 7	15	4	6		27367.5	yes	
P62995	Transformer-2 protein homolog beta	17	4	10		33666.7	yes	
P09661	U2 small nuclear ribonucleoprotein A'	28	5	7		28417.1		
O00566	U3 small nucleolar ribonucleoprotein protein MPP10	33	15	38		78866.8		
O75643	U5 small nuclear ribonucleoprotein 200 kDa helicase	4	6	11		244513.8		
Q96MU7	YTH domain-containing protein 1	9	7	12		84700.6		
P62280	40S ribosomal protein S11	42	8	18	*Transcriptional Regulation/Translation*	18431.3		
P62277	40S ribosomal protein S13	30	4	7		17223.3		
P62263	40S ribosomal protein S14	36	4	9		16272.9		
P62249	40S ribosomal protein S16	45	7	14		16445.9		
P08708	40S ribosomal protein S17	55	5	14		15550.5		

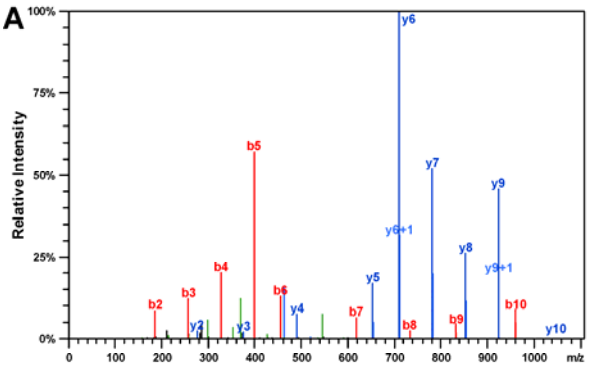
P62269	40S ribosomal protein S18	36	6	12		17719.3		
P15880	40S ribosomal protein S2	23	6	10		31325.2		
P62266	40S ribosomal protein S23	34	7	14		15807.7		
P62701	40S ribosomal protein S4	21	5	8		29599.3		
P46782	40S ribosomal protein S5	35	7	11		22877.0		
P62241	40S ribosomal protein S8	42	7	11		24206.4		
P46781	40S ribosomal protein S9	53	13	29		22592.5		
Q8NHW5	60S acidic ribosomal protein P0-like	18	5	9		34365.1		
Q96L21	60S ribosomal protein L10-like	21	5	14		24519.2		
P30050	60S ribosomal protein L12	49	6	10		17819.1		
P26373	60S ribosomal protein L13	33	7	11		24262.2		
P40429	60S ribosomal protein L13a	33	12	26		23577.9		
P50914	60S ribosomal protein L14	37	8	22		23432.3		
P61313	60S ribosomal protein L15	39	8	18		24146.5		
P18621	60S ribosomal protein L17	31	5	11		21397.4		
Q07020	60S ribosomal protein L18	35	6	19		21635.2		
Q02543	60S ribosomal protein L18a	52	10	23		20762.6		
P84098	60S ribosomal protein L19	32	7	24		23467.4		
P46778	60S ribosomal protein L21	34	5	12		18565.0		
P62829	60S ribosomal protein L23	55	7	21		14865.9		
P62750	60S ribosomal protein L23a	33	6	14		17696.2		
P83731	60S ribosomal protein L24	41	7	15		17779.5		
P61353	60S ribosomal protein L27	36	4	9		15798.4		
P46776	60S ribosomal protein L27a	34	5	10		16561.4		
P46779	60S ribosomal protein L28	49	8	16		15747.9		
P39023	60S ribosomal protein L3	24	9	18		46109.5		
P62888	60S ribosomal protein L30	51	4	9		12784.7		
P62910	60S ribosomal protein L32	35	4	14		15860.4		
P49207	60S ribosomal protein L34	29	5	9		13293.1		
Q9Y3U8	60S ribosomal protein L36	35	5	12		12254.2		
P61927	60S ribosomal protein L37	20	4	5		11078.2		
P61513	60S ribosomal protein L37a	59	5	12		10275.4		
P36578	60S ribosomal protein L4	36	13	28		47699.1		
P46777	60S ribosomal protein L5	37	9	20		34363.5		
Q02878	60S ribosomal protein L6	34	10	26		32729.3		
P18124	60S ribosomal protein L7	40	10	22		29227.7		
P62424	60S ribosomal protein L7a	40	9	22		29996.3		
P62917	60S ribosomal protein L8	22	6	11		28024.8		
P32969	60S ribosomal protein L9	55	6	17		21863.7		

P11387	DNA topoisomerase 1	6	4	6		90729.7		
Q03701	CCAAT/enhancer-binding protein ζ	11	8	14		120992.2		
Q13185	Chromobox protein homolog 3	17	4	7		20812.0		
O75367	Core histone macro-H2A.1	28	7	14		39618.9		
O95602	DNA-directed RNA polymerase I subunit RPA1	4	4	6		194814.5		
P68104	Elongation factor 1-α 1 (EF-1α1)	27	9	25		50141.2		yes
P13639	Elongation factor 2 (EF-2)	17	10	24		95340.1		yes
P60842	Eukaryotic initiation factor 4A-I	21	7	12		46155.3		yes
P05198	Eukaryotic translation initiation factor 2 subunit 1	27	8	15		36112.7		
P63241	Eukaryotic translation initiation factor 5A-1	39	4	8		16832.7		
Q5SSJ5	Heterochromatin protein 1-binding protein 3	16	7	11		61208.7		
Q99729	Heterogeneous nuclear ribonucleoprotein A/B	11	4	7		36225.2		yes
O14979	Heterogeneous nuclear ribonucleoprotein D-like	15	4	8		46438.6		
Q14103	Heterogeneous nuclear ribonucleoprotein D0	23	8	15		38434.5		yes
P52926	High mobility group protein HMGA2	58	5	17		11831.9	yes	
Q12905	Interleukin enhancer-binding factor 2	62	16	80		43062.7		yes
Q12906	Interleukin enhancer-binding factor 3	15	12	18		95338.9		yes
P43243	Matrin-3	50	30	107		94626.7	yes	
Q9BQG0	Myb-binding protein 1A	28	31	80		148858.3		
Q13765	Nascent polypeptide-associated complex subunit α	26	4	7		23383.3		
P17480	Nucleolar transcription factor 1	20	14	33		89409.6	yes	
P55209	Nucleosome assembly protein 1-like 1	36	8	24		45375.0		
Q9H307	Pinin	37	22	56		81613.2		
P51531	Probable global transcription activator SNF2L2	9	10	22		181283.0		
Q8IZL8	Proline-, glutamic acid- and leucine-rich protein 1	15	11	22		119700.6		
P35659	Protein DEK	22	8	15		42675.9	yes	
Q8N7H5	RNA polymerase II-associated factor 1 homolog	15	6	12		59976.0		
Q15424	Scaffold attachment factor B1	21	14	30		102642.8		
Q92922	SWI/SNF complex subunit SMARCC1	22	18	42		122867.4		
Q969G3	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily E member 1	30	11	29		46649.7	yes	
Q9Y2W1	Thyroid hormone receptor-associated protein 3	13	12	33		108668.9		
P51532	Transcription activator BRG1	11	14	28		184649.6		
Q5BKZ1	Zinc finger protein 326	33	14	34		65653.5	yes	
Q9NVI7	ATPase family AAA domain-containing protein 3A	7	4	6	*unknown*	71370.2		
P55081	Microfibrillar-associated protein 1	17	5	10		51958.7		
Q9Y3T9	Nucleolar complex protein 2 homolog	30	19	70		84907.1		
O94880	PHD finger protein 14	6	4	7		100055.1		

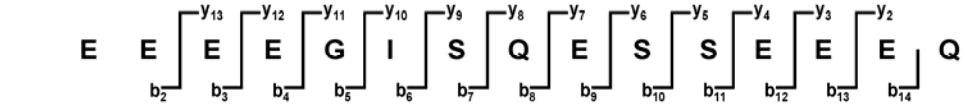
Q96GQ7	Probable ATP-dependent RNA helicase DDX27	17	12	20		89838.1		
Q9H0S4	Probable ATP-dependent RNA helicase DDX47	11	4	9		50648.4		
Q9Y4W2	Protein LAS1 homolog	22	13	28		83065.5		
Q9BXY0	Protein MAK16 homolog	50	12	30		35370.0		
Q5JTH9	RRP12-like protein	5	4	6		143705.1		
Q15061	WD repeat-containing protein 43	37	20	106		74890.8		

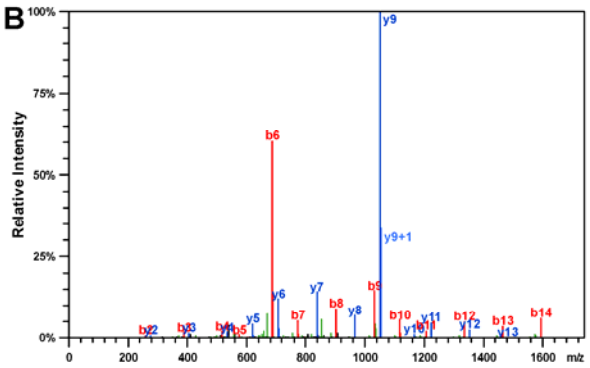
Figure 4.3 Representative HPLC-ESI$^+$-MS/MS spectra of tryptic peptides used in the identification of histone H1D (A), HMG B1 (B), and XRCC-6 protein (C).



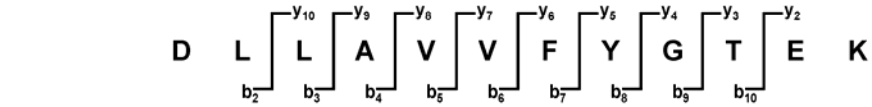


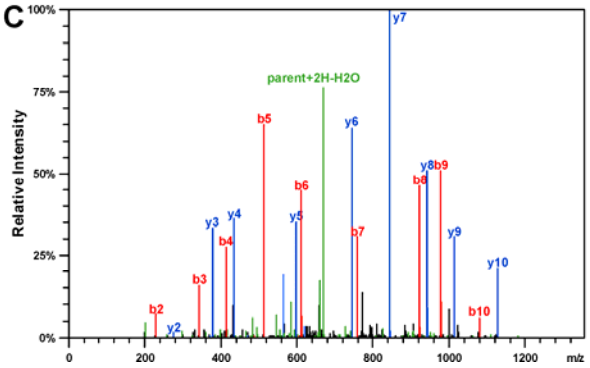
Sequence	#	B Ions	Y Ions	+1
A	1	72.05	1107.57	11
L	2	185.13	1036.53	10
A	3	256.17	923.45	9
A	4	327.2	852.41	8
A	5	398.24	781.37	7
G	6	455.26	710.34	6
Y	7	618.32	653.31	5
D	8	733.35	490.25	4
V	9	832.42	375.22	3
E	10	961.46	276.16	2
K	11	1107.57	147.11	1





Sequence	#	B Ions	Y Ions	+1
E	1	130.12	1738.68	15
E	2	259.09	1609.63	14
E	3	388.14	1480.59	13
E	4	517.18	1351.55	12
G	5	574.2	1222.51	11
I	6	687.28	1165.49	10
S	7	774.32	1052.4	9
Q	8	902.37	965.37	8
E	9	1031.42	837.31	7
S	10	1118.45	708.27	6
S	11	1205.48	621.24	5
E	12	1334.52	534.2	4
E	13	1463.57	405.16	3
E	14	1592.61	276.12	2
Q	15	1738.68	147.08	1





Sequence	#	B Ions	Y Ions	+1
D	1	130.12	1354.73	12
L	2	229.12	1239.70	11
L	3	342.20	1126.61	10
A	4	413.24	1013.53	9
V	5	512.31	942.49	8
V	6	611.38	843.42	7
F	7	758.44	744.36	6
Y	8	921.51	597.29	5
G	9	978.53	434.22	4
T	10	1079.58	377.20	3
E	11	1208.62	276.16	2
K	12	1354.73	147.11	1

matrin-3, and interleukin enhancer-binding factors. However, since DNA isolated by our modified phenol/chloroform method has minimal RNA contamination (see representative chromatogram in Figure 4.5 where no RNA ribonucleosides were detected in enzymatic digests) *(177)*, we hypothesize these proteins possess additional DNA-binding capabilities which can be triggered by DNA-damaging agents such as cisplatin.

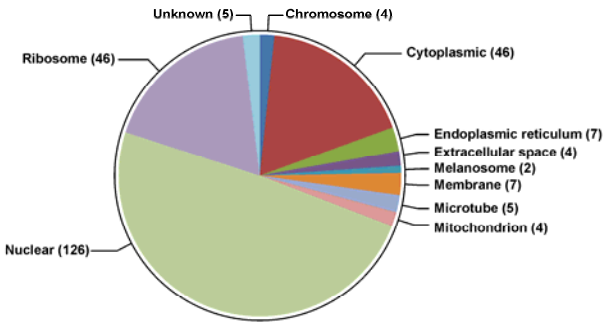
4.3.3 Western Blot Analysis of Cross-Linked Proteins

To confirm the results of the mass spectrometry based proteomics and to discover additional proteins participating in DPC formation, proteins co-purified with chromosomal DNA isolated from cisplatin-treated cells were subjected to western blot analysis. Commercial antibodies against EF-1α1, PARP, Ref-1, nucleolin, actin, GAPDH, Fen-1, AGT, and XRCC-1 were employed. These proteins were selected because they were either among the gene products identified by mass spectrometry analyses (EF-1α1, PARP, GAPDH, nucleolin, and actin, Table 4.1) or have been previously found to form cisplatin-induced DPC in our *in vitro* studies employing cell free protein extracts (Ref-1, AGT, Fen-1, and XRCC-1) (*see Chapter III*) *(35;36)*. Aliquots (30 μg) of DNA isolated from cisplatin-treated HT1080 cells (10, 50, 100, 250, or 500 μM) were taken and heated with SDS-containing gel loading buffer (1 h at 70°C) to release the proteins *via* platination migration (Scheme 4.2). The resulting proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes for western blot analysis using specific antibodies mentioned above.

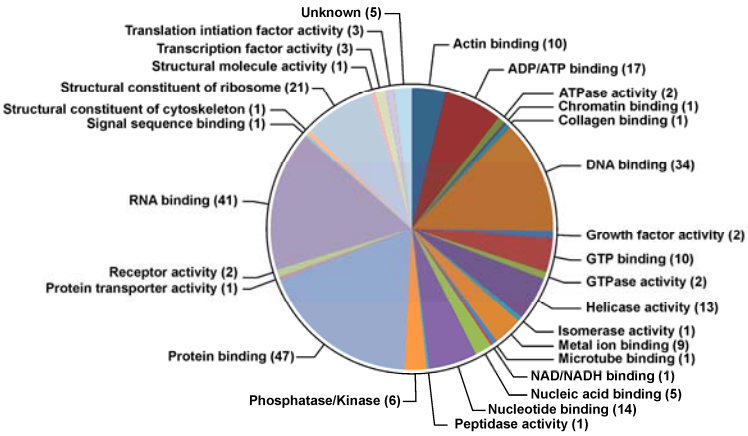
Western blotting experiments confirmed the identities of five gene products identified from mass spectrometry based proteomics: EF-1α1, PARP, nucleolin, GAPDH, and actin (Figure 4.6A). In addition, a concentration-dependent DPC formation involving four additional proteins: Ref-1, AGT, Fen-1 and XRCC-1, was observed. Among these, nucleolin displayed the greatest cross-linking efficiency, with approximately 10 % of total proteins trapped on DNA following

Figure 4.4 GO annotations for proteins involved in cisplatin-induced DPC formation in human HT1080 cells: cellular distribution (A), molecular function (B), and biological process (C). The numbers of proteins in each category is labeled on the charts.

A. Cellular distrubution



B. Molecular function



C. Biological process

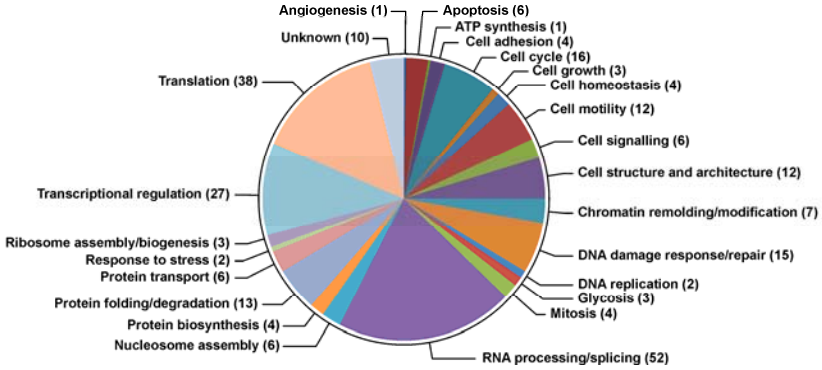
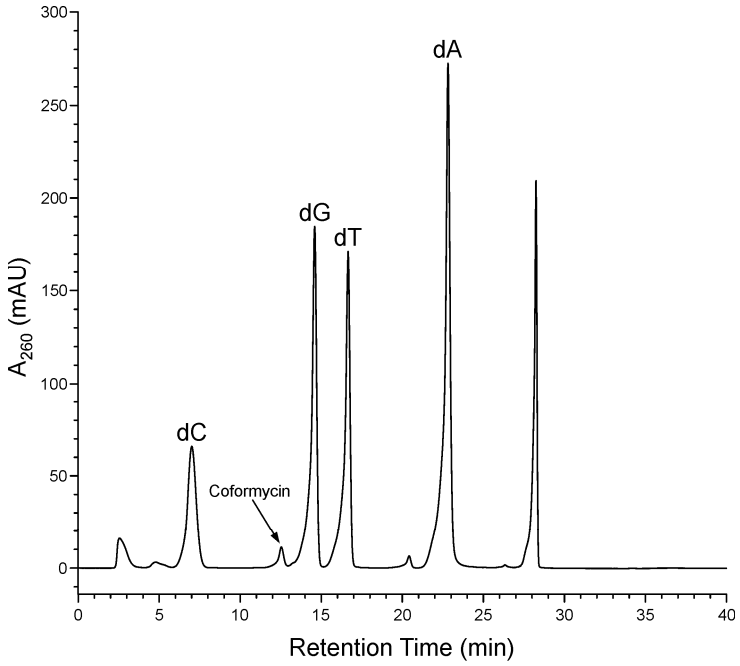


Figure 4.5 HPLC separation of nucleoside mixtures resulting from enzymatic digestion of DNA that was isolated from cisplatin-treated HT1080 cells. After harvesting cell nuclei, DNA containing covalently trapped proteins was isolated by a phenol/chloroform extraction protocol optimized for the extraction of DPCs. The identities of nucleoside was confirmed by HPLC analysis of authentic standards. Coformycin was added to inhibit the contaminating deaminase activity present in commercial alkaline phosphatase. Accurate DNA quantification was accomplished by comparing dG peak areas in all samples to a calibration curve obtaining by injecting known amounts of dG standard solutions.



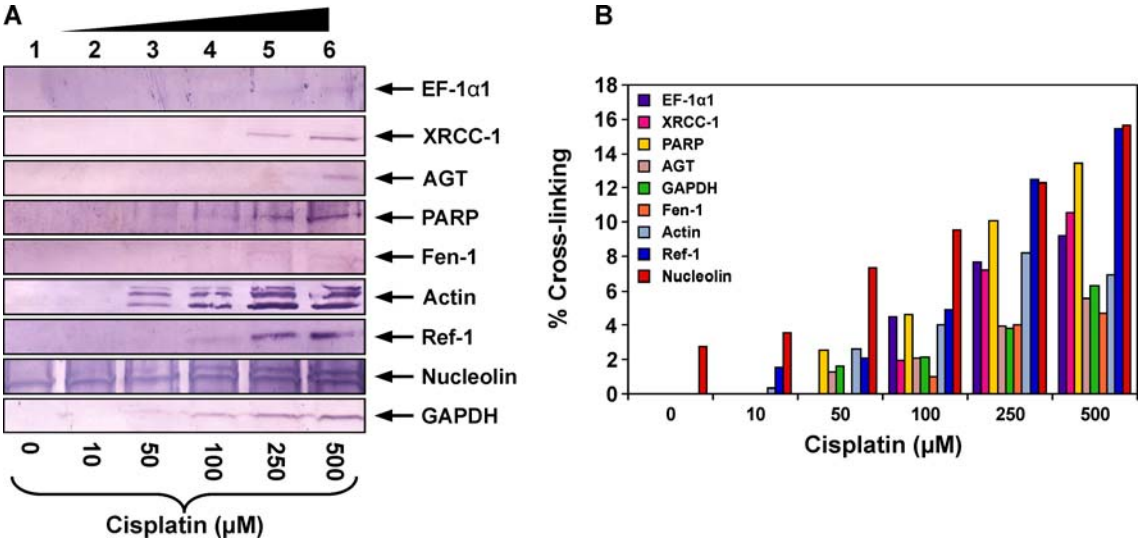
treatment with 100 μM cisplatin (Figure 4.6B). We hypothesize that our ability to detect DPCs involving AGT, Fen-1, Ref-1, and XRCC-1 protein by western blot analyses but not by mass spectrometry is a result of a greater sensitivity of the former methodology.

4.3.4 HPLC-ESI$^+$-MS/MS Analysis of dG-Pt-Lys Conjugates as Evidence for DPC Formation

To confirm the formation of covalent DNA-protein conjuates in cisplatin-treated cells, HT1080 cells (~10^6) were treated with 0 or 100 μM cisplatin, and the chromosomal DNA was extracted as described above. Equal DNA amounts (50 μg) were taken from each sample and subjected to enzymatic hydrolysis to yield protein-nucleoside conjugates, followed by enzymatic digestestion to amino acids in the presence of trypsin and proteinase K. Since our previous studies with recombinant AGT protein and *in vitro* cross-linking revealed that cisplatin is able to cross-link N7 of dG and amino group of lysine (*see Chapter II*), we have focused on lysine-guanine conjugates. Enzymatic digests were subjected to offline HPLC purification and HPLC-ESI$^+$-MS/MS analysis of dG-Pt-Lys.

Representative extracted ion chromatograms for HPLC-ESI$^+$-MS/MS analysis of dG-Pt-Lys in samples from cisplatin-treated and control HT1080 cells are shown in Figure 4.7. dG-Pt-Lys was detected in DNA samples from cisplatin-treated cells (Figure 4.7C), but not from untreated cells (Figure 4.7B), consistent with Lys-Gua conjugate formation as a result of cisplatin treatment. These data indicate that cisplatin-induced DNA-protein cross-linking takes place between the N7 position of guanine in DNA and the side-chain ε-amino group of lysine in proteins. Based on our previous studies with recombinant proteins (*see Chapter II*), we hypothesize that other nucleophilic amino acid, such as cysteine, histidine, glutamic acid, and arginine, also participates in cisplatin-induced DNA-protein cross-linking (*see Chapter II*),

Figure 4.6 Western blot analysis of cisplatin-induced DPCs in HT1080 cells. Following treatment with 0 (lane 1), 10 (lane 2), 50 (lane 3), 100 (lane 4), 250 (lane 5), or 500 μM cisplatin (lane 6), DNA and cross-linked proteins were isolated by phenol/chloroform extraction. Samples were normalized for DNA content, and proteins from 30 μg DNA were released by heating and separated by SDS-PAGE, followed by transfer to nitrocellulose membranes. Western blotting was performed using primary antibodies specific for EF-1α1, AGT, Fen-1, nucleolin, actin, GAPDH, PARP, Ref-1, andXRCC-1 (A). The efficiency of DPC formation in the presence of cisplatin was estimated by densitometric analysis of proteins associated with DNA following cisplatin treatment and those present in whole cell protein lysates (B).



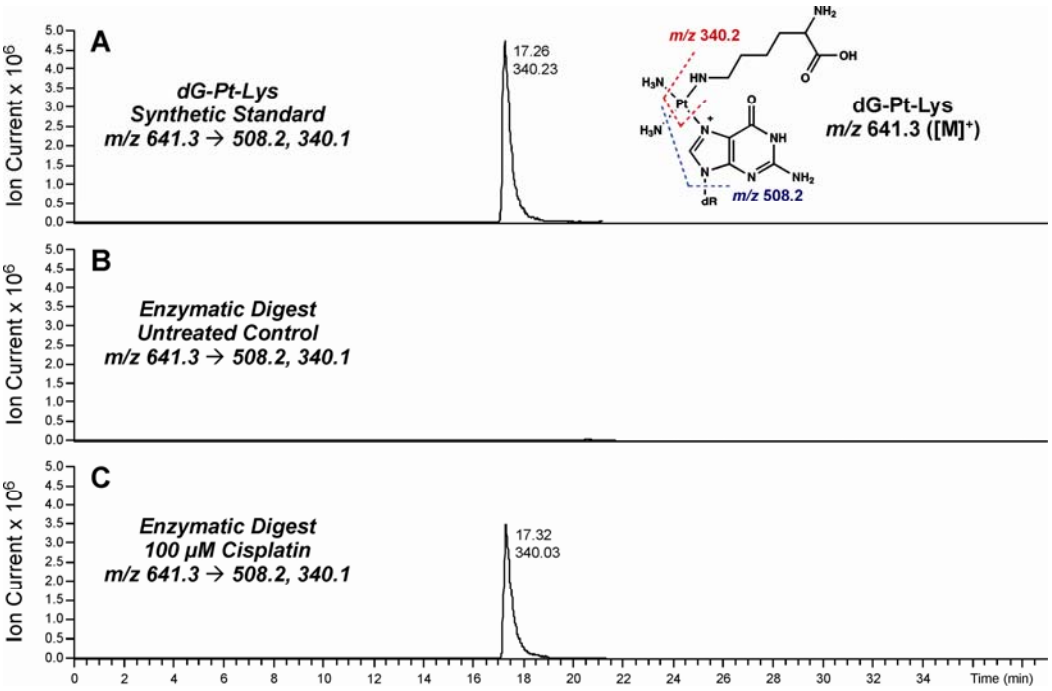
although they were not analyzed in the present study due to the lack of the corresponding standards.

4.4 Discussion

DNA-DNA cross-linking appears to play a central role in toxicity of cisplatin and its analogues *(111)*. Furthermore, the toxicity of the corresponding DNA-DNA lesions is mediated by DNA-protein interactions *(172)*. For example, Lippard *et al* have shown that binding of HMG group proteins to cross-linked adducts prevents their repair by the NER mechanism and sequesters transcriptional factors *(109;180)*. This raises a possibility that proteins can bind to cisplatin-DNA monoadducts and become covalently trapped on DNA strands. Indeed, platinum compounds have been shown to induce DNA-protein lesions *(32;71-74;174)*. However, little is known the free range of proteins targeted for cross-linking to DNA in the presence of cisplatin. Our previous studies with recombinant AGT protein and nuclear protein extracts revealed that cisplatin is an extremely efficient DNA-protein cross-linking agent *in vitro* (*see Chapters II and III*). However, the DPC formation by platinum compounds in human cell culture has not been studied.

In order to study DPC formation by cisplatin in cells, a methodology was required to selectively extract chromosomal DNA containing platinated DPC lesions. Our laboratory recently has developed a simple and effective approach for DPC isolation from cells using a modified phenol/chloroform extraction *(177)*. In the case of thermally labile adducts such as those derived from nitrogen mustards, neutral thermal hydrolysis can be used to selectively release DPC from the DNA backbone in the form of protein-guanine conjugates *(177)*. In contrast, platinum-N7 guanine does not undergo spontaneous depurination *(181)*. However, cisplatin cross-linked proteins can be partially released *via* platination migration from amino acids of the protein to

Figure 4.7 HPLC-ESI$^+$-MS/MS analysis of dG-Pt-Lys conjugates in total proteolytic digests of cisplatin-induced DPCs. HT1080 cells were treated with cisplatin to induce DNA-protein cross-links. Following extraction of the chromosomal DNA containing covalent DPCs, the cross-linked proteins were subjected to enzymatic hydrolysis to release amino acid-nucleobase conjugates. Shown are the HPLC-MS/MS traces of synthetic dG-Pt-Lys (A); enzymatic digests of DPC mixtures from HT1080 cells incubated in the absence of cisplatin (negative control) (B); enzymatic digests of DPC mixtures treated with 100 μM cisplatin (C).



nucleobases in DNA (Scheme 4.2 and Chapter II).

Following isolation of chromosomal DNA from cisplatin-treated cells, it was subjected to heating for 1 h at 70 ^{o}C. We have previously shown that under these conditions, proteins bound to DNA *via* platination are partially displaced by proximal nucleobases on genomic DNA, releasing the intact proteins (Scheme 4.2 and Chapter II). We combined this methodology with mass spectrometry-based proteomics and immunoblotting to examine DNA-protein cross-linking in human fibrosarcoma HT1080 cells treated with a toxic concentration of cisplatin (100 μM). Nearly 256 proteins were found to form cross-links to chromosomal DNA in the presence of cisplatin (Table 4.1). Close inspection of Table 4.1 reveals a broad range of proteins with a variety of cellular functions, including DNA damage response and repair (e.g. HMG proteins, histone proteins, PARP-1, and XRCC proteins), transcriptional regulation (e.g. 40S ribosomal proteins, 60S ribosomal proteins, EF-1α1, and matrin-3), RNA processing (e.g. ATP-dependent RNA helicase, heterogeneous nuclear ribonucleoproteins, poly(rC)-binding protein 1, and putative rRNA methyltransferase 3), cell signaling and architecture (e.g. actin, keratin, lamine, and vimentin), and regulation of cell cycle (e.g. GAPDH, nucleolin, nucleophosmin, and T-complex proteins). The majority of the identified proteins are known DNA-binding proteins (e.g. HMG proteins, histone proteins, PARP-1, and XRCC proteins), RNA-binding proteins (e.g. heterogeneous nuclear ribonucleoproteins, 40S ribosomal proteins, 60S ribosomal proteins, and arginine/serine-rich splicing factors), and protein-binding proteins (e.g. keratin, lamin, vimentin, and galectin-1) which are present in the nucleus (Figure 4.4C).

Previous studies employed affinity methodology to show that several DNA-binding proteins, including HMG proteins 1, 2, and E, cytokeratins, and histone proteins, can become cross-linked to DNA in the presence of cisplatin *(73)*. All of these have been detected in our proteomics screen of cisplatin-induced DPCs (Table 4.1). In addition, our studies established the identities of many additional nuclear proteins that participate in DNA-protein cross-linking

formation in the presence of cisplatin and determined atomic connectivity of the resulting macromolecular conjugates.

Approximately 20 proteins identified in the present work, including vimentin, histone H1D, HMG-A2, arginine/serine rich splicing factor, matrin-3, and zinc finger protein 326, have been previously found to form DPCs in HT1080 cells treated with mechlorethamine (50 μM). As was the case of mechlorethamine-induced cross-linking, we also observed a large number of RNA-binding proteins (Table 4.1). Our observation of less than 10% overlap between the two protein lists suggests that different cross-linking mechanisms are involved in cisplatin-induced platination and alkylation in the presence of nitrogen mustards. Furthermore, although 2-fold higher concentration (100 μM) of cisplatin was employed, ~ 6 fold (256) more proteins were identified to cross-link DNA compared with mechlorethamine, indicating that cisplatin is a more effective DPC-inducing agent than mechlorethamine.

Twenty eight proteins identified in the present study (Table 4.1), including actin, annexin A2, GAPDH, T-complex protein, HMG-B1, PARP-1, XRCC-5 and 6, and EF-1α1, were among the proteins that formed DPCs *in vitro* when cell-free nuclear protein extracts from human cervical carcinoma (HeLa) cells were incubated with biotinylated DNA duplexes in the presence of cisplatin (*see Chapter III*). The limited overlap (11%) between our cell culture and *in vitro* results may be explained by differences in relative protein abundance in Hela *(157)* and HT1080 cells *(178)* and by differences in the experimental design. Our *in vitro* studies utilized a biotinylated short synthetic oligonucleotide duplex derived from a region of the *K-ras* gene (*see Chapter III*). Since such oligonucleotide is unlikely to participate in sequence-specific interactions with DNA-binding proteins due to the lack of chromosomal structure, it is not surprising that lower efficiency of DPC formation was observed (*see Chapter III*). In contrast, the present study employed intact human cells where DNA is organized in chromatin and involves specific DNA-protein interactions, potentially facilitating DPC formation.

While the contributions of DNA-protein cross-linking to the biological activity of cisplatin remains to be established, these bulky lesions are expected to block DNA replication, transcription, and repair. Three important DNA repair pathways have been reported to contribute to the repair of cisplatin-induced DNA-DNA cross-links (*e.g.* 1,2-intrastrand cross-links), including nucleotide excision repair (NER) *(175;182;183)*, mismatch repair (MMR) *(184;185)*, and DNA-dependent protein kinase (DNA-PK) pathway *(186)*. However, repair pathways responsible for removal of cisplatin-induced DPCs remain undefined. It has been proposed that DPCs formed as a consequence of cellular exposure to formaldehyde and other *bis*-electrophiles can be repaired by NER *(45)*, homologous recombination (HR) *(167)*, and proteolytic degradation *(9)*. It has been proposed that proteolytic degradation breaks the proteins component of DPC to short peptides, followed by NER removal of the resulting DNA-peptide lesions *(1)*. NER can also directly excise DPCs involving relatively low molecular weight protein (<14 KDa). Alternatively, homologous recombination repair (HR) can directly remove cross-links involving larger proteins (>14 KDa). Indeed, a recent study by Chvalova and colleagues *(110)* observed the failure of human NER system to remove proteins cross-linked to DNA by cisplatin, suggesting that the majority of DPCs induced by cisplatin are expected to be repaired by HR, due to their significant size. Additionally the exact cross-link structure may affect their repair, and the repair mechanism may require more than one repair pathway *(1;168)*. Further studies involving site-specific cisplatin lesions are needed to determine the mechanisms of their repair and their effects on DNA replication.

In conclusion, the present study demonstrates that clinically relevant doses of cisplatin (0.1 ~ 500 μM) *(187)* can induce DNA-protein cross-links involving a large number of cellular proteins (256) in human fibrosarcoma cells. Proteins were identified by a combination of mass spectrometry-based proteomics and immunological detection. Many of the identified proteins are involved in a variety of cellular processes such as chromatin remodeling, translation, DNA

replication, DNA damage response, DNA repair, RNA processing, and transcriptional regulation. If not efficiently repaired, these bulky DPC lesions are expected to cause chromosomal double-strand breaks or be misread by DNA polymerases to induce mutations, ultimately triggering programmed cell death or genotoxic outcomes. Studies involving site-specifically damaged DNA are currently underway in our laboratory to obtain more details on the consequences of DPCs induced by antitumor platinum agents in human cells.

V. Formation of 8-oxo-dG and oxazolone lesions with *p53* derived DNA sequences following photooxidation in the presence of riboflavin

5.1 Introduction

Reactive oxygen species (ROS) and reactive nitrogen species (RNS), *e.g.* hydrogen peroxide, hydroxyl radical, superoxide, peroxynitrite, and singlet oxygen, are physiologically produced in tissues as a result of normal aerobic metabolism, immune response, and inflammation *(188-190)*. ROS and RNS play a dual role in a cell: while required for certain cellular processes such as signal transduction and protection against pathogens, they can seriously damage cells when overproduced, leading to oxidative stress, degradation of cellular biomolecules, toxicity, and cancer *(189)*.

Oxidative degradation of DNA has received a lot of attention due to its potential role in aging, cancer, and some neurodegenerative diseases *(191-198)*. Guanine bases within DNA are specifically targeted by oxidants due to their low ionization potential *(80)*. Oxidation of guanine gives rise to a complex mixture of products, including 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxo-dG), spiroiminodihydantoin, guanidinohydantoin, 2,6-diamino-4-hydroxyl-5-formamidopyrimidine (FapyG), 2-amino-5-[2-deoxy-*β*-D-*erythro*-pentofuranosyl)amino]-4*H*-imidazol-4-one (imidazolone) and its hydrolysis product, 2,2-diamino-4-[(2-deoxy-*β*-D-*erythro*-pentofuranosyl)amino]-5(2*H*)-oxazolone (oxazolone) (Figure 1.2) *(199-205)*. These oxidative lesions are formed by a common mechanism involving one-electron oxidation of G to form a guanine radical cation (Scheme 1.4). The resulting electron hole can migrate along the π-stack of base pairs within the DNA duplex until it reaches a site with low ionization potential (*e.g.* GG and GGG repeats) *(86;206-211)*, where the radical cation irreversibly reacts with water or molecular oxygen to form stable adducts (Scheme 1.4).

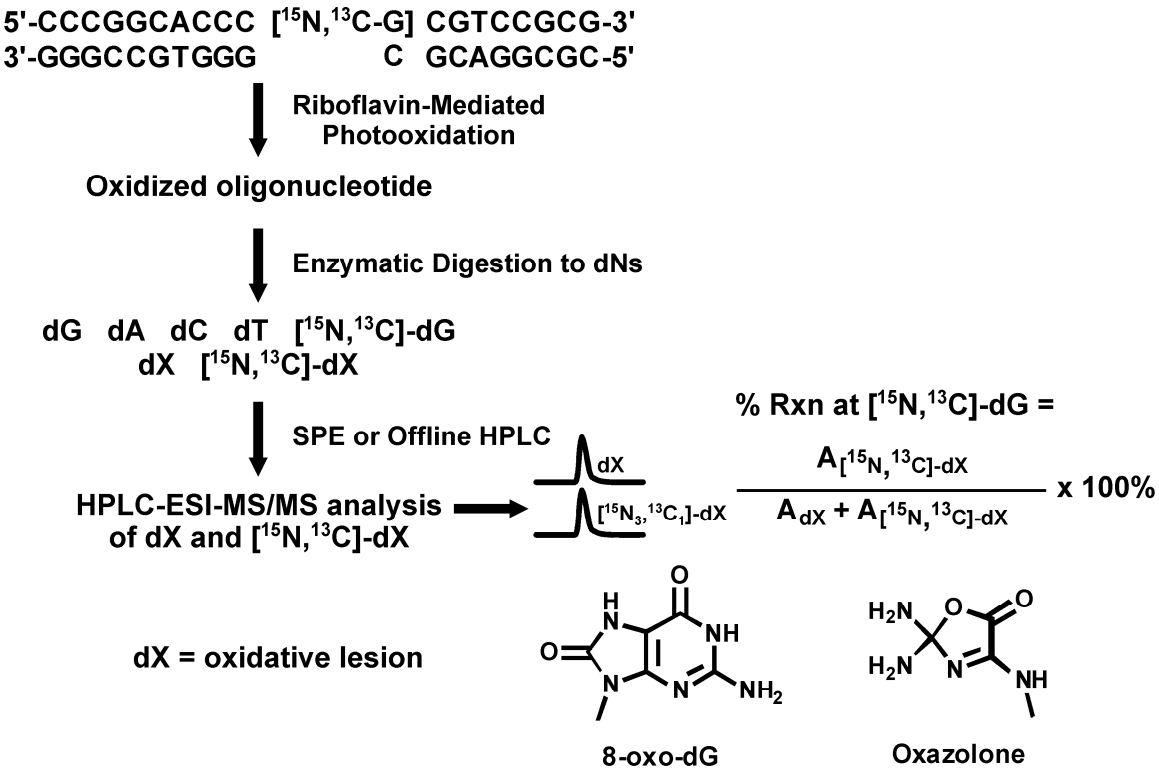
One type of DNA oxidation relevant to biological systems is photooxidation mediated by endogenous porphyrin and flavin-containing chromophores. For example, riboflavin (vitamin B2) binds to DNA *via* intercalation of the flavin moiety between G:C base pairs *(212)*. Exposure of DNA to UV irradiation in the presence of riboflavin enables nucleic bases to successively absorb two photons, whose cumulative energy exceeds the ionization threshold, thus giving rise to the ejection of an electron and consequently to the formation of the radical cations *(213;214)*. The latter undergoes further reactions to form stable adducts such as 8-oxo-dG and oxazolone (Figure 1.2, Scheme 1.4). Alternatively, excited riboflavin can transfer energy to triplet oxygen, leading to the formation of singlet oxygen *(213;214)*. Since riboflavin is ubiquitous in the human body, exposure to sunlight and/or artificial UV sources may greatly increase the opportunities of forming DNA oxidative adducts.

The goal of the present investigation was to analyze the distribution of specific oxidative adducts along DNA duplexes derived from frequently mutated regions of the *p53* gene following photooxidation in the presence of riboflavin. Stable isotope labeling of DNA-HPLC-ESI$^+$-MS/MS approach (ILD-MS), recently developed by our laboratory *(215)*, was employed to directly quantify 8-oxo-dG and oxazolone formation at specific nucleobases within DNA sequence (Scheme 5.1). Unlike previous studies that have relied on gel electrophoresis analysis of lesion-induced DNA strand breaks *(216)*, our approach, for the first time, analyzes sequence distribution of specific oxidative DNA lesions.

5.2 Materials and Methods

Chemicals **-** Ammonium acetate, ammonium formate, sodium cacodylic acid, sodium chloride, desferrioxamine, acetonitrile, zinc chloride, Tris base, methanol, sodium bicarbonate, nuclease P1, alkaline phosphatase, and riboflavin were purchased from Sigma-Aldrich (Milwaukee, WI).

Scheme 5.1 Strategy for quantitation of oxidative guanine lesions originating from specific sites within DNA sequence.



Triethylamine and hydrochloric acid were obtained from Fisher Scientific (Hanover Park, IL). [1,7,NH_2-$^{15}N_3$-2-^{13}C]-dG and [1,7,NH_2-$^{15}N_3$-2-^{13}C]-dG phosphoramidites were prepared as described previously *(217)*. Oxazolone, 8-oxo-dG, and their $^{15}N_3$,$^{13}C_1$-labelled analogues were prepared from dG and [1,7,NH_2-$^{15}N_3$-2-^{13}C]-dG as described elsewhere *(218)*.

DNA Oligodeoxynucleotides - DNA duplexes containing *p53* codons 153-159, 243-250, and 269-275 were selected for this study (Table 5.1). These contain the dominant lung cancers mutational "hot spots" at codons 157, 158, 245, 248, 249, and 273 (Figure 1.3). For each sequence of interest, a series of DNA strands were synthesized containing the [1,7,NH_2-$^{15}N_3$-2-^{13}C] stable isotope tag at a different guanine base (Table 5.1). ^{Me}C was incorporated in each strand of DNA at all physiologically methylated CG sites *(219)*. DNA oligodeoxynucleotides were synthesized by standard phosphoramidite chemistry using an Applied Biosystems ABI 394 instrument (Foster City, CA). $^{15}N_3$,$^{13}C_1$-dG was introduced by using [1,7,NH_2-^{15}N-2-^{13}C]-dG phosphoramidite prepared as described previously *(217)*. DNA oligomers were purified by HPLC and structurally characterized by UV spectroscopy and mass spectrometry as described previously *(96)*. The identity and purity of each strand was confirmed by HPLC-ESI^--MS (Table 5.1). DNA quantification was based on HPLC-UV analysis of dG in enzymatic digests of DNA as described previously *(220)*. Each isotopically labeled DNA oligomer was annealed to an equimolar amount of the corresponding unlabeled complementary strand.

Riboflavin Mediated Photooxidation - Riboflavin solution in 10 mM sodium cacodylate (62.5 μM, pH 7) was placed on ice and vigorously bubbled with oxygen for 1 minute. DNA duplexes (9 nmoles in 45 μL 10 mM Tris/50 mM NaCl buffer, pH 8.0) were mixed with the oxygenated riboflavin solution to reach a final concentration of 50 μM riboflavin and 40 μM DNA. Samples were transferred to glass vials and suspended in ice-cold water. The solution was irradiated for 20

Table 5.1 DNA sequences used in this study.

ID	Sequence	Calculated Molecular Weight	Observed Molecular Weight
[$^{15}N_3$,$^{13}C_1$]-*p53* exon5-G1-Me	CC^{Me}C[$^{15}N_3$,$^{13}C_1$-G]GCACC^{Me}CG^{Me}CGTC^{Me}CG^{Me}CG	5786.0	5785.7
[$^{15}N_3$,$^{13}C_1$]-*p53* exon5-G2-Me	CC^{Me}CG[$^{15}N_3$,$^{13}C_1$-G]CACC^{Me}CG^{Me}CGTC^{Me}CG^{Me}CG	5785.9	5785.7
[$^{15}N_3$,$^{13}C_1$]-*p53* exon5-G3-Me	CC^{Me}CGGCACC^{Me}C[$^{15}N_3$,$^{13}C_1$-G]^{Me}CGTC^{Me}CG^{Me}CG	5785.8	5785.7
[$^{15}N_3$,$^{13}C_1$]-*p53* exon5-G4-Me	CC^{Me}CGGCACC^{Me}CG^{Me}C[$^{15}N_3$,$^{13}C_1$-G]TC^{Me}CG^{Me}CG	5784.9	5785.7
[$^{15}N_3$,$^{13}C_1$]-*p53* exon5-G5-Me	CC^{Me}CGGCACC^{Me}CG^{Me}CGTC^{Me}C[$^{15}N_3$,$^{13}C_1$-G]^{Me}CG	5785.9	5785.7
(-)-*p53* exon5-Me	^{Me}CG^{Me}CGGA^{Me}CG^{Me}CGGGTGC^{Me}CGGG	5981.4	5981.9
[$^{15}N_3$,$^{13}C_1$]-*p53* exon7-G4-Me	ATGGG^{Me}C[$^{15}N_3$,$^{13}C_1$-G]GCATGAAC^{Me}CGGAGGCCCA	7773.7	7774.1
[$^{15}N_3$,$^{13}C_1$]-*p53* exon7-G5-Me	ATGGG^{Me}CG[$^{15}N_3$,$^{13}C_1$-G]CATGAAC^{Me}CGGAGGCCCA	7773.8	7774.1
[$^{15}N_3$,$^{13}C_1$]-*p53* exon7-G6-Me	ATGGG^{Me}CGGCAT[$^{15}N_3$,$^{13}C_1$-G]AAC^{Me}CGGAGGCCCA	7773.7	7774.1
[$^{15}N_3$,$^{13}C_1$]-*p53* exon7-G7-Me	ATGGG^{Me}CGGCATGAAC^{Me}C[$^{15}N_3$,$^{13}C_1$-G]GAGGCCCA	7774.1	7774.1
[$^{15}N_3$,$^{13}C_1$]-*p53* exon7-G8-Me	ATGGG^{Me}CGGCATGAAC^{Me}CG[$^{15}N_3$,$^{13}C_1$-G]AGGCCCA	7774.2	7774.1
[$^{15}N_3$,$^{13}C_1$]-*p53* exon7-G9-Me	ATGGG^{Me}CGGCATGAAC^{Me}CGGA[$^{15}N_3$,$^{13}C_1$-G]GCCCA	7774.0	7774.1
(-)-*p53* exon7-Me	TGGGCCC^{Me}CGGTTCATGC^{Me}CGCCCAT	7614.2	7614.0
[$^{15}N_3$,$^{13}C_1$]-*p53* exon8-G2-Me	GCTTT[$^{15}N_3$,$^{13}C_1$-G]AGGTG^{Me}CGTGTTTGTG	6548.3	6548.9
[$^{15}N_3$,$^{13}C_1$]-*p53* exon8-G3-Me	GCTTTGA[$^{15}N_3$,$^{13}C_1$-G]GTG^{Me}CGTGTTTGTG	6548.3	6548.9
[$^{15}N_3$,$^{13}C_1$]-*p53* exon8-G4-Me	GCTTTGAG[$^{15}N_3$,$^{13}C_1$-G]TG^{Me}CGTGTTTGTG	6548.3	6548.9
[$^{15}N_3$,$^{13}C_1$]-*p53* exon8-G5-Me	GCTTTGAGGT[$^{15}N_3$,$^{13}C_1$-G]^{Me}CGTGTTTGTG	6548.3	6548.9
[$^{15}N_3$,$^{13}C_1$]-*p53* exon8-G6-Me	GCTTTGAGGTG^{Me}C[$^{15}N_3$,$^{13}C_1$-G]TGTTTGTG	6548.3	6548.9
[$^{15}N_3$,$^{13}C_1$]-*p53* exon8-G7-Me	GCTTTGAGGTG^{Me}CGT[$^{15}N_3$,$^{13}C_1$-G]TTTGTG	6548.3	6548.9
(-)-*p53* exon8-Me	CACAAACA^{Me}CGCACCTCAAAGC	6336.2	6336.0
[$^{15}N_3$,$^{13}C_1$]-*p53* exon7-codon245	ATGGGC[$^{15}N_3$,$^{13}C_1$-G]GCATGAAC	4646.1	4645.4
[$^{15}N_3$,$^{13}C_1$]-*p53* exon7-codon245-Me	ATGGG^{Me}C[$^{15}N_3$,$^{13}C_1$-G]GCATGAAC	4660.1	4659.5
(-)-*p53* exon7-codon245	GTTCATGCCGCCCAT	4504.0	4503.1
(-)-*p53* exon7-codon245-Me	GTTCATGC^{Me}CGCCCAT	4518.0	4517.4

minutes with a 60-watt tungsten bulb positioned 2 cm from the vial. Following photooxidation, samples were promptly transferred to low actinic vials and immediately frozen in dry ice, then stored at -80 °C until enzymatic digestion.

Enzymatic Hydrolysis of Oxidized DNA **-** Samples were thawed and diluted with buffer to reach the final concentration of 25 mM ammonium acetate/2.5 mM zinc chloride/5 mM desferrioxamine (DFO), pH 5.3, in a final volume of 450 μL. DNA hydrolysis was performed in the presence of nuclease P1 (18 U) and alkaline phosphatase (60 U) for 2 hours at 37 °C. The completeness of hydrolysis was confirmed by HPLC-UV. Samples were divided into two equal aliquots for separate analyses of 8-oxo-dG and oxazolone.

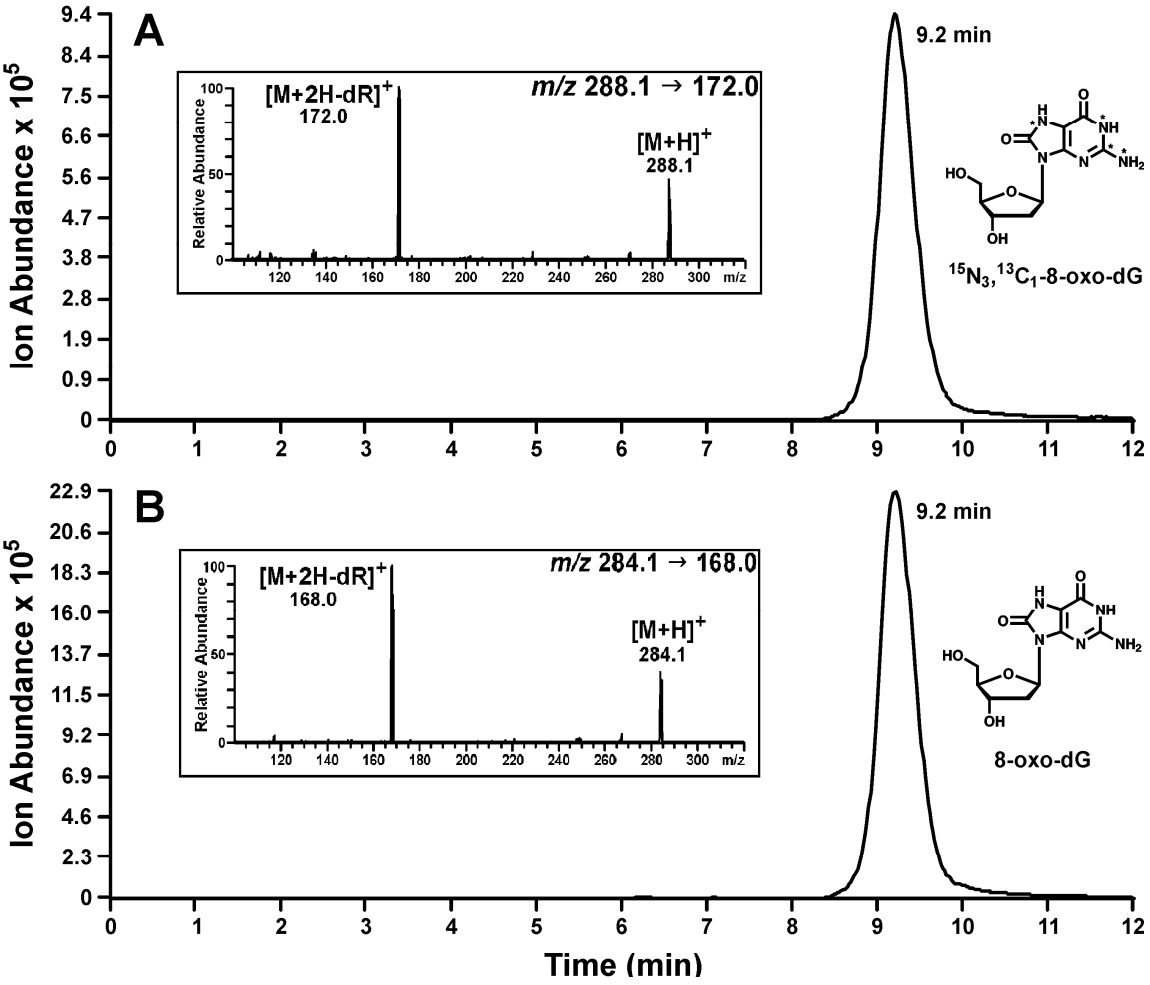
Sample Preparation for HPLC-ESI$^+$-MS/MS Analysis of 8-oxo-dG **-** 8-oxo-dG and $^{15}N_3$,$^{13}C_1$-8-oxo-dG were isolated from DNA hydrolysates by offline HPLC using an Agilent 1100 HPLC system. A Synergi Hydro-RP [4.6 x 250 mm, 4 um, Phenomenex, Torrance, CA] column was eluted at 1 mL/minute with a gradient of acetonitrile (solvent B) and 10 mM ammonium formate, pH 4.2 containing 6% methanol (solvent A). Solvent composition was maintained at 0% B from 0 to 32 min, increased to 50% B from 32 to 36 min, maintained at 50% B for 4 min, and brought back to 0% B by 43 min, followed by equilibration at 0% B for 17 min. HPLC fractions containing 8-oxo-dG and $^{15}N_3$,$^{13}C_1$-8-oxo-dG (32.6 – 35.2 min) were collected using an Agilent 1100 fraction collector, immediately frozen in dry ice, and dried under reduced pressure.

Sample Preparation for HPLC-ESI$^+$-MS/MS analysis of Oxazolone **-** Enzymatic hydrolysates of photooxidized DNA were incubated for 18 hours at 25 °C in the dark to allow for spontaneous convertion of imidazolone to oxazolone. Following incubation, oxazolone and its ^{15}N,^{13}C analog were isolated by solid phase extraction (SPE) as follows. Extract-Clean Carbograph solid phase

extraction cartridges (150 mg/4 mL, Grace, Deerfield, IL) were prepared with two volumes of methanol, followed by two volumes of water. Samples were diluted to 1 mL with water and loaded onto the cartridges. Cartridges were washed with 1 mL of water, and oxazolone was eluted with 3 mL of 20% methanol in water. The SPE fractions containing the analyte were dried under vacuum, re-dissolved in 20 uL of water, and transferred to autosampler vials for HPLC-ESI-MS/MS analysis.

Capillary HPLC-ESI$^+$-MS/MS of 8-Oxo-dG and Oxazolone - Quantitative analysis of 8-oxo-dG and oxazolone was performed with a Thermo TSQ Quantum Ultra mass spectrometer (Palo Alto, CA) interfaced with a Waters nanoAcquity UPLC system (Milford, MA). For HPLC-ESI$^+$-MS/MS analysis of 8-oxo-dG, an Agilent Extend-C18 column (0.5 x 150 mm, 3.5 um, Agilent, Santa Clara, CA) was eluted with a flow rate of 11 μL/min at 25 °C. The solvent composition was maintained at 13% methanol in 10 mM ammonium formate, pH 4.2 (isocractic elution). Using this method, 8-oxo-dG eluted at ~ 9.2 min. Electrospray ionization was typically achieved at a spray voltage of 3.2 kV and a capillary temperature of 250 °C. Quantitative analyses were performed in the selected reaction monitoring (SRM) mode. The first quadrupole was set to isolate the protonated molecules of 8-oxo-dG ($[M + H]^+$, *m/z* 284.1) and $^{15}N_3$,$^{13}C_1$-8-oxo-dG (*m/z* 288.1), and their fragmentation was induced in the second quadrupole serving as a collision cell. CID was performed with Ar as a collision gas (1.0 mTorr) at a collision energy of 14 V. The third quadrupole was set to detect the product ions corresponding to the neutral loss of deoxyribose ($[M + 2H – dR]^+$): *m/z* 168.1 for 8-oxo-dG and *m/z* 172.1 for $^{15}N_3$,$^{13}C_1$-8-oxo-dG (Figure 5.1). The lower limit of detection for 8-oxo-dG standard was 3 fmol (S/N = 10). The MS parameters were optimized for maximum response during infusing of a standard solution of 8-oxo-dG and may vary slightly between experiments.

Figure 5.1 HPLC-ESI$^+$-MS/MS analysis of 8-oxo-dG (A) and $^{15}N_3$, $^{13}C_1$–8-oxo-dG (B) in enzymatic hydrolysates of a $^{15}N_3$, $^{13}C_1$–dG containing DNA duplex subjected to photooxidation in the presence of riboflavin.



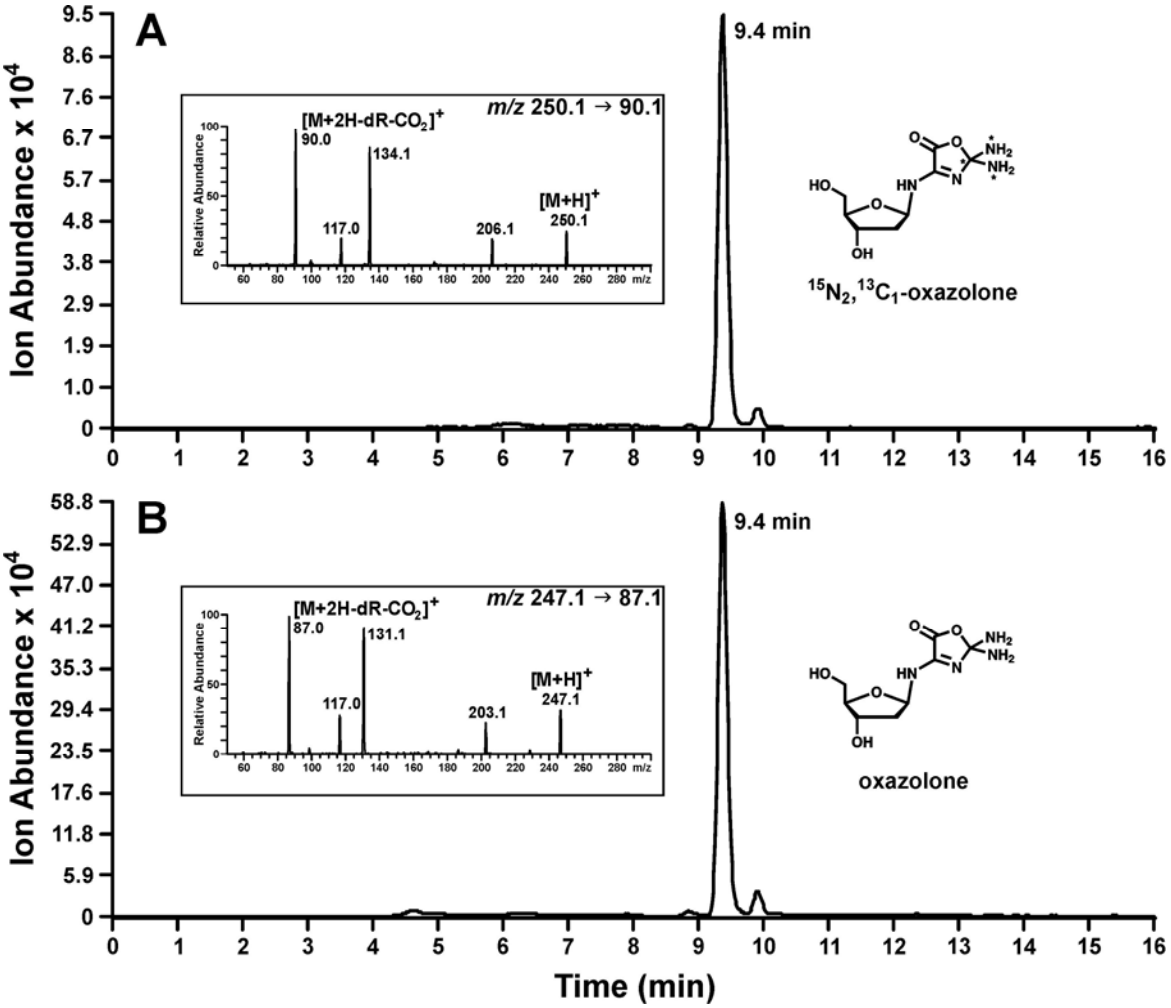
For analysis of oxazolone, a Thermo-Finnigan Hypercarb column (0.5 x 100 mm, 5 μm, Thermo Scientific, Waltharm, MA) was eluted at a flow rate of 12 μL/min with a gradient of 3:1 isopropanol: acetonitrile (solvent B) in 0.05% acetic acid (solvent A). HPLC solvent composition was gradually changed as follows: 0 min, 1.5% B; 7.1 min, 9.5% B; 7.6 min, 1.5% B; 16 min, 1.5% B. Using these conditions, oxazolone eluted at ~9.4 min. Typically, electrospray ionization (ESI) was achieved at a spray voltage of 2.8 kV and a capillary temperature of 250 °C. Quantitative analyses were performed in the SRM mode. The first quadrupole was set to isolate the protonated molecules of oxazolone ($[M + H]^+$, *m/z* 247.1) and $^{15}N_2$,$^{13}C_1$-oxazolone (*m/z* 250.1), and their fragmentation was induced in the second quadrupole serving as a collision cell. Typically, collision-induced-dissociation (CID) was performed with Ar as a collision gas (1.0 mTorr) at a collision energy of 14 V. The third quadrupole was set to detect the product ions corresponding to the neutral loss of deoxyribose and CO_2 ($[M + 2H - dR - CO_2]^+$): *m/z* 87.1 for oxazolone and *m/z* 90.1 for $^{15}N_2$,$^{13}C_1$-oxazolone (Figure 5.2). The lower limit of detection for oxazolone standard was 5 fmol (S/N = 10). The mass spectrometer parameters were optimized for maximum response during infusing of a standard solution of oxazolone and may vary slightly between experiments.

Calculations and Statistical Analysis - The extent of oxidative adduct formation at the $^{15}N_3$,$^{13}C_1$-labeled guanine (G_x) was calculated from the corresponding amounts of adduct using the equation:

$$\% \text{ oxidation at } G_x = A_{Labeled}/(A_{Labeled} + A_{Unlabeled}) * 100\%$$

Where $A_{Labeled}$ and $A_{Unlabeled}$ are the areas under the HPLC-ESI-MS/MS peaks corresponding to the labeled and unlabeled adducts, respectively.

Figure 5.2 HPLC-ESI$^+$-MS/MS chromatogram of oxazolone (A) and $^{15}N_2$, $^{13}C_1$–oxazolone (B) in enzymatic hydrolysates of a $^{15}N_3$, $^{13}C_1$–dG containing DNA duplex subjected to photooxidation.



One-way analysis of variance (ANOVA) *(221)* was used to examine differences in mean percent reactivity among groups. If the overall F-test was significant, pairwise group comparisons were made using two-sided, two-sample t-tests, using the Bonferroni correction for multiple comparisons. Each group's mean percent reactivity was also compared to a theoretical "random" reactivity value, using a two-sided, one-sample t-test, with Bonferroni-adjusted *p*-values. The theoretical value is given by 1/n, where n is the total number of groups in each set. For all t-tests, standard deviation was estimated by the square root of mean squared error (MSE) from ANOVA. All statistical analyses were conducted in SAS (Statistical Analysis Software) version 9.2. The significance level was set at 5%.

5.3 Results

5.3.1 Stable Isotope Labeling Approach

The present study employed stable isotope labeling HPLC-MS/MS approach developed in our laboratory (Scheme 5.1) *(215)* to map the distribution of two specific guanine oxidative products, 8-oxo-dG and oxazolone, along DNA duplexes representing frequently mutated regions of the *p53* tumor suppressor gene (Table 5.1).

For each target sequence, a series of DNA oligomers were synthesized containing the $^{15}N_3$,$^{13}C_1$ stable isotope tag at a different guanine base (Table 5.1). Each site-specifically labeled oligomer was annealed to the complementary strand, and the resulting DNA duplexes were subjected to photooxidation in the presence of riboflavin. Previous studies in our laboratory *(222)* established that following 20 min irradiation on ice in the presence of riboflavin, ~ 4% of total DNA strands contain 8-oxo-dG and ~2% of strands contain oxazolone, ensuring "single hit" reactions. Given our HPLC-ESI$^+$-MS/MS detection limits for 8-oxo-dG and oxazolone (3-5 fmol),

we estimated that photooxidation of 3 nmol of DNA should produce sufficient amounts of 8-oxo-dG and oxazolone to be accurately quantified by our methodology.

Oxidized DNA was enzymatically digested to 2'-deoxyribonucleosides, and then split into two equal parts (Scheme 5.1). One half of the digest was separated by off-line HPLC to isolate 8-oxo-dG and $^{15}N_3$,$^{13}C_1$ -8-oxo-dG, while the other half was incubated in the dark at 25 °C for 18 hours to allow for spontaneous conversion of imidazolone to oxazolone, followed by solid phase extraction. HPLC-ESI$^+$-MS/MS was used to determine the relative amounts of 8-oxo-dG and oxazolone formed at $^{15}N_3$,$^{13}C_1$-labeled guanine as compared to oxidative adducts originating from other (unlabeled) guanine bases within the duplex (*e.g.* Figure 5.1).

Representative HPLC-ESI$^+$-MS/MS traces corresponding to $^{15}N_3$,$^{13}C_1$-8-oxo-dG and unlabeled 8-oxo-dG generated following photooxidation of an $^{15}N_3$,$^{13}C_1$-dG containing DNA duplex are shown in Figure 5.1. HPLC-ESI$^+$-MS/MS quantification is based on selected reaction monitoring of the transition *m/z* 284.1 $[M + H]^+ \rightarrow$ *m/z* 168.0 $[M + 2H - dR]^+$) for 8-oxo-dG (Figure 5.1B) and the corresponding transition: *m/z* 288.1 $\rightarrow$ *m/z* 172.0 for $^{15}N_3$,$^{13}C_1$-8-oxo-dG (Figure 5.1A). The extent of 8-oxo-dG formation at the $^{15}N_3$,$^{13}C_1$ labeled position was calculated directly from the areas of HPLC-ESI$^+$-MS/MS peaks corresponding to the labeled and unlabeled adducts. A similar procedure was followed for oligomers containing $^{15}N_3$,$^{13}C_1$-dG at different sites, making it possible to accurately determine the amount of 8-oxo-dG at each site along the DNA duplex.

HPLC-ESI$^+$-MS/MS quantitation of oxazolone was conducted analogously using isotope ratio mass spectrometry (Figure 5.2). It should be noted the ^{15}N labeled nitrogen from the N-2 position of $^{15}N_3$,$^{13}C_1$-deoxyguanosine is lost during the reaction sequence that produces oxazolone (Scheme 1.4) *(218)*, resulting in an $^{15}N_2$,$^{13}C_1$-labeled product. HPLC-ESI$^+$-MS/MS quantification of oxazolone and $^{15}N_2$,$^{13}C_1$- oxazolone is based on a loss of deoxyribose and CO_2

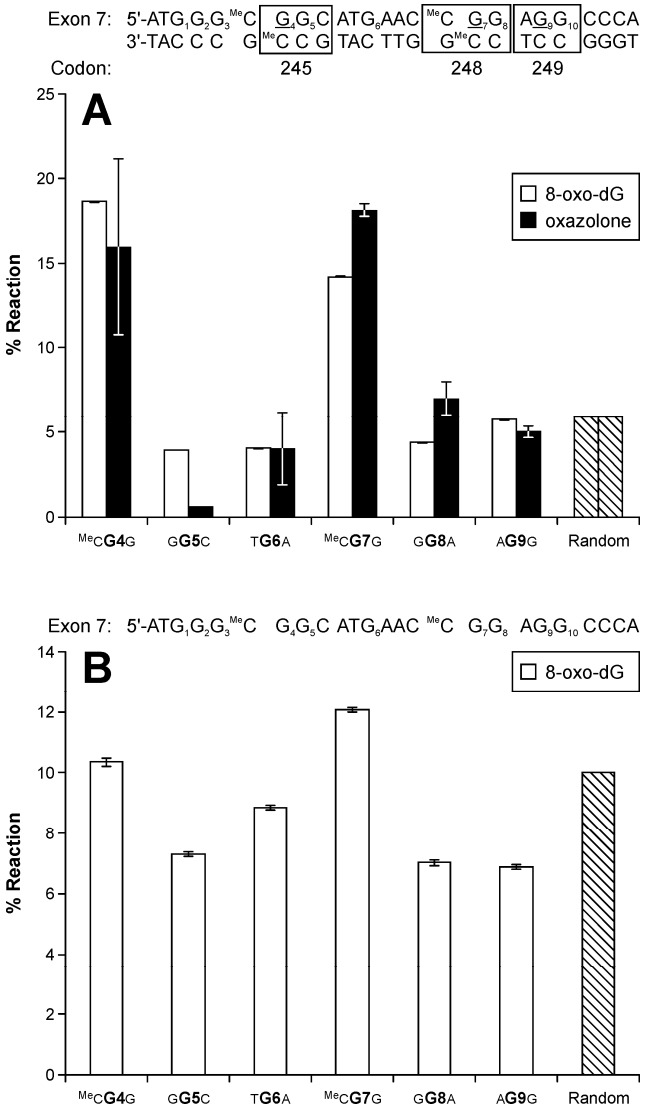
from the $[M + H]^+$ ions subjected to collision induced dissociation (*m/z* 247.1 $[M + H]^+ \rightarrow$ *m/z* 87.2 $[M + 2H - dR - CO_2]^+$ and *m/z* 250.1 → *m/z* 90.1, respectively) (Figure 5.2).

5.3.2 Distribution of 8-oxo-dG and oxazolone in riboflavin mediated photooxidation in *p53* Exon-7-derived DNA duplex

Base substitution mutations at codons 245 (***G****GC* → ***T****GC*), 248 (*C****G****G* → *C****T****G*), and 249 (*A****G****G* → *A****T****G*) of the *p53* tumor suppressor gene are common in smoking-associated lung cancer (Figure 1.3) *(223)*. To examine whether these sites are preferential targets for photooxidation, a series of double-stranded oligodeoxynucleotides were synthesized representing a portion of *p53* gene sequence containing codons 243-250 (5'-ATG$_1$ G$_2$G$_3$ ^{Me}C **G$_4$G$_5$**C AT**G$_6$** AAC ^{Me}C**G$_7$G$_8$** A**G$_9$**G$_{10}$ CCC A-3', where ^{Me}C = 5-methylcytosine) (Table 5.1). $^{15}N_3$,$^{13}C_1$-dG was introduced at one of the highlighted positions (**G$_4$**, **G$_5$**, **G$_6$**, **G$_7$**, **G$_8$**, or **G$_9$**, Table 5.1), and ^{Me}C was inserted in both strands at the endogenously methylated CG sites, 5'- ^{Me}C**G$_4$** and 5'- ^{Me}C**G$_7$** (Table 5.1). Following riboflavin-mediated photooxidation and DNA hydrolysis, the extent of 8-oxo-dG and oxazolone originating from each location of interest along this duplex was calculated based on isotope ratio HPLC-ESI$^+$-MS/MS as described above (Scheme 5.1).

As shown in Table 5.1, *p53* exon-7-derived DNA duplexes contain a total of 17 guanines, therefore a completely random reaction should lead to 5.88% of oxidation occurring at each individual G (100%/17 = 5.88%). One of the guanines examined (**G$_9$**) shows reactivity similar to the theoretical value (5.71%, $p > 0.5$) (white bars in Figure 5.3A). However, the remaining five guanines show significant differences between adduct yields at different guanine bases. The general reactivity order is **G$_7$** ≈ **G$_4$** > **G$_8$** > **G$_9$** > **G$_6$** > **G$_5$** (Figure 5.3A). It should be noted that the sites of the highest 8-oxo-dG adduct formation, **G$_4$** (18.6%, $p < 0.0001$) and **G$_7$** (14.2%, $p < 0.001$), are located within endogenously methylated CG dinucleotides and a flanked by a 3'-G

Figure 5.3 Distribution of 8-oxo-dG (white bars) and oxazolone (black bars) along *p53* exon 7 derived DNA duplexes (A) and single stranded DNA (B) as determined by stable isotope labeling. Riboflavin concentration was 62.5 μM, and samples were purged with oxygen for 1 minute, followed by irradiation for 20 min at 0ºC. The known *p53* lung cancer mutation "hotspots" are given in underlined type.



(*5'-^{Me}C**G**G-3'* sequence context). In contrast, nearly random formation of 8-oxo-G adducts was observed for the same single-stranded sequence (Figure 5.3B), suggesting that electron hole transfer requires double stranded DNA structure.

A similar distribution pattern was obtained for oxazolone (black bars in Figure 5.3A), with the majority of adducts originating from **G_4** (15.9%, $p < 0.001$) and **G_7** (18.1%, $p = 0.003$), followed by **G_8** (*G**G_8**A,* 7.0%), **G_9** (*A**G_9**G,* 5.0%), **G_6** (*T**G_6**A,* 4.0%), and **G_5** (*G**G_5**C,* 0.6%). Interestingly, the most frequently photooxidized positions in the *p53* exon 7-derived sequences (**G_4** and **G_7**) coincide with the known lung cancer mutational hotspots within *p53* codons 245 and 248 (Figure 1.3).

5.3.3 Distribution of 8-oxo-dG and oxazolone in riboflavin mediated photooxidation in *p53* Exon-5-derived DNA duplex

Codons 154 (G***G***C), 157 (***G***TC), and 158 (C***G***C) of the *p53* tumor suppressor gene are among the major mutational "hotspots" in smoking-induced lung cancer (Figure 1.3) *(224)*. To examine the formation of 8-oxo-dG and oxazolone at these frequently mutated guanines and surrounding sequence, a series of stable isotope labeled DNA duplexes were prepared representing *p53* codons 153-158 (Table 5.1). $^{15}N_3$,$^{13}C_1$-dG was sequentially incorporated at **G_1, G_2, G_3, G_4,** or **G_5**, and ^{Me}C was placed at all physiologically methylated sites: 5'-CC^{Me}C **G_1G_2**C ACC ^{Me}C**G_3**^{Me}C **G_4**TC ^{Me}C**G_5**^{Me}C G_6 [(+) strand] (Table 5.1). Following riboflavin-mediated photooxidation and DNA hydrolysis, the relative amounts of 8-oxo-dG and oxazolone formed at each guanine of interest was determined by HPLC-ESI$^+$-MS/MS as described above (Scheme 5.1).

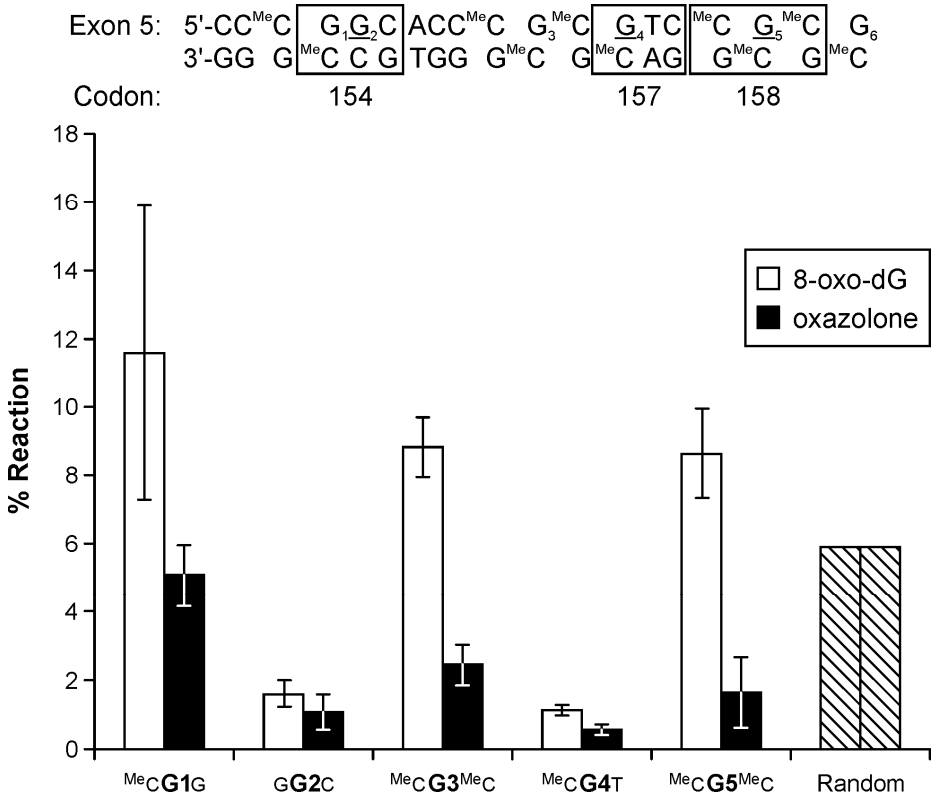
Since *p53* exon-5-derived duplex 5'-CC^{Me}C **G_1G_2**C ACC ^{Me}C**G_3**^{Me}C **G_4**TC ^{Me}C**G_5**^{Me}C G_6 contains 17 guanine bases, completely random reaction should result in 5.88% of adducts originating from each guanine (100%/17 = 5.88%). However, we found that the distribution of 8-oxo-dG and oxazolone adducts along this sequence was not uniform, with 8-oxo-dG adduct

formation following the order **G₁** (*^{Me}CG₁G₂,*) > **G₃** (*^{Me}CG₃^{Me}C*) ≈ **G₅** (*^{Me}CG₅^{Me}C*) >> **G₂** (*G₁G₂C*) ≈ **G₄** (*^{Me}CG₄T*) (Figure 5.4, white bars). A strong sequence preference for the formation of 8-oxo-dG was observed at **G₁**, which is the 5' guanine within GG repeat (11.6% of total reaction). Increased adduct numbers were also observed at **G₃** (8.8%) and **G₅** (8.6%), which have identical sequence context (*^{Me}CG^{Me}C*). In contrast, 8-oxo-dG adduct formation at **G₂** and **G₄** (1.1-1.6%) was significantly lower than the theoretical value ($p < 0.05$). A similar reactivity order (**G₁** (*^{Me}CG₁G₂*) > **G₃** (*^{Me}CG₃^{Me}C*) > **G₅** (*^{Me}CG₅^{Me}C*) > **G₂** (*G₁G₂C*) > **G₄** (*^{Me}CG₄T*)) was observed for oxazolone adducts, although the overall adduct yields were lower than those observed for 8-oxo-dG (black bars in Figure 5.4). Taken together, these results indicate that the distribution of photooxidative adducts within *p53* exon 5-derived duplex is dependent on the local sequence context. Interestingly, one of the preferentially oxidized sites (**G₅**) coincides with the major lung cancer mutational "hotspot" at *p53* codon 158. However, the other two readily oxidized sites (**G₁** and **G₃**) are less frequently mutated in lung cancer (Figures 1.3 and 5.4).

5.3.4 Distribution of 8-oxo-dG and oxazolone in riboflavin mediated photooxidation in *p53* Exon-8-derived DNA duplex

Another important lung cancer mutational "hotspot" is codon 273 located within *p53* exon 8 (*CGT* → *CTT*) (Figure 1.3) *(225;226)*. To investigate the distribution of 8-oxo-dG and oxazolone adducts in this critical region of the *p53* gene, a series of synthetic DNA duplexes, 5'-G_1CT TT**G₂** A**G₃G₄** T**G₅**^{Me}C **G₆**T**G₇** TTT G_8TG_9 (+ strand) (Table 5.1), were prepared representing *p53* codons 269-275. $^{15}N_3$,$^{13}C_1$-dG was introduced at one of the highlighted positions (**G₂**, **G₃**, **G₄**, **G₅**, **G₆**, or **G₇**, Table 5.1), and ^{Me}C was inserted at the physiologically methylated site, 5'-^{Me}C**G₆** (Table 5.1). Following riboflavin-mediated photooxidation and DNA hydrolysis,

Figure 5.4 Distribution of 8-oxo-dG (white bars) and oxazolone lesions (black bars) along DNA duplexes derived from *p53* exon 5 as determined by stable isotope labeling. Riboflavin concentration was 62.5 μM, and samples were purged with oxygen for 1 minute, followed by irradiation for 20 min at 0ºC. The known *p53* lung cancer mutation "hotspots" are given in underlined type.



stable isotope labeling HPLC-ESI$^+$-MS/MS was employed to quantify the formation of 8-oxo-dG and oxazolone at each location of interest (Scheme 5.1).

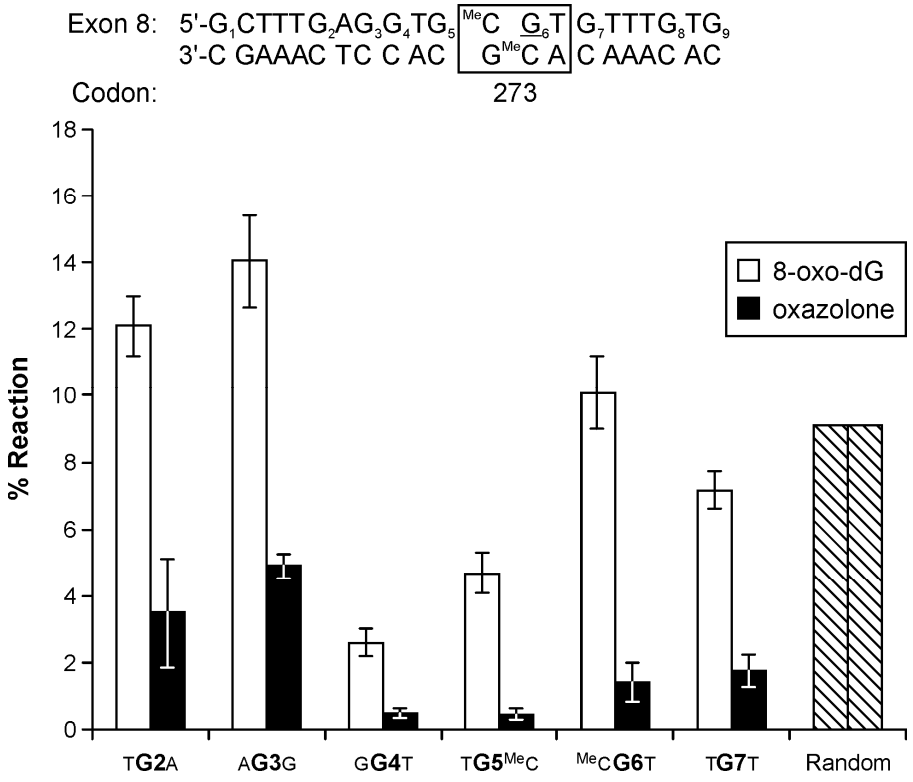
Because the *p53* exon 8-derived DNA duplex contains a total of 11 guanines, in the event of base sequence not imposing any effect on adduct formation, riboflavin-mediated photooxidation should result in 9.09% of the adducts produced at each guanine (100%/11 = 9.09%). We found that for this DNA duplex, the highest number of 8-oxo-dG adducts was produced at **G$_3$** (*A***G$_3$***G$_4$* in *p53* codon 271, 14.0%, $p < 0.0001$) (Figure 5.5, white bars). The second and third most reactive sites were **G$_2$** (*T***G$_2$***A* codon 270*,* 12.1%, $p < 0.0001$), and **G$_6$** (^{Me}*C***G$_6$***T in* codon 273*,* 10.1%) (white bars in Figure 5.5). In contrast, the reactivities of **G$_4$** (*G$_3$***G$_4$***T,* 2.6%), **G$_5$** (*T***G$_5$**^{Me}*C*, 4.7%), and **G$_7$** (*T***G$_7$***T*, 7.2%) were significantly less than the random value ($p \leq 0.02$).

Oxazolone adduct formation in the *p53* exon 8-derived sequence follows the same general order (**G$_3$** (*A***G$_3$***G$_4$*, 4.9%) > **G$_2$** (*T***G$_2$***A*, 3.5%) > **G$_6$** (^{Me}*C***G$_5$**^{Me}*C*, 1.4%) ≈ **G$_7$** (*T***G$_7$***T*, 1.7%) > G4 (*G$_3$***G$_4$***T*, 0.5%) ≈ G5 (*T***G$_5$**^{Me}*C*, 0.5%)), although oxazolone adduct yields are uniformly low at all five guanine bases tested (black bars in Figure 5.5). We hypothesize that this is due to the preferential formation of oxazolone lesions at terminal guanines (G_1 and G_9), which were not isotopically labeled in this study. For both adducts, the site of the highest reactivity (**G$_3$** in *A***G$_3$***G$_4$*) does not coincide with the known mutational hotspot for G → T transversions in smoking-related lung cancer (**G$_6$** in *p53* codon 273) (Figures 1.3 and 5.5), suggesting that other factors are responsible for increased mutagenesis at this site.

5.3.5 Effects of neighboring ^{Me}C on the formation of 8-oxo-dG and oxazolone in *p53*-derived DNA duplex

Our quantitative results shown in Figures 5.3-5.5 reveal that the yields of photooxidation-induced 8-oxo-dG and oxazolone lesions are elevated at endogenously methylated CG sequences,

Figure 5.5 Distribution of 8-oxo-dG (white bars) and oxazolone (black bars) lesions along DNA duplexes derived from *p53* exon 8 as determined by stable isotope labeling. Riboflavin concentration was 62.5 μM, and samples were purged with oxygen for 1 minute, followed by irradiation for 20 min at 0ºC. The known *p53* lung cancer mutation "hotspots" are given in underlined type.



e.g. those within *p53* codons 157, 158, 245, 249, and 273. Because each methylated CG dinucleotide contains two ^{Me}C nucleobases (the 5'-neighboring ^{Me}C, and the base-paired ^{Me}C), either one of them may contribute to the increased susceptibility of target guanine towards riboflavin-mediated photooxidation.

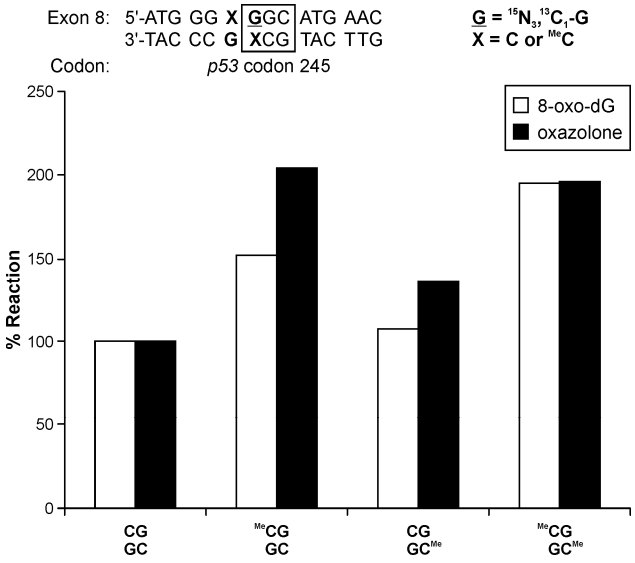
In order to separately analyze the effects of the 5'-flanking and base-paired ^{Me}C on reactivity of guanine bases within CG dinucleotides, synthetic DNA duplexes were constructed containing methylated cytosine opposite the labeled G, immediately 5' to the target guanine, or in both DNA strands as observed physiologically (Table 5.1). Following photooxidation, the effects of cytosine methylation on oxidative adduct yields at the $^{15}N_3$,$^{13}C_1$-labeled position were determined by mass spectrometry of enzymatic digests. The relative reactivity data was normalized to the values obtained for unmethylated CG dinucleotide (100%).

We found that 8-oxo-dG adduct formation at the $^{15}N_3$,$^{13}C_1$-labeled G in codon 245 was elevated when ^{Me}C precedes the target guanine (152 %, $p < 0.0001$) and in fully methylated dinucleotide (194.6%, $p < 0.0001$), but is unaffected by ^{Me}C at the base-paired position (107.5%, $p = 0.16$) (white bars in Figure 5.6). Similar results were obtained for oxazolone (black bars in Figure 5.6), with a 2-fold reactivity increase in the presence of 5'-neighboring ^{Me}C (203.74%, $p < 0.0001$) and in fully methylated dinucleotides (196.07%, $p < 0.0001$), but a smaller effect observed for G:^{Me}C base pairs (135.78%, $p = 0.75$). Taken together, these results indicate that riboflavin-mediated photooxidation of guanine is facilitated in the presence of 5'-neighboring ^{Me}C.

5.4 Discussion

The *p53* tumor suppressor gene is a major target for genetic damage in smoking-induced cancer *(223)*, with approximately 56% of total lung tumors of smokers bearing *p53* mutations

Figure 5.6 Effect of neighboring ^{Me}C on the formation of 8-oxo-dG (white bars) and oxazolone (black bars) within *p53* codon 245 in double-stranded DNA sequences derived from *p53* exon 7 as determined by stable isotope labeling. Riboflavin concentration was 62.5 μM, and samples were purged with oxygen for 1 minute, followed by irradiation for 20 min at 0ºC.



(92). The p53 protein functions as a cell cycle checkpoint to respond to DNA damage and to assist DNA repair processes prior to replication and mitosis, thus limiting the occurrence of mutations *(227)*. The majority of the *p53* base substitution mutations observed in smoking-associated lung tumors are G → T transitions frequently occurring at *p53* exons 5, 7, and 8 (Figure 1.3) *(94;95)*. Prominent *p53* mutational "hotspots" include endogenously methylated ^{Me}CG dinucleotides within codons 154 (*G**G**C* → *G**T**C*), 157 (***G**TC* → ***T**TC*), 158 (*C**G**C* → *C**T**C*), 245 (***G**GC* → ***T**GC*), 248 (*C**G**G* → *C**T**G*), 249 (*A**G**G* → *A**T**G*), and 273 (*C**G**T* → *C**T**T*) (Figure 1.3). Because *p53* exons 5-8 encode the sequence-dependent DNA binding domain of the *p53* protein, mutations in these regions directly inactivate the p53 protein, leading to an increased probability of further genetic damage *(92)*.

One possible mechanism for the increased mutagenesis at ^{Me}CG sites within the *p53* gene involves targeted binding of metabolically activated tobacco carcinogens to these sequences. For example, previous studies in our laboratory revealed that bulky N^2-guanine adducts of polycyclic aromatic hydrocarbons (PAH) present in tobacco smoke are preferentially formed at endogenously methylated CG dinucleotides *(96-99)*. Our recent study indicates that PAH diolepoxides preferentially form physical complexes with ^{Me}C:G base pairs, leading to increased adduct yields at these sites *(100)*. However, other constituents of tobacco smoke such as reactive oxygen species may also target ^{Me}CG dinucleotides, contributing to mutagenesis.

Oxidative DNA damage is a likely contributor to the pathogenesis of lung cancer *(101)*. Smokers excrete increased amounts of 8-oxo-dG in their urine, which is reversed upon smoking cessation *(228)*. 8-oxo-dG levels are also increased in leukocyte DNA of smokers as compared with non-smoking controls *(101)*. This oxidative DNA damage can be induced by reactive oxygen species initially present in tobacco smoke and those generated from redox cycling of tobacco carcinogens *(229-231)*. The increased cellular load of oxidative DNA lesions in the tissues of smokers likely plays a role in lung tumor induction. It has been shown that smokers

who exhibit low levels of hOGG1, a repair protein responsible for the removal of 8-oxo-dG and other oxidative lesions, are at an increased risk of lung cancer *(195)*.

Both 8-oxo-dG and oxazolone have been shown to cause G → T transversion mutations in both prokaryotic and eukaryotic cells *(102)*. Furthermore, cytosine methylation has been demonstrated to facilitate guanine oxidation at the base-paired position. For example, Kawai and co-workers found that the rate of one-electron oxidation of guanine was accelerated when G was base-paired with ^{Me}C as compared to unmethylated G:C base-pair *(103)*. Therefore, we hypothesized that oxidative damage may contribute to G → T transversions observed at endogenously methylated CG dinucleotides of *p53* exons 5, 7, and 8.

The present study was initiated to investigate the formation of 8-oxo-dG and oxazolone formations in critical regions of the *p53* tumor suppressor gene (Table 5.1). Our major goal was to determine the effects of local sequence context and the presence of 5'-neighboring ^{Me}C on reactivity of guanine nucleobases within *p53* gene sequence. Our results (Figures 5.3-5.5) indicate that riboflavin-induced 8-oxo-dG and oxazolone lesions are produced non-randomly in all DNA sequences examined. In general, 8-oxo-dG yields follow the following order: ^{Me}C**G**G ≈ A**G**G ≈ ^{Me}C**G**T ≈ ^{Me}C**G**C ≈ A**G**G > T**G**T > T**G**A ≈ G**G**C ≈ C**G**T ≈ T**G**C ≈ G**G**A ≈ G**G**T. Similar, but less pronounced sequence preferences were observed for oxazolone (Figures 5.3-5.5).

The observed preferential reactivity of 5' Gs within guanine repeats towards riboflavin-mediated photooxidation is consistent with previous studies *(86;206)*. Once formed, electron holes generated at random sites within DNA sequence can migrate through DNA duplex towards the sites with the lowest redox potential, getting preferentially trapped at guanine residues located at 5'-position to another guanine (e.g. 5'-GG or 5'-GGG) *(206;216)*. *Ab initio* calculations by Saito and coworkers suggested that these are the most electron-rich sites in the B-form of DNA *(86;206)* and therefore may act as thermodynamic "sinks" for oxidative damage following long-

range charge transport from other sites in the helix (G-G stacking rule) *(209)*. Margolin *et.al.* quantified riboflavin-induced damage in 5'-^{32}P-labeled oligodeoxynucleotides containing guanine in all possible three-base sequence contexts (5'-XGY-3') based on sequencing gels by converting guanine lesions to strand breaks *(87;88)*. The relative amount of oxidative lesions originating from each guanine was then plotted against their corresponding ionization potentials (IPs). A strong inverse correlation was observed between the sequence-specific IP and the reactivity of guanine, suggesting that sequence-specific IPs were the major determinants of guanine's reactivity towards one-electron oxidation in DNA duplexes *(87;88)*. In contrast, single stranded DNA did not show the same sequence selectivity due to limited opportunities for π- π staking in a random coil structure *(87;88)*.

Another factor that governs the reactivity of individual guanines upon riboflavin-mediated photooxidation of *p53* derived duplexes is the presence of endogenous 5-methylcytosine. For example, endogenously methylated CG dinucleotides at codons 154, 156, 158, 245, 248, and 273 are preferentially oxidized as compared to neighboring guanines (Figures 5.3-5.5). One possible explanation is that the presence of electron donating 5-methyl group on cytosine lowers the redox potential of its partner guanine and stabilizes the resulting radical cation through Watson-Crick hydrogen bonds *(103)*. Indeed, when electron-donating methyl group is introduced to the C5 position of cytosine, increased rate of one-electron oxidation relative to unsubstituted G:C base-pair was observed *(103)*. An opposite effect (reduced reativity) was observed when electron-withdrawing Br substitutuent was introduced on cytosine *(103)*. However, our experiments with partially methylated sequences revealed that the base-paired ^{Me}C has a less pronounced effect on oxidative adduct yields than the 5'-neighboring ^{Me}C, which doubles the number of 8-oxo-dG and oxazolone lesions (Figure 5.6). An alternative explanation involves favorable physical interactions between riboflavin and methylated ^{Me}CG dinucleotides. Riboflavin is known to intercalate into GC and CG base pairs *(212)*, and the presence of 5-methyl

group on cytosine increases its molecular polarizability *(232)*, potentially facilitating π- π interactions between DNA and the intercalated photosensitizer. Upon irradiation, the intercalated riboflavin is expected to abstract an electron from guanine and induce the subsequent photooxidation events. Either mechanism will result in the preferential photooxidation of guanines in ^{Me}CG dinucleotides.

Our study was the first to examine the formation of specific oxidative adducts along DNA sequences. We found that the overall distribution patterns of 8-oxo-dG and oxazolone adducts were similar (Figures 5.3-5.5), probably a result of their formation from a common radical cation intermediate. However, the overall yields of oxazolone adducts in exon 5- and exon 8-derived sequences were low at all guanine bases examined (black bars in Figures 5.4 and 5.5). This can be explained by the preferential formation of oxazolone lesions at the unlabeled guanine bases located at the duplex termini (*e.g.* DNA sequences in Figure 5.4 and 5.5). For example, Lee and Dedon reported that guanine oxidation in DNA by carbonate radical anion is preferred at the end of the sequence (end-effects) due to increased solvent exposure *(233)*.

The fate of the radical cation produced upon initial electron abstraction from guanine is influenced by solvent accessibility and the presence of oxygen. Reaction with free oxygen leads to guanine radical cation oxidation to imidazolone/oxazolone, while reactions with water result in 8-oxo-dG formation (Scheme 1.4). 8-oxo-dG can further react with oxygen to form the hydantoin compounds or react again with water to form imidazolone/oxazolone *(202;205)*. Therefore different amounts of oxygen present can affect the relative amounts of imidazolone/oxazolone and 8-oxo-dG formed. Furthermore, the nature of the oxidant involved in DNA oxidation is also expected to affect the sequence-dependence of guanine oxidation. For example, Margolin and co-workers observed a totally different DNA oxidation behavior when they examined the guanine reactivity toward another one-electron oxidant nitroperoxycarbonate ($ONOOCO_2^-$) arising from inflammation *(87;88)*. Unlike the observation from riboflavin-induced oxidation, $ONOOCO_2^-$

preferentially oxidized guanines with the highest IP in GC-containing context, and such GC-containing motif appeared to be the determinant rather than IPs in $ONOOCO_2^-$-induced oxidation *(87;88)*. Therefore, sequence-dependence of guanine oxidation should be considered as a function of multiple factors, such as the nature of the oxidant, kinetics and stability of the resulting radical cation, and its subsequent chemical behaviors and/or fate *(87)*.

In conclusion, stable isotope labeling HPLC-MS/MS methodology was for the first time employed in this study to examine the formation of specific DNA oxidative lesions (8-oxo-dG and oxazolone) along DNA sequences derived *p53* tumor suppressor gene following one-electron photooxidation in the presence of riboflavin. We demonstrated that the formation of both oxidative lesions was sequence-selective rather than random, and was especially favored at the 5'-Gs in guanine repeats and guanines within ^{Me}CG dinucleotides, probably due to low guanine IP at 5'-**G**G-3' and/or the preferential intercalation of riboflavin at ^{Me}CG dinucleotides. While these results are in general consistent with those previously determined by the gel electrophoresis, our methodology avoids the high background of strand breaks resulting from 2'-deoxyribose oxidation and provides additional details about the structural identities of both oxidative lesions. Importantly, we found that the most frequently adducted positions, $\mathbf{G_5}$ in exon 5, $\mathbf{G_4}$ and $\mathbf{G_7}$ in exon 7, and $\mathbf{G_6}$ in exon 8 coincide with the known *p53* lung cancer mutation "hotspots" at codons 158 (*C**G**C*), 245 (***G**GC*), 248 (*C**G**G*), and 273 (*C**G**T*), respectively, suggesting that oxidative DNA damage is a likely contributor to the pathogenesis of lung cancer.

VI. CONCLUSIONS

6.1 DNA-protein Cross-linking by Cisplatin

DPCs are generated when proteins become covalently cross-linked to chromosomal DNA in the presence of *bis*-electrophiles, and are expected to block the binding and progression of protein complexes, and therefore interfere with critical DNA metabolic processes. Despite their potential threats to cells, the biological effects of DPCs (i.e., cytotoxicity, mutagenesis and carcinogenesis) are not well understood, and their identities as well as their atomic connections have not been well characterized, mainly due to the reasons that many cross-linking agents can generate a variety of other DNA lesions in addition to DPCs and different agents are able to induce cross-links by different mechanism leading to different DNA-protein cross-linking chemistries. In the current study, the ability of the antitumor agent cisplatin to induce DNA-protein cross-linking was examined both cell-free nuclear protein extracts and cultured cells using a combination of mass spectrometry, gel shift assay, proteomic analysis, immunological detection methods, and cytotoxicity assays.

Previous studies have employed antibodies against specific proteins, showing that several DNA-binding proteins, including HMG proteins 1, 2, and E, cytokeratins, and histone proteins, can become cross-linked to DNA in the presence of cisplatin *(73)*. Recently, Lippard *et al.* employed photoreactive aryl azide group-containing PtBPn complexes and biotinylated DNA probes to affinity purify and identify 14 nuclear proteins that bind to cisplatin-DNA adducts, including PARP-1, HMG B1/B2/B3, DNA ligase III, XRCC-1, ATP-dependent DNA helicase (Ku) 70/86, and human DNA mismatch repair protein (hMutS) *(160-162)*. However, little information is provided about other nuclear proteins involved in cisplatin-induced DNA-protein cross-linking. Furthermore, the covalent structures of resulting DPCs lesions and their cellular

abundance have not been determined. We therefore sought to characterize cisplatin-induced DNA-protein cross-linking from three perspectives: recombinant proteins (Chapter II), *in vitro* (Chapter III) and *in vivo* (Chapter IV).

Denaturing gel electrophoresis detected a concentration-dependent formation of DNA-protein conjugates following incubation of human recombinant AGT or GAPDH protein with ^{32}P-labeled oligonucleotides in the presence of cisplatin (Chapter II, Figure 2.1). Mass spectrometry analysis of human recombinant AGT protein that had been treated with cisplatin-induced dG-Pt-Cl monoadduct as a model of monoplatinated DNA confirmed these results (Chapter II, Figures 2.2 and 2.3). To further gain insight into the chemical nature of cisplatin-induced cross-links, HPLC-ESI+-MS/MS analysis of dG-Pt-AGT complexes obtained upon incubation of AGT with dG-Pt-Cl monoadduct demonstrated that cross-linking could occur at a total of six observed sites within this protein, including Glu^{110}, Lys^{125}, Cys^{145}, His^{146}, Arg^{147}, and Cys^{150} (Chapter II, Figure 2.4). Among them, Glu^{110} is located in proximity to the reactive site pocket (IPCHRV), which makes it possible be cross-linked by platinum, Lys^{125} is located at the DNA-binding domain which interacts with double stranded DNA *via* helix-turn-helix (HTH) motif, and Cys^{145}, His^{146}, Arg^{147}, and Cys^{150} are other nucleophilic active site residues that are located in the immediate proximity to DNA in the AGT-DNA complex (Chapter II, Figure 2.5). Subsequent analysis of amino acid-nucleobase conjugates resulting from total digestion of AGT-guanine cross-linking reactions revealed one of the expected structures of the cross-linked lesions as 1,1-*cis*-diammine-2-(5-amino-5-carboxypentyl)amino-2-(2'-deoxyguanosine-7-yl)-platinum(II) (dG-Pt-Lys) (Chapter II, Figure 2.8). Interestingly, our results indicated that dG-Pt-AGT complexes would likely experience a platination migration from protein to guanine in the presence of excess dG under heating, suggesting a higher thermodynamic stability of Pt-N7 guanine bond (Chapter II, Figures 2.6 and 2.7).

After establishing that cisplatin induce AGT-DNA cross-links in experiments with

purified protein, we set out to characterize DNA-protein cross-linking by cisplatin *in vitro*. Our laboratory previously developed a novel approach which coupled affinity capture enrichment with mass spectrometry-based proteomics to identify nuclear extract proteins which became cross-linked to biotinylated DNA oligomers in the presence of *bis*-electrophiles, including antitumor nitrogen mustards and DEB *(35;36)*. Using this methodology (Chapter III, Scheme 3.2), we have identified 131 nuclear proteins from human cervical carcinoma (Hela) cells which formed DPCs *in vitro* in the presence of cisplatin (Chapter III, Table 3.1 and Figures 3.1-3.3). The identified proteins displayed great diversity in terms of cellular function, transcriptional regulation, DNA replication and repair, cell cycle and homeostasis, and architectural/structural support (Chapter III, Figure 3.5). The identities of proteins observed from proteomics screen were later confirmed *via* western blotting using commercially-available antibodies (Chapter III, Figure 3.6). Importantly, HPLC-ESI$^+$-MS/MS analysis of total proteolytic digests of cross-linked proteins revealed the presence of dG-Pt-Lys conjugates (Chapter III, Figure 3.7), in agreement with our previous findings that cisplatin-induced DPC formation involves the N7-position of guanine bases within DNA and lysine residues within proteins (Chapter II, Figure 2.8). Taken together, the experiments described in Chapter III established that cisplatin can induce covalent cross-linking between a number of nuclear proteins and DNA, potentially contributing to its cytotoxic and mutagenic effects associated with cisplatin treatment in the clinic.

In order to characterize DNA-protein cross-linking by cisplatin in human fibrosarcoma (HT1080) cells, we employed a modified phenol/chloroform DNA extraction strategy, developed by our laboratory previously *(177)*, for the isolation of DPC lesions from HT1080 cells. Following drug treatment, DPCs were isolated by a modified phenol/chloroform DNA extraction incorporating proteasome inhibitors. Proteins were released from DNA by thermal platination transfer and identified by mass spectrometry-based proteomics and immunological detection. Over 250 nuclear proteins were captured on chromosomal DNA following treatment with

cisplatin (Chapter IV, Table 4.1). Four additional proteins were identified by western blotting (Chapter IV, Figure 4.6). Compared to only 10% of the total nuclear proteins in human fibroblasts *(179)*, over 48% of the identified proteins were classified in the GO database as nuclear proteins, suggesting that close proximity to DNA and interactions with chromosomal DNA contribute significantly to the likelihood of a particular protein becoming cross-linked to DNA. Additionally, approximately 40% of the identified proteins were also grouped as DNA-binding, RNA-binding, and protein-binding proteins (Chapter IV, Figure 4.4) involved in some aspects of DNA replication, transcriptional regulation, DNA damage response and repair, cell cycle control, and architectural support. HPLC-ESI$^+$-MS/MS analysis of total proteolytic digests revealed the formation dG-Pt-Lys conjugates between the N7 guanine of DNA and the ε-amino group of lysine (Chapter IV, Figure 4.7). Although cisplatin-induced DPCs spontaneously release proteins to form DNA-DNA cross-links upon heating, they appear to be stable enough under physiological conditions to inhibit DNA replication and transcription, contributing to the biological effects of cisplatin (*see Chapter II*).

In summary, we have shown that cisplatin at clinically relevant concentrations *(187)* is capable of inducing DPCs in addition to DNA-DNA cross-links, both in human recombinant proteins, *in vitro* and in mammalian cell culture. Nuclear proteins, identified by a combination of mass spectrometry-based proteomics and immunological detection, encompass a variety of cellular functions, potentially contributing to the genotoxic and cytotoxic effects of cisplatin. Future work is necessary for a better understanding of the biological effects of common *bis*-electrophiles and other cross-linking agents.

6.2 Formation of 8-oxo-dG and oxazolone lesions with *p53* derived DNA sequences following photooxidation in the presence of riboflavin

Reactive oxygen species produced as part of normal cellular metabolism and immune response can damage cellular DNA, giving rise to promutagenic nucleobase lesions *(188-190)*. Since the biological impact of a given oxidative adduct is influenced by its position within a gene sequence, many previous studies focused on determining the distribution of oxidative lesions along DNA sequences. However, since these studies have relied on gel electrophoresis to locate the sites of oxidative damage, they could not analyze the distribution of structurally defined nucleobase lesions along DNA sequences and suffered from the high background of direct strand breaks induced by sugar oxidation.

The *p53* tumor suppressor gene is a major target for genetic damage in smoking-induced cancer *(223)*. The majority of the *p53* base substitution mutations observed in smoking-associated lung tumors are G → T transversions and G → A transversions frequently occurring at *p53* exons 5, 7, and 8 *(94;95)*. Prominent *p53* mutational "hotspots" include endogenously methylated ^{Me}CG dinucleotides within codons 154 (*G**G**C* → *G**T**C*), 157 (***G**TC* → ***T**TC*), 158 (*C**G**C* → *C**T**C*), 245 (***G**GC* → ***T**GC*), 248 (*C**G**G* → *C**T**G*), 249 (*A**G**G* → *A**T**G*), and 273 (*C**G**T* → *C**T**T*). One possible mechanism for the increased mutagenesis at ^{Me}CG sites within the *p53* gene involves targeted binding of metabolically activated tobacco carcinogens to these sequences *(96-99)*.

Oxidative DNA damage is a likely contributor to the pathogenesis of lung cancer *(101)*. Smokers excrete increased amounts of 8-oxo-dG in their urine, which is reversed upon smoking cessation *(228)*. Both 8-oxo-dG and oxazolone have been shown to cause G → T transversion mutations in both prokaryotic and eukaryotic cells *(102)*. Furthermore, cytosine methylation has been demonstrated to facilitate guanine oxidation at the base-paired position *(103)*. Therefore, we hypothesized that oxidative damage may contribute to G → T transversions observed at the lung cancer mutational hotspots within endogenously methylated CG dinuclotides at *p53* exons 5, 7,

and 8.

In current study, stable isotope labeling of DNA-mass spectrometry (ILD-MS) approach (Chapter V, Scheme 5.1) was employed to map the formation of 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxo-dG) and 2,2-diamino-4-[2-deoxy-*β*-D-*erythro*-pentofuranosyl)amino]-2,5-dihydrooxazol-5-one (oxazolone) lesions along DNA sequences derived from the *p53* tumor suppressor gene (Chapter V, Table 5.1). In each duplex, one of the guanine bases was labeled with [1,7,NH_2-$^{15}N_3$-2-^{13}C]-guanine which served as an isotope "tag" to enable specific quantification of guanine lesions originating from that position (Chapter V, Figures 5.1 and 5.2). Following photooxidation in the presence of riboflavin, DNA was enzymatically digested to 2'-deoxynucleosides, and the formation of 8-oxo-dG and oxazolone at each site of interest was quantified by isotope ratios obtained from capillary HPLC-ESI^+-MS/MS (Chapter V, Scheme 5.1).

Our results indicated that in double stranded DNA, both oxidative lesions were generated non-randomly, and their distribution was strongly influenced by the local DNA sequence (Chapter V, Figures 5.3-5.6). In particular, the 5' Gs in guanine repeats and guanines within ^{Me}CG dinucleotides were preferentially targeted for photooxidation in the presence of riboflavin. This can be explained by the low ionization potential of 5'-guanine at 5'-**G**G sites and/or the preferential intercalation of riboflavin at ^{Me}CG sites. While these results are in general consistent with those previously determined by the gel electrophoresis, our methodology avoids the high background of strand breaks resulting from 2'-deoxyribose oxidation and thus provides more details about the structural identities of both oxidative lesions. Furthermore, the most frequently adducted position, $\mathbf{G_5}$ in exon 5, $\mathbf{G_4}$ and $\mathbf{G_7}$ in exon7, and $\mathbf{G_6}$ in exon 8, coincide with the known *p53* lung cancer mutational "hotspots" at *p53* codons 158 (*C**G**C*), 245 (***G**GC*), 248 (*C**G**G*), and 273 (*C**G**T*), respectively, suggesting that oxidative DNA damage may contribute to mutagenesis in the *p53* gene.

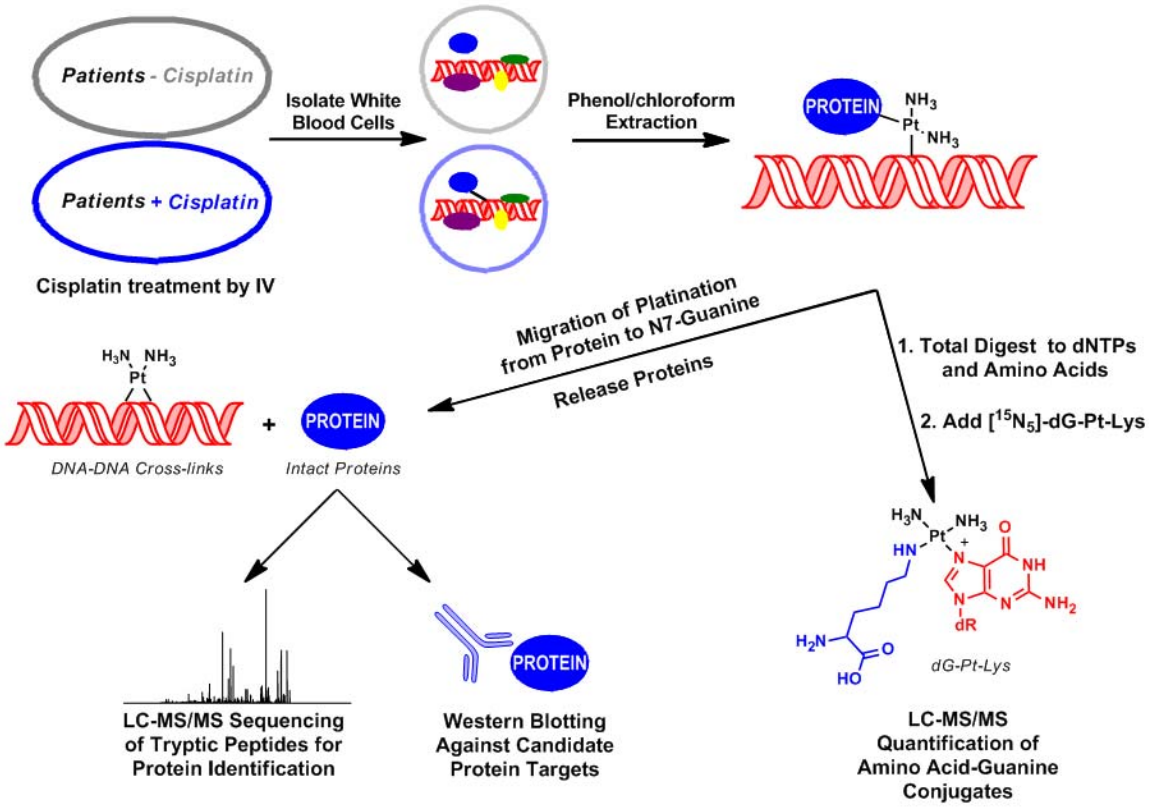
VII. FUTURE DIRECTIONS

7.1 Detection and Quantitation of DPCs in White Blood Cells of Cancer Patients Receiving Cisplatin Therapy

While the therapeutic efficacy of cisplatin is mainly attributed to its ability to form DNA-DNA cross-links, knowledge of the identities and the amounts of cisplatin-induced DPCs in the tissues of patients undergoing cisplatin therapy could help us to fully understand the origins of the on-target and off-target toxicity of this useful drug. Previous studies have shown that binding of certain nuclear proteins to cisplatin-induced DNA lesions was required for its biological effects *(172)*. For example, the preferential binding of HMG group proteins to cisplatin-induced DNA-DNA cross-links prevents their repair and sequesters transcriptional factors *(109;180)*. These and other proteins may also bind to cisplatin-DNA monoadducts, potentially leading to their irreversible trapping on DNA and the formation of covalent DNA-protein lesions. Although our results presented in this Thesis demonstrate that cisplatin forms DPCs *in vitro* and mammalian cell culture (*see Chapters II-IV*), questions still remain whether similar cross-linking occurs in cancer patients treated with cisplatin *(129)*.

We propose to employ our modified phenol/chloroform DNA extraction strategy (*see Chapter IV*) to isolate cisplatin-induced DPCs from white blood cells of cancer patients receiving cisplatin chemotherapy (Scheme 7.1). These cross-linked proteins can then be identified using mass spectrometry-based proteomics as described in *Chapter IV*. Furthermore, quantitative isotope dilution HPLC-ESI-MS/MS methodology can be developed for quantifying cisplatin-induced DPCs lesions patients' blood (Scheme 7.1). ^{15}N-labeled dG-Pt-Lys can be used as internal standard to investigate the formation and persistence of DPC lesions in blood of cancer patients receiving cisplatin treatment.

Scheme 7.1 Experimental scheme for the quantitative analysis of cisplatin-induced DPCs in human blood.

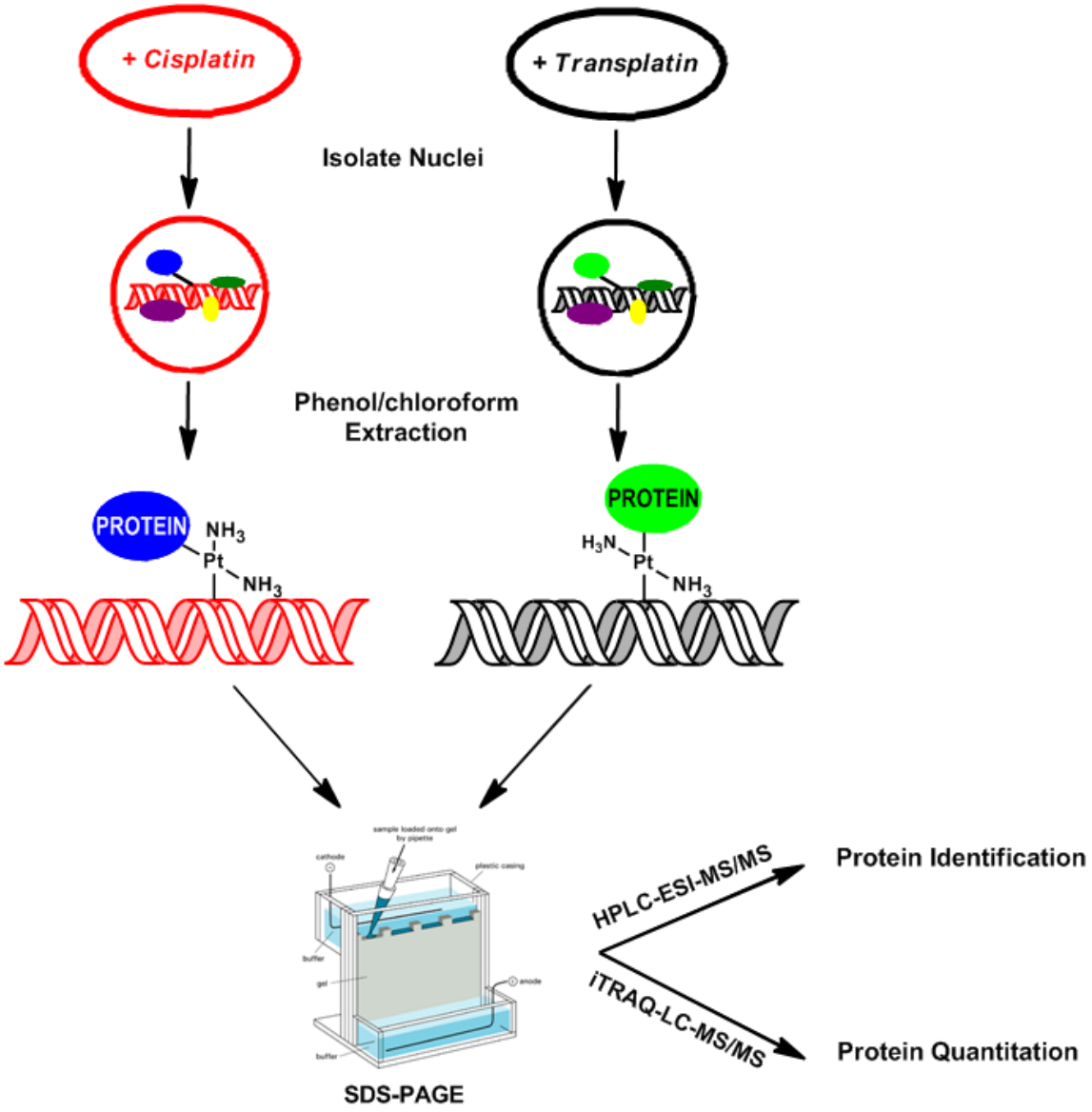


7.2 Comparative Proteomic Studies of DPCs Induced by Cisplatin and Transplatin in Mammalian Cells

DNA is the well-established target of platinum-based compounds in cells, and the resulting DNA adduct formation (i.e. 1,2-GG intrastrand cross-links) is considered to be important for platinum-mediated cytotoxicity *(234)*. The extent and the nature of these DNA adducts have been used to explain the differences in biological effects and clinical efficacies between cisplatin and transplatin (Figure 1.1). Cisplatin is capable of forming a variety of DNA adducts, including 1,2-GG intrastrand cross-links (major), 1,2-AG intrastrand cross-links, 1,3-GXG intrastrand cross-links, and some minor 1,2-GG interstrand cross-links (minor). In contrast, the genomic constraints of transplatin prevent the formation of 1,2-GG intrastrand cross-link, which is considered the basis of its lower antitumor activity *(234)*.

Recently, Chvalova and co-workers used a gel shift assay to compare the abilities of cisplatin and tranplatin to form DPCs *(110)*. The amount of radioactivity associated with DPCs bands induced by cisplatin was remarkably higher than that by transplatin, suggesting that cisplatin exhibits a considerably higher efficiency to form DPCs than transplatin *(110)*. This raises a possibility that transplatin is clinically ineffective due to its relative inability to form DPCs to specific proteins. However, the identities of these protein targets have not been established. The ultimate goal of such studies is to contribute to the development of more efficient and selective platinum-based therapies. To identify the proteins participating in DPC formation by cisplatin or transplatin in mammalian cells, the experimental strategy described in *Chapter III* can be used (Scheme 7.2). HT1080 human fibrosarcoma cells will be treated with 10 μM concentrations of cisplatin or 10 μM transplatin under physiological conditions. Chromosomal DNA containing covalently cross-linked proteins will be isolated by phenol/chloroform extraction as described

Scheme 7.2 Experimental scheme for comparative proteomic analysis.



above *(177)*. Proteins will be released by heating and subjected to SDS-PAGE analysis, serving as a visual comparison among two sets of cross-linked proteins. Mass spectrometry-based proteomic analysis will provide detailed information about the identities of cross-linked proteins among cisplatin-treated and transplatin-treated samples. Finally, the use of 4- or 8-plex iTRAQ reagents will enable a relative quantitation of DPC formation by cisplatin and transplatin, respectively *(235)*.

7.3 Examine the Biological Effects of DPC: Replication, Transcription, and Repair

As mentioned above, DPC lesions are expected to interrupt normal DNA-protein interactions, which are critical to cellular processes and DNA metabolism, such as DNA replication and transcription *(1)*. However, the contribution of DPCs to the observed cellular effects of DPC-inducing agents is poorly understood, mainly because these agents can form many other types of DNA damage. The fact that DPCs only constitute a small portion (less than 3%) of total DNA adducts following exposure to *bis*-electrophiles also makes it difficult to investigate cellular repair pathways responsible for their removal *(1)*. Therefore, it is necessary to develop methodology for selectively introducing site-specific DPC-containing DNA both *in vitro* and *in vivo* in order to investigate their repair and biological effects in the absence of other types of DNA damage.

Our previous studies have characterized AGT-DNA cross-linking reaction in the presence of cisplatin (*see Chapter II*), providing a possible way to build site-specifically modified DNA repair substrates. Synthetic oligodeoxyribonucleotides containing a single guanine in their sequence will be reacted with 1 molar equivalent of cisplatin to form DNA monoadducts and isolated by reversed phase HPLC. The resulting oligomers containing N7-guanine cisplatin

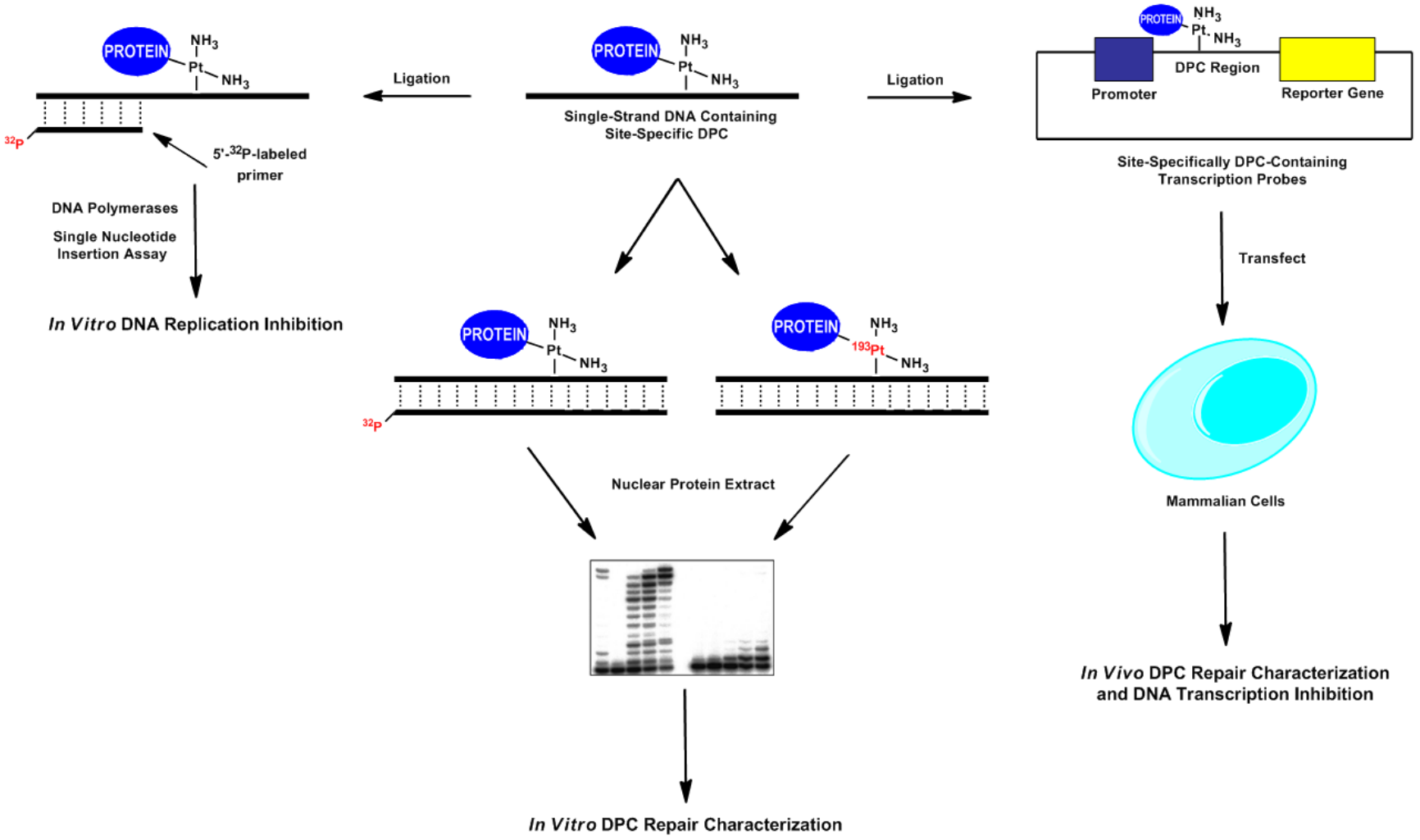
monoadduct will be incubated with target proteins (AGT, GAPDH, HMG, histone, or PARP), leading to the formation of site-specific DPC adducts (Scheme 7.3).

DPC-containing single-stranded oligomers will be annealed with radioactively labeled primers, followed by primer extension and single nucleotide insertion assay in the presence of all four dNTPs and polymerases (hPol η, hPol κ and hPol ι) *(236)*. The extension products will be separated by gel electrophoresis and then visualized using a phosphorImager. DPC lesions will stall replicative polymerases or will be bypassed in an error-prone or error-free lesion, providing insight into how DPCs affect DNA replication *in vitro* (Scheme 7.3).

To examine DPC repair *in vitro*, synthetic DNA duplexes containing site-specific cisplatin DPC lesions will be either ^{32}P- or ^{193}Pt-radiolabeled, and then incubated with nuclear protein extract from Hela cells. The resulting excision products will be analyzed on gel electrophoresis. Examination of radioactive bands resulting from repair reactions on the DPC substrate will help us identify mammalian DNA repair pathways involved in DPC repair since different repair products can be expected depending on different mechanisms (*e.g.* NER, HR, or proteolytic degradation). Alternatively, a combination of ^{193}Pt-labeling, SDS-PAGE, and mass spectrometry-based proteomics can be employed to identify specific proteins that participate in DPC damage recognition and/or repair (Scheme 7.3).

To examine DPC repair in mammalian cell culture, either ^{32}P- or ^{193}Pt-labeled DPC containing DNA can be ligated into a circular DNA plasmid, and the entire vector will be transfected into cells *(237)*. Then their *in vivo* repair can be similarly assessed by SDS-PAGE and mass spectrometry-based proteomics. Furthermore, the possible transcription inhibition can be examined based on the use of DPC-containing DNA plasmids constructed from a recombinant mammalian expression vector containing a reporter gene under the control of promoter for

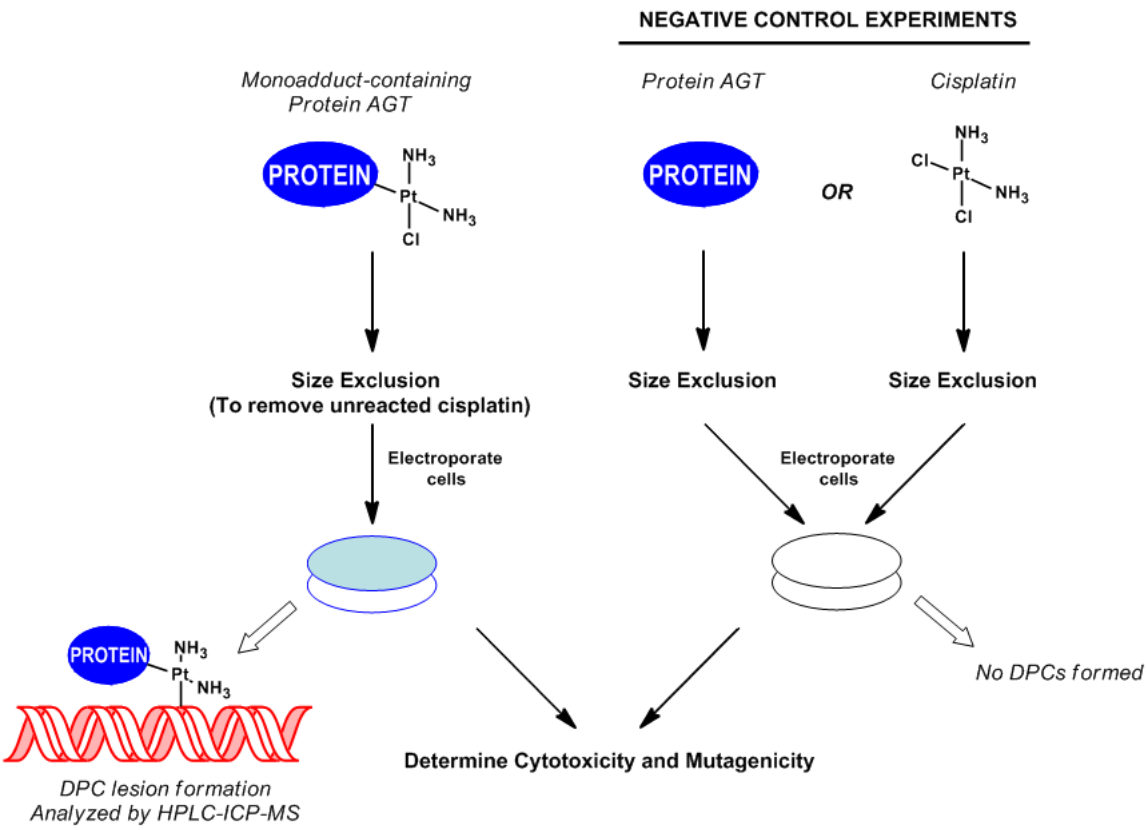
Scheme 7.3 Scheme for constructing site-specifically DPC-containing DNA for investigating its biological effects of DPC.



subsequent transfection into mammalian cells. The expression of such a reporter gene can serve as a transcription probe under specific DPC-containing lesions (Scheme 7.3) *(237)*.

Alternatively, AGT proteins can be incubated with cisplatin to generate AGT-monoadducts. After removing any unreacted cisplatin by size exclusion, these protein reagents will be electroporated into cultured mammalian cells, where they can further react with chromosomal DNA to form DPCs (Scheme 7.4), and the resulting cytotoxicity and mutagenicity will be examined. DPCs formation can be quantified by HPLC inductively coupled plasma mass spectrometry (HPLC-ICP-MS) analysis of chromosomal DNA. Importantly, unlike cisplatin and other *bis*-electrophiles, these protein monoadduct reagents are unable to produce other types of DNA damage which could contribute to cytotoxicity and mutagenicity (Scheme 7.4). This approach enables the analysis of biological outcomes associated specifically with DPCs.

Scheme 7.4 Strategy for the selective induction of DPCs in mammalian cells.



VIII. Bibliography

(1) Barker, S., Weinfeld, M., and Murray, D. (2005) DNA-protein cross-links: their induction, repair, and biological consequences. *Mutat. Res. 589*(2), 111-135.

(2) Reardon, J. T., Cheng, Y., and Sancar, A. (2006) Repair of DNA-protein cross-links in mammalian cells. *Cell Cycle 5*(13), 1366-1370.

(3) Ide, H., Shoulkamy, M. I., Nakano, T., Miyamoto-Matsubara, M., and Salem, A. M. (2010) Repair and biochemical effects of DNA-protein crosslinks. *Mutat. Res.*

(4) Yang, S. W., Burgin, A. B., Jr., Huizenga, B. N., Robertson, C. A., Yao, K. C., and Nash, H. A. (1996) A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. *Proc. Natl. Acad. Sci. U. S. A 93*(21), 11534-11539.

(5) Subramanian, D., Rosenstein, B. S., and Muller, M. T. (1998) Ultraviolet-induced DNA damage stimulates topoisomerase I-DNA complex formation in vivo: possible relationship with DNA repair. *Cancer Res. 58*(5), 976-984.

(6) Neale, M. J., Pan, J., and Keeney, S. (2005) Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature 436*(7053), 1053-1057.

(7) Hashimoto, M., Greenberg, M. M., Kow, Y. W., Hwang, J. T., and Cunningham, R. P. (2001) The 2-deoxyribonolactone lesion produced in DNA by neocarzinostatin and other damaging agents forms cross-links with the base-excision repair enzyme endonuclease III. *J. Am. Chem. Soc. 123*(13), 3161-3162.

(8) Ma, T. H., and Harris, M. M. (1988) Review of the genotoxicity of formaldehyde. *Mutat. Res. 196*(1), 37-59.

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

(10) Heck, H. D., Casanova, M., and Starr, T. B. (1990) Formaldehyde toxicity--new understanding. *Crit Rev. Toxicol. 20*(6), 397-426.

(11) Merk, O., and Speit, G. (1998) Significance of formaldehyde-induced DNA-protein crosslinks for mutagenesis. *Environ. Mol. Mutagen. 32*(3), 260-268.

(12) McGhee, J. D., and von Hippel, P. H. (1975) Formaldehyde as a probe of DNA structure. I. Reaction with exocyclic amino groups of DNA bases. *Biochemistry 14*(6), 1281-1296.

(13) McGhee, J. D., and von Hippel, P. H. (1975) Formaldehyde as a probe of DNA structure. II. Reaction with endocyclic imino groups of DNA bases. *Biochemistry 14*(6), 1297-1303.

(14) Casanova, M., Morgan, K. T., Steinhagen, W. H., Everitt, J. I., Popp, J. A., and Heck, H. D. (1991) Covalent binding of inhaled formaldehyde to DNA in the respiratory tract of rhesus monkeys: pharmacokinetics, rat-to-monkey interspecies scaling, and extrapolation to man. *Fundam. Appl. Toxicol. 17*(2), 409-428.

(15) Casanova, M., Morgan, K. T., Gross, E. A., Moss, O. R., and Heck, H. A. (1994) DNA-protein cross-links and cell replication at specific sites in the nose of F344 rats exposed subchronically to formaldehyde. *Fundam. Appl. Toxicol. 23*(4), 525-536.

(16) Shaham, J., Bomstein, Y., Meltzer, A., Kaufman, Z., Palma, E., and Ribak, J. (1996) DNA--protein crosslinks, a biomarker of exposure to formaldehyde--in vitro and in vivo studies. *Carcinogenesis 17*(1), 121-125.

(17) O'Connor, P. M., and Fox, B. W. (1989) Isolation and characterization of proteins cross-linked to DNA by the antitumor agent methylene dimethanesulfonate and its hydrolytic product formaldehyde. *J. Biol. Chem. 264*(11), 6391-6397.

(18) Hubal, E. A., Schlosser, P. M., Conolly, R. B., and Kimbell, J. S. (1997) Comparison of inhaled formaldehyde dosimetry predictions with DNA-protein cross-link measurements in the rat nasal passages. *Toxicol. Appl. Pharmacol. 143*(1), 47-55.

(19) Salnikow, K., and Zhitkovich, A. (2008) Genetic and epigenetic mechanisms in metal carcinogenesis and cocarcinogenesis: nickel, arsenic, and chromium. *Chem. Res. Toxicol. 21*(1), 28-44.

(20) Mattagajasingh, S. N., Misra, B. R., and Misra, H. P. (2008) Carcinogenic chromium(VI)-induced protein oxidation and lipid peroxidation: implications in DNA-protein crosslinking. *J. Appl. Toxicol. 28*(8), 987-997.

(21) Mattagajasingh, S. N., and Misra, H. P. (1996) Mechanisms of the carcinogenic chromium(VI)-induced DNA-protein cross-linking and their characterization in cultured intact human cells. *J. Biol. Chem. 271*(52), 33550-33560.

(22) Zhitkovich, A., Voitkun, V., Kluz, T., and Costa, M. (1998) Utilization of DNA-protein cross-links as a biomarker of chromium exposure. *Environ. Health Perspect. 106 Suppl 4*, 969-974.

(23) Chakrabarti, S. K., Bai, C., and Subramanian, K. S. (2001) DNA-protein crosslinks induced by nickel compounds in isolated rat lymphocytes: role of reactive oxygen species and specific amino acids. *Toxicol. Appl. Pharmacol. 170*(3), 153-165.

(24) Barker, S., Weinfeld, M., Zheng, J., Li, L., and Murray, D. (2005) Identification of mammalian proteins cross-linked to DNA by ionizing radiation. *J. Biol. Chem. 280*(40), 33826-33838.

(25) Weir Lipton, M. S., Fuciarelli, A. F., Springer, D. L., and Edmonds, C. G. (1996) Characterization of radiation-induced thymine-tyrosine crosslinks by electrospray ionization mass spectrometry. *Radiat. Res. 145*(6), 681-686.

(26) Zhang, H., and Wheeler, K. T. (1993) Radiation-induced DNA damage in tumors and normal tissues. I. Feasibility of estimating the hypoxic fraction. *Radiat. Res. 136*(1), 77-88.

(27) Zhang, H., and Wheeler, K. T. (1994) Radiation-induced DNA damage in tumors and normal tissues. II. Influence of dose, residual DNA damage and physiological factors in oxygenated cells. *Radiat. Res. 140*(3), 321-326.

(28) Zhang, H., Koch, C. J., Wallen, C. A., and Wheeler, K. T. (1995) Radiation-induced DNA damage in tumors and normal tissues. III. Oxygen dependence of the formation of strand breaks and DNA-protein crosslinks. *Radiat. Res. 142*(2), 163-168.

(29) Johansen, M. E., Muller, J. G., Xu, X., and Burrows, C. J. (2005) Oxidatively induced DNA-protein cross-linking between single-stranded binding protein and oligodeoxynucleotides containing 8-oxo-7,8-dihydro-2'-deoxyguanosine. *Biochemistry 44*(15), 5660-5671.

(30) Xu, X., Muller, J. G., Ye, Y., and Burrows, C. J. (2008) DNA-protein cross-links between guanine and lysine depend on the mechanism of oxidation for formation of C5 vs C8 guanosine adducts. *J. Am. Chem. Soc. 130*(2), 703-709.

(31) Xu, X., Fleming, A. M., Muller, J. G., and Burrows, C. J. (2008) Formation of tricyclic [4.3.3.0] adducts between 8-oxoguanosine and tyrosine under conditions of oxidative DNA-protein cross-linking. *J. Am. Chem. Soc. 130*(31), 10080-10081.

(32) Zwelling, L. A., Anderson, T., and Kohn, K. W. (1979) DNA-protein and DNA interstrand cross-linking by *cis*-platinum(II) and *trans*-platinum(II) diamminedichloride in L1210 mouse leukemia-cells and relation to cytotoxicity. *Cancer Research 39*(2), 365-369.

(33) Masuda, K., Nakamura, T., Mizota, T., Mori, J., and Shimomura, K. (1988) Interstrand DNA-DNA and DNA-protein cross-links by a new antitumor antibiotic, FK973, in L1210 cells. *Cancer Res. 48*(18), 5172-5177.

(34) Ewig, R. A., and Kohn, K. W. (1978) DNA-protein cross-linking and DNA interstrand cross-linking by haloethylnitrosoureas in L1210 cells. *Cancer Res. 38*(10), 3197-3203.

(35) Michaelson-Richie, E. D., Loeber, R. L., Codreanu, S. G., Ming, X., Liebler, D. C., Campbell, C., and Tretyakova, N. Y. (2010) DNA-protein cross-linking by 1,2,3,4-diepoxybutane. *J. Proteome. Res. 9*(9), 4356-4367.

(36) Loeber, R. L., Michaelson-Richie, E. D., Codreanu, S. G., Liebler, D. C., Campbell, C. R., and Tretyakova, N. Y. (2009) Proteomic analysis of DNA-protein cross-linking by antitumor nitrogen mustards. *Chem. Res. Toxicol. 22*(6), 1151-1162.

(37) Cupo, D. Y., and Wetterhahn, K. E. (1985) Binding of chromium to chromatin and DNA from liver and kidney of rats treated with sodium dichromate and chromium(III) chloride in vivo. *Cancer Res. 45*(3), 1146-1151.

(38) Oleinick, N. L., Chiu, S. M., Ramakrishnan, N., and Xue, L. Y. (1987) The formation, identification, and significance of DNA-protein cross-links in mammalian cells. *Br. J. Cancer Suppl 8*, 135-140.

(39) Baker, D. J., Wuenschell, G., Xia, L., Termini, J., Bates, S. E., Riggs, A. D., and O'Connor, T. R. (2007) Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells. *J. Biol. Chem. 282*(31), 22592-22604.

(40) Wood, R. D. (1999) DNA damage recognition during nucleotide excision repair in mammalian cells. *Biochimie 81*(1-2), 39-44.

(41) Minko, I. G., Zou, Y., and Lloyd, R. S. (2002) Incision of DNA-protein crosslinks by UvrABC nuclease suggests a potential repair pathway involving nucleotide excision repair. *Proc. Natl. Acad. Sci. U. S. A 99*(4), 1905-1909.

(42) Minko, I. G., Kurtz, A. J., Croteau, D. L., Van Houten, B., Harris, T. M., and Lloyd, R. S. (2005) Initiation of repair of DNA-polypeptide cross-links by the UvrABC nuclease. *Biochemistry 44*(8), 3000-3009.

(43) Reardon, J. T., and Sancar, A. (2006) Repair of DNA-polypeptide crosslinks by human excision nuclease. *Proc. Natl. Acad. Sci. U. S. A 103*(11), 4056-4061.

(44) Thompson, L. H. (1996) Evidence that mammalian cells possess homologous recombinational repair pathways. *Mutat. Res. 363*(2), 77-88.

(45) Nakano, T., Morishita, S., Katafuchi, A., Matsubara, M., Horikawa, Y., Terato, H., Salem, A. M. H., Izumi, S., Pack, S. P., Makino, K., and Ide, H. (2007) Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNA-protein crosslinks. *Molecular Cell 28*(1), 147-158.

(46) Nakano, T., Katafuchi, A., Matsubara, M., Terato, H., Tsuboi, T., Masuda, T., Tatsumoto, T., Pack, S. P., Makino, K., Croteau, D. L., Van Houten, B., Iijima, K., Tauchi, H., and Ide, H. (2009) Homologous recombination but not nucleotide excision repair plays a pivotal role in tolerance of DNA-protein cross-links in mammalian cells. *J. Biol. Chem. 284*(40), 27065-27076.

(47) Zhitkovich, A., and Costa, M. (1992) A simple, sensitive assay to detect DNA-protein crosslinks in intact cells and in vivo. *Carcinogenesis 13*(8), 1485-1489.

(48) Chen, H. J., Chiu, W. L., Lin, W. P., and Yang, S. S. (2008) Investigation of DNA-protein cross-link formation between lysozyme and oxanine by mass spectrometry. *Chembiochem. 9*(7), 1074-1081.

(49) Wisniewski, J. R. (2008) Mass spectrometry-based proteomics: principles, perspectives, and challenges. *Arch. Pathol. Lab Med. 132*, 1566-1569.

(50) Jelitto, B., Vangala, R. R., and Laib, R. J. (1989) Species differences in DNA damage by butadiene: role of diepoxybutane. *Arch. Toxicol. Suppl 13*, 246-249.

(51) Ewig, R. A., and Kohn, K. W. (1977) DNA damage and repair in mouse leukemia L1210 cells treated with nitrogen mustard, 1,3-bis(2-chloroethyl)-1-nitrosourea, and other nitrosoureas. *Cancer Res. 37*(7 Pt 1), 2114-2122.

(52) Fornace, A. J., Jr., and Little, J. B. (1977) DNA crosslinking induced by x-rays and chemical agents. *Biochim. Biophys. Acta 477*(4), 343-355.

(53) Kasparkova, J., Novakova, O., Vrana, O., Intini, F., Natile, G., and Brabec, V. (2006) Molecular aspects of antitumor effects of a new platinum(IV) drug. *Mol. Pharmacol. 70*(5), 1708-1719.

(54) Loeber, R., Michaelson, E., Fang, Q., Campbell, C., Pegg, A. E., and Tretyakova, N. (2008) Cross-linking of the DNA repair protein Omicron6-alkylguanine DNA alkyltransferase to DNA in the presence of antitumor nitrogen mustards. *Chem. Res. Toxicol. 21*(4), 787-795.

(55) Loeber, R., Rajesh, M., Fang, Q., Pegg, A. E., and Tretyakova, N. (2006) Cross-linking of the human DNA repair protein O6-alkylguanine DNA alkyltransferase to DNA in the presence of 1,2,3,4-diepoxybutane. *Chem. Res. Toxicol. 19*(5), 645-654.

(56) Kohn, K. W. (1981) DNA damage in mammalian cells. *BioScience 31*, 593-597.

(57) Thomas, C. B., Kohn, K. W., and Bonner, W. M. (1978) Characterization of DNA-protein cross-links formed by treatment of L1210 cells and nuclei with *bis*(2-chloroethyl)methylamine (nitrogen mustard). *Biochemistry 17*(19), 3954-3958.

(58) Ostling, O., and Johanson, K. J. (1984) Microelectrophoretic study of radiation-induced DNA damages in individual mammalian cells. *Biochem. Biophys. Res. Commun. 123*(1), 291-298.

(59) Merk, O., Reiser, K., and Speit, G. (2000) Analysis of chromate-induced DNA-protein crosslinks with the comet assay. *Mutat. Res. 471*(1-2), 71-80.

(60) Costa, M., Zhitkovich, A., Harris, M., Paustenbach, D., and Gargas, M. (1997) DNA-protein cross-links produced by various chemicals in cultured human lymphoma cells. *J. Toxicol. Environ. Health 50*(5), 433-449.

(61) Goggin, M., Loeber, R., Park, S., Walker, V., Wickliffe, J., and Tretyakova, N. (2007) HPLC-ESI+-MS/MS analysis of N7-guanine-N7-guanine DNA cross-links in tissues of mice exposed to 1,3-butadiene. *Chem. Res. Toxicol. 20*(5), 839-847.

(62) Goggin, M., Anderson, C., Park, S., Swenberg, J., Walker, V., and Tretyakova, N. (2008) Quantitative high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry analysis of the adenine-guanine cross-links of 1,2,3,4-diepoxybutane in tissues of butadiene-exposed B6C3F1 mice. *Chem. Res. Toxicol. 21*(5), 1163-1170.

(63) Seneviratne, U., Antsypovich, S., Goggin, M., Dorr, D. Q., Guza, R., Moser, A., Thompson, C., York, D. M., and Tretyakova, N. (2010) Exocyclic deoxyadenosine adducts of 1,2,3,4-diepoxybutane: synthesis, structural elucidation, and mechanistic studies. *Chem. Res. Toxicol. 23*(1), 118-133.

(64) Doneanu, C. E., Gafken, P. R., Bennett, S. E., and Barofsky, D. F. (2004) Mass spectrometry of UV-cross-linked protein-nucleic acid complexes: identification of

amino acid residues in the single-stranded DNA-binding domain of human replication protein A. *Anal. Chem. 76*(19), 5667-5676.

(65) Bochkarev, A., Pfuetzner, R. A., Edwards, A. M., and Frappier, L. (1997) Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature 385*(6612), 176-181.

(66) Kelleher, N. L. (2004) Top-down proteomics. *Analytical Chemistry 76*(11), 196A-203A.

(67) Armirotti, A., and Damonte, G. (2010) Achievements and perspectives of top-down proteomics. *Proteomics 10*(20), 3566-3576.

(68) Rosenberg, B. (1985) Fundamental studies with cisplatin. *Cancer 55*(10), 2303-23l6.

(69) Brabec, V. (2002) DNA modifications by antitumor platinum and ruthenium compounds: their recognition and repair. *Prog. Nucleic Acid Res. Mol. Biol. 71*, 1-68.

(70) Eastman, A. (1987) The formation, isolation and characterization of DNA adducts produced by anticancer platinum complexes. *Pharmacol. Ther. 34*(2), 155-166.

(71) Lippard, S. J., and Hoeschele, J. D. (1979) Binding of *cis*- and *trans*-dichlorodiammineplatinum(II) to the nucleosome core. *Proc. Natl. Acad. Sci. U. S. A 76*(12), 6091-6095.

(72) Banjar, Z. M., Hnilica, L. S., Briggs, R. C., Stein, J., and Stein, G. (1984) *Cis*- and *trans*-diamminedichloroplatinum(II)-mediated cross-linking of chromosomal non-histone proteins to DNA in HeLa cells. *Biochemistry 23*(9), 1921-1926.

(73) Olinski, R., Wedrychowski, A., Schmidt, W. N., Briggs, R. C., and Hnilica, L. S. (1987) *In vivo* DNA-protein cross-linking by *cis*- and *trans*-diamminedichloroplatinum(II). *Cancer Res. 47*(1), 201-205.

(74) Yamamoto, J., Miyagi, Y., Kawanishi, K., Yamada, S., Miyagi, Y., Kodama, J., Yoshinouchi, M., and Kudo, T. (1999) Effect of cisplatin on cell death and DNA crosslinking in rat mammary adenocarcinoma *in vitro*. *Acta Med. Okayama 53*(5), 201-208.

(75) Valko, M., Leibfritz, D., Moncol, J., Cronin, M. T. D., Mazur, M., and Telser, J. (2007) Free radicals and antioxidants in normal physiological functions and human disease. *International Journal of Biochemistry & Cell Biology 39*(1), 44-84.

(76) Ames, B. N., Shigenaga, M. K., and Hagen, T. M. (1993) Oxidants, Antioxidants, and the Degenerative Diseases of Aging. *Proceedings of the National Academy of Sciences of the United States of America 90*(17), 7915-7922.

(77) Jaruga, P., and Dizdaroglu, M. (1996) Repair of products of oxidative DNA base damage in human cells. *Nucleic Acids Research 24*(8), 1389-1394.

(78) Richter, C. (1995) Oxidative damage to mitochondrial-DNA and its relationship to aging. *International Journal of Biochemistry & Cell Biology 27*(7), 647-653.

(79) Seidel, C. A. M., Schulz, A., and Sauer, M. H. M. (1996) Nucleobase-specific quenching of fluorescent dyes .1. Nucleobase one-electron redox potentials and their correlation with static and dynamic quenching efficiencies. *Journal of Physical Chemistry 100*(13), 5541-5553.

(80) Steenken, S., and Jovanovic, S. V. (1997) How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution. *Journal of the American Chemical Society 119*(3), 617-618.

(81) Pratviel, G., and Meunier, B. (2006) Guanine oxidation: One- and two-electron reactions. *Chemistry-A European Journal 12*(23), 6018-6030.

(82) Neeley, W. L., and Essigmann, J. M. (2006) Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chemical Research in Toxicology 19*(4), 491-505.

(83) Cadet, J., Douki, T., and Ravanat, J. L. (2006) One-electron oxidation of DNA and inflammation processes. *Nature Chemical Biology 2*(7), 348-349.

(84) Hall, D. B., Holmlin, R. E., and Barton, J. K. (1996) Oxidative DNA damage through long-range electron transfer. *Nature 382*(6593), 731-735.

(85) Giese, B. (2002) Long-distance electron transfer through DNA. *Annual Review of Biochemistry 71*, 51-70.

(86) Saito, I., Nakamura, T., Nakatani, K., Yoshioka, Y., Yamaguchi, K., and Sugiyama, H. (1998) Mapping of the hot spots for DNA damage by one-electron oxidation: Efficacy of GG doublets and GGG triplets as a trap in long-range hole migration. *Journal of the American Chemical Society 120*(48), 12686-12687.

(87) Margolin, Y., Cloutier, J. F., Shafirovich, V., Geacintov, N. E., and Dedon, P. C. (2006) Paradoxical hotspots for guanine oxidation by a chemical mediator of inflammation. *Nature Chemical Biology 2*(7), 365-366.

(88) Margolin, Y., Shafirovich, V., Geacintov, N. E., Demott, M. S., and Dedon, P. C. (2008) DNA sequence context as a determinant of the quantity and chemistry of guanine oxidation produced by hydroxyl radicals and one-electron oxidants. *Journal of Biological Chemistry 283*(51), 35569-35578.

(89) Gazdar, A. F., Bader, S., Hung, J., Kishimoto, Y., Sekido, Y., Sugio, K., Virmani, A., Fleming, J., Carbone, D. P., and Minna, J. D. (1994) Molecular-genetic changes found in human lung-cancer and its precursor lesions. *Cold Spring Harbor Symposia on Quantitative Biology 59*, 565-572.

(90) Zochbauer-Muller, S., Gazdar, A. F., and Minna, J. D. (2002) Molecular pathogenesis of lung cancer. *Annual Review of Physiology 64*, 681-708.

(91) Harris, C. C. (1996) Structure and function of the *p53* tumor suppressor gene: Clues for rational cancer therapeutic strategies. *Journal of the National Cancer Institute 88*(20), 1442-1455.

(92) Greenblatt, M. S., Bennett, W. P., Hollstein, M., and Harris, C. C. (1994) Mutations in the *p53* tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Research 54*(18), 4855-4878.

(93) Harris, C. C. (1996) *p53* tumor suppressor gene: at the crossroads of molecular carcinogenesis, molecular epidemiology, and cancer risk assessment. *Environmental Health Perspectives 104*, 435-439.

(94) Hussain, S. P., and Harris, C. C. (1999) *p53* mutation spectrum and load: the generation of hypotheses linking the exposure of endogenous or exogenous carcinogens

to human cancer. *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis 428*(1-2), 23-32.

(95) Pfeifer, G. P., Denissenko, M. F., Olivier, M., Tretyakova, N., Hecht, S. S., and Hainaut, P. (2002) Tobacco smoke carcinogens, DNA damage and *p53* mutations in smoking-associated cancers. *Oncogene 21*(48), 7435-7451.

(96) Matter, B., Wang, G., Jones, R., and Tretyakova, N. (2004) Formation of diastereomeric benzo[a]pyrene diol epoxide-guanine adducts in *p53* gene-derived DNA sequences. *Chem. Res. Toxicol. 17*(6), 731-741.

(97) Denissenko, M. F., Pao, A., Tang, M., and Pfeifer, G. P. (1996) Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in *p53*. *Science 274*(5286), 430-432.

(98) Denissenko, M. F., Chen, J. X., Tang, M. S., and Pfeifer, G. P. (1997) Cytosine methylation determines hot spots of DNA damage in the human *p53* gene. *Proc. Natl. Acad. Sci. U. S. A 94*(8), 3893-3898.

(99) Guza, R., Ma, L., Fang, Q., Pegg, A. E., and Tretyakova, N. (2009) Cytosine methylation effects on the repair of O^6-methylguanines within CG dinucleotides. *J. Biol. Chem. 284*(34), 22601-22610.

(100) Guza, R., Kotandeniya, D., Murphy, K., Dissanayake, T., Lin, C., Giambasu, G. M., Lad, R. R., Wojciechowski, F., Amin, S., Sturla, S. J., Hudson, R. H., York, D. M., Jankowiak, R., Jones, R., and Tretyakova, N. Y. (2011) Influence of C-5 substituted cytosine and related nucleoside analogs on the formation of benzo[a]pyrene diol epoxide-dG adducts at CG base pairs of DNA. *Nucleic Acids Res.*

(101) Gackowski, D., Speina, E., Zielinska, M., Kowalewski, J., Rozalski, R., Siomek, A., Paciorek, T., Tudek, B., and Olinski, R. (2003) Products of oxidative DNA damage and repair as possible biomarkers of susceptibility to lung cancer. *Cancer Research 63*(16), 4899-4902.

(102) Henderson, P. T., Delaney, J. C., Gu, F., Tannenbaum, S. R., and Essigmann, J. M. (2002) Oxidation of 7,8-dihydro-8-oxoguanine affords lesions that are potent sources of replication errors *in vivo*. *Biochemistry 41*(3), 914-921.

(103) Kawai, K., Wata, Y., Hara, M., Tojo, S., and Majima, T. (2002) Regulation of one-electron oxidation rate of guanine by base pairing with cytosine derivatives. *Journal of the American Chemical Society 124*(14), 3586-3590.

(104) Bain, G., Maandag, E. C. R., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Krop, I., Schlissel, M. S., Feeney, A. J., Vanroon, M., Vandervalk, M., Teriele, H. P. J., Berns, A., and Murre, C. (1994) E2A proteins are required for proper B-cell development and initiation of immunoglobulin gene rearrangements. *Cell 79*(5), 885-892.

(105) Dynlacht, B. D. (1997) Regulation of transcription by proteins that control the cell cycle. *Nature 389*(6647), 149-152.

(106) Accili, D., and Arden, K. C. (2004) FoxOs at the crossroads of cellular metabolism, differentiation, and transformation. *Cell 117*(4), 421-426.

(107) Kloster, M., Kostrhunova, H., Zaludova, R., Malina, J., Kasparkova, J., Brabec, V., and Farrell, N. (2004) Trifunctional dinuclear platinum complexes as DNA-protein cross-linking agents. *Biochemistry 43*(24), 7776-7786.

(108) Baker, J. M., Parish, J. H., and Curtis, J. P. (1984) DNA-DNA and DNA-protein cross-linking and repair in Neurospora crassa following exposure to nitrogen mustard. *Mutat. Res. 132*(5-6), 171-179.

(109) Jamieson, E. R., and Lippard, S. J. (1999) Structure, recognition, and processing of cisplatin-DNA adducts. *Chemical Reviews 99*(9), 2467-2498.

(110) Chvalova, K., Brabec, V., and Kasparkova, J. (2007) Mechanism of the formation of DNA-protein cross-links by antitumor cisplatin. *Nucleic Acids Research 35*(6), 1812-1821.

(111) Jung, Y. W., and Lippard, S. J. (2007) Direct cellular responses to platinum-induced DNA damage. *Chemical Reviews 107*(5), 1387-1407.

(112) Sherman, S. E., and Lippard, S. J. (1987) Structural aspects of platinum anticancer drug-interactions with DNA. *Chemical Reviews 87*(5), 1153-1181.

(113) Zwelling, L. A., Bradley, M. O., Sharkey, N. A., Anderson, T., and Kohn, K. W. (1979) Mutagenicity, cytotoxicity and DNA crosslinking in V79 Chinese-hamster cells treated with *cis*-Pt(II) diamminedichloride and *trans*-Pt(II) diamminedichloride. *Mutation Research 67*(3), 271-280.

(114) Pegg, A. E. (2000) Repair of O^6-alkylguanine by alkyltransferases. *Mutat. Res. 462*(2-3), 83-100.

(115) Daniels, D. S., Woo, T. T., Luu, K. X., Noll, D. M., Clarke, N. D., Pegg, A. E., and Tainer, J. A. (2004) DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nature Structural & Molecular Biology 11*(8), 714-720.

(116) Guengerich, F. P. (2005) Principles of covalent binding of reactive metabolites and examples of activation of *bis*-electrophiles by conjugation. *Archives of Biochemistry and Biophysics 433*(2), 369-378.

(117) Liu, L. P., Hachey, D. L., Valadez, G., Williams, K. M., Guengerich, F. P., Loktionova, N. A., Kanugula, S., and Pegg, A. E. (2004) Characterization of a mutagenic DNA adduct formed from 1,2-dibromoethane by O^6-alkylguanine-DNA alkyltransferase. *Journal of Biological Chemistry 279*(6), 4250-4259.

(118) Valadez, J. G., Liu, L. P., Loktionova, N. A., Pegg, A. E., and Guengerich, F. P. (2004) Activation of *bis*-electrophiles to mutagenic conjugates by human O^6-alkylguanine-DNA alkyltransferase. *Chemical Research in Toxicology 17*(7), 972-982.

(119) Kalapila, A. G., Loktionova, N. A., and Pegg, A. E. (2008) Alkyltransferase-mediated toxicity of 1,3-butadiene diepoxide. *Chemical Research in Toxicology 21*(9), 1851-1861.

(120) Bender, K., Federwisch, M., Loggen, U., Nehls, P., and Rajewsky, M. F. (1996) Binding and repair of O^6-ethylguanine in double-stranded oligodeoxynucleotides by recombinant human O^6-alkylguanine-DNA alkyltransferase do not exhibit significant dependence on sequence context. *Nucleic Acids Research 24*(11), 2087-2094.

(121) Eastman, A. (1982) Separation and characterization of products resulting from the reaction of *cis*-diamminedichloroplatinum(II) with deoxyribonucleosides. *Biochemistry 21*(26), 6732-6736.

(122) Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual* (1989) Cold Spring Harbor Press.

(123) Yates, J. R., III, Eng, J. K., and McCormack, A. L. (1995) Mining genomes: correlating tandem mass spectra of modified and unmodified peptides to sequences in nucleotide databases. *Anal. Chem. 67*(18), 3202-3210.

(124) Yates, J. R., III, Eng, J. K., McCormack, A. L., and Schieltz, D. (1995) Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal. Chem. 67*(8), 1426-1436.

(125) Daniels, D. S., Mol, C. D., Arvai, A. S., Kanugula, S., Pegg, A. E., and Tainer, J. A. (2000) Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. *Embo Journal 19*(7), 1719-1730.

(126) Vanboom, S. S. G. E., and Reedijk, J. (1993) Unprecedented migration of $[Pt(dien)]^{2+}$ (dien = 1,5-diamino-3-azapentane) from sulfur to guanosine-N7 in *S*-guanosyl-L-homocysteine (sgh). *Journal of the Chemical Society-Chemical Communications*(18), 1397-1398.

(127) Barnham, K. J., Djuran, M. I., Murdoch, P. D., and Sadler, P. J. (1994) Intermolecular displacement of *S*-bound L-methionine on platinum(II) by guanosine 5'-monophosphate - Implications for the mechanism of action of anticancer drugs. *Journal of the Chemical Society-Chemical Communications*(6), 721-722.

(128) Teuben, J. M., and Reedijk, J. (2000) Reaction of DNA oligonucleotides with [Pt(dien)GSMe](2+) (GSMe = S-methylated glutathione) and cis-[Pt(NH3)(2)(GSMe)2](2+): evidence of oligonucleotide platination *via* sulfur-coordinated platinum intermediates. *Journal of Biological Inorganic Chemistry 5*(4), 463-468.

(129) Reedijk, J. (1999) Why does cisplatin reach guanine-N7 with competing *S*-donor ligands available in the cell? *Chemical Reviews 99*(9), 2499-2510.

(130) Zhao, T., and King, F. L. (2009) Direct determination of the primary binding site of cisplatin on cytochrome C by mass spectrometry. *J. Am. Soc. Mass Spectrom. 20*(6), 1141-1147.

(131) Zhao, T., and King, F. L. (2010) A mass spectrometric comparison of the interactions of cisplatin and transplatin with myoglobin. *J. Inorg. Biochem. 104*(2), 186-192.

(132) Moreno-Gordaliza, E., Canas, B., Palacios, M. A., and Gomez-Gomez, M. M. (2009) Top-down mass spectrometric approach for the full characterization of insulin-cisplatin adducts. *Analytical Chemistry 81*(9), 3507-3516.

(133) Moreno-Gordaliza, E., Canas, B., Palacios, M. A., and Gomez-Gomez, M. M. (2010) Novel insights into the bottom-up mass spectrometry proteomics approach for the characterization of Pt-binding proteins: The insulin-cisplatin case study. *Analyst 135*(6), 1288-1298.

(134) Zhao, T., and King, F. L. (2011) Mass-spectrometric characterization of cisplatin binding sites on native and denatured ubiquitin. *J. Biol. Inorg. Chem.*

(135) Will, J., Wolters, D. A., and Sheldrick, W. S. (2008) Characterisation of cisplatin binding sites in human serum proteins using hyphenated multidimensional liquid chromatography and ESI tandem mass spectrometry. *Chemmedchem 3*(11), 1696-1707.

(136) Tubbs, J. L., Pegg, A. E., and Tainer, J. A. (2007) DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O^6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *Dna Repair 6*(8), 1100-1115.

(137) Deubel, D. V. (2002) On the competition of the purine bases, functionalities of peptide side chains, and protecting agents for the coordination sites of dicationic cisplatin derivatives. *Journal of the American Chemical Society 124*(20), 5834-5842.

(138) Pearson, R. G. (1987) Recent advances in the concept of hard and soft acids and bases. *Journal of Chemical Education 64*(7), 561-567.

(139) Jessberger, R., and Berg, P. (1991) Repair of deletions and double-strand gaps by homologous recombination in a mammalian *in vitro* system. *Mol. Cell Biol. 11*(1), 445-457.

(140) Shin, N. Y., Liu, Q., Stamer, S. L., and Liebler, D. C. (2007) Protein targets of reactive electrophiles in human liver microsomes. *Chem. Res. Toxicol. 20*(6), 859-867.

(141) Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem. 74*(20), 5383-5392.

(142) Zhang, B., Chambers, M. C., and Tabb, D. L. (2007) Proteomic parsimony through bipartite graph analysis improves accuracy and transparency. *J. Proteome. Res. 6*(9), 3549-3557.

(143) Eastman, A. (1983) Characterization of the adducts produced in DNA by *cis*-diamminedichloroplatinum(II) and *cis*-dichloro(ethylenediamine)platinum(II). *Biochemistry 22*(16), 3927-3933.

(144) Royer-Pokora, B., Gordon, L. K., and Haseltine, W. A. (1981) Use of exonuclease III to determine the site of stable lesions in defined sequences of DNA: the cyclobutane pyrimidine dimer and *cis*- and *trans*- dichlorodiammine platinum II examples. *Nucleic Acids Res. 9*(18), 4595-4609.

(145) Johnson, N. P., Macquet, J. P., Wiebers, J. L., and Monsarrat, B. (1982) Structures of the adducts formed between [Pt(dien)Cl]Cl and DNA *in vitro*. *Nucleic Acids Res. 10*(17), 5255-5271.

(146) Jagtap, P., and Szabo, C. (2005) Poly(ADP-ribose) polymerase and the therapeutic effects of its inhibitors. *Nat. Rev. Drug Discov. 4*(5), 421-440.

(147) Gao, C., Guo, H., Mi, Z., Wai, P. Y., and Kuo, P. C. (2005) Transcriptional regulatory functions of heterogeneous nuclear ribonucleoprotein-U and -A/B in endotoxin-mediated macrophage expression of osteopontin. *J. Immunol. 175*(1), 523-530.

(148) Sirover, M. A. (2005) New nuclear functions of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in mammalian cells. *J. Cell Biochem. 95*(1), 45-52.

(149) Xanthoudakis, S., and Curran, T. (1996) Redox regulation of AP-1: a link between transcription factor signaling and DNA repair. *Adv. Exp. Med. Biol. 387*, 69-75.

(150) Coulombe, R. A., Jr., Drew, G. L., and Stermitz, F. R. (1999) Pyrrolizidine alkaloids crosslink DNA with actin. *Toxicol. Appl. Pharmacol. 154*(2), 198-202.

(151) Burgers, P. M. (2009) Polymerase dynamics at the eukaryotic DNA replication fork. *J. Biol. Chem. 284*(7), 4041-4045.

(152) Nguewa, P. A., Fuertes, M. A., Valladares, B., Alonso, C., and Perez, J. M. (2005) Poly(ADP-ribose) polymerases: homology, structural domains and functions. Novel therapeutical applications. *Prog. Biophys. Mol. Biol. 88*(1), 143-172.

(153) Finger, L. D., Blanchard, M. S., Theimer, C. A., Sengerova, B., Singh, P., Chavez, V., Liu, F., Grasby, J. A., and Shen, B. (2009) The 3'-flap pocket of human flap endonuclease 1 is critical for substrate binding and catalysis. *J. Biol. Chem. 284*(33), 22184-22194.

(154) Mukherjee, S., Law, S. M., and Feig, M. (2009) Deciphering the mismatch recognition cycle in MutS and MSH2-MSH6 using normal-mode analysis. *Biophys. J. 96*(5), 1707-1720.

(155) Coghlin, C., Carpenter, B., Dundas, S. R., Lawrie, L. C., Telfer, C., and Murray, G. I. (2006) Characterization and over-expression of chaperonin t-complex proteins in colorectal cancer. *J. Pathol. 210*(3), 351-357.

(156) McGinnis, E., and Williams, L. S. (1972) Role of histidine transfer ribonucleic acid in regulation of synthesis of histidyl-transfer ribonucleic acid synthetase of Salmonella typhimurium. *J. Bacteriol. 109*(2), 505-511.

(157) Fountoulakis, M., Tsangaris, G., Oh, J. E., Maris, A., and Lubec, G. (2004) Protein profile of the HeLa cell line. *Journal of Chromatography A 1038*(1-2), 247-265.

(158) Zlatanova, J., Yaneva, J., and Leuba, S. H. (1998) Proteins that specifically recognize cisplatin-damaged DNA: a clue to anticancer activity of cisplatin. *FASEB J. 12*(10), 791-799.

(159) Scovell, W. M., Muirhead, N., and Kroos, L. R. (1987) *Cis*-diamminedichloroplatinum(II) selectively cross-links high mobility group proteins 1 and 2 to DNA in micrococcal nuclease accessible regions of chromatin. *Biochem. Biophys. Res. Commun. 142*(3), 826-835.

(160) Zhang, C. X., Chang, P. V., and Lippard, S. J. (2004) Identification of nuclear proteins that interact with platinum-modified DNA by photoaffinity labeling. *J. Am. Chem. Soc. 126*(21), 6536-6537.

(161) Zhu, G., and Lippard, S. J. (2009) Photoaffinity labeling reveals nuclear proteins that uniquely recognize cisplatin-DNA interstrand cross-links. *Biochemistry 48*(22), 4916-4925.

(162) Zhu, G., Chang, P., and Lippard, S. J. (2010) Recognition of platinum-DNA damage by poly(ADP-ribose) polymerase-1. *Biochemistry 49*(29), 6177-6183.

(163) Qiu, H. B., and Wang, Y. S. (2009) Exploring DNA-binding proteins with *in vivo* chemical cross-linking and mass spectrometry. *Journal of Proteome Research 8*(4), 1983-1991.

(164) Ray, R., and Miller, D. M. (1991) Cloning and characterization of a human c-myc promoter-binding protein. *Molecular and Cellular Biology 11*(4), 2154-2161.

(165) Chu, G. (1994) Cellular-responses to cisplatin - the roles of DNA-binding proteins and DNA-repair. *Journal of Biological Chemistry 269*(2), 787-790.

(166) Fram, R. J., Cusick, P. S., Wilson, J. M., and Marinus, M. G. (1985) Mismatch repair of *cis*-diamminedichloroplatinum(II)-induced DNA damage. *Molecular Pharmacology 28*(1), 51-55.

(167) Nakano, T., Katafuchi, A., Matsubara, M., Terato, H., Tsuboi, T., Masuda, T., Tatsumoto, T., Pack, S. P., Makino, K., Croteau, D. L., Van Houten, B., Iijima, K., Tauchi, H., and Ide, H. (2009) Homologous recombination but not nucleotide excision repair plays a pivotal role in tolerance of DNA-protein cross-links in mammalian cells. *Journal of Biological Chemistry 284*(40), 27065-27076.

(168) de Graaf, B., Clore, A., and McCullough, A. K. (2009) Cellular pathways for DNA repair and damage tolerance of formaldehyde-induced DNA-protein crosslinks. *Dna Repair 8*(10), 1207-1214.

(169) Bosl, G. J., Bajorin, D. F., Sheinfeld, J., and Motzer, R. *Cancer of the testis, 6th ed.* (2001) Lippincott Williams & Wilkins, Philadelphia, PA.

(170) Boulikas, T., and Vougiouka, M. (2004) Recent clinical trials using cisplatin, carboplatin and their combination chemotherapy drugs (review). *Oncol. Rep. 11*(3), 559-595.

(171) Miller, S. E., and House, D. A. (1991) The hydrolysis products of *cis*-diamminedichloroplatinum(II). *Inorganica Chimica Acta 187*(2), 125-132.

(172) Kartalou, M., and Essigmann, J. M. (2001) Recognition of cisplatin adducts by cellular proteins. *Mutat. Res. 478*(1-2), 1-21.

(173) Barry, M. A., Behnke, C. A., and Eastman, A. (1990) Activation of programmed cell-death (apoptosis) by cisplatin, other cnticancer drugs, toxins and hyperthermia. *Biochemical Pharmacology 40*(10), 2353-2362.

(174) Wozniak, K., and Walter, Z. (2000) Induction of DNA-protein cross-links by platinum compounds. *Zeitschrift fur Naturforschung C-A Journal of Biosciences 55*(9-10), 731-736.

(175) Bruhn, S. L., Pil, P. M., Essigmann, J. M., Housman, D. E., and Lippard, S. J. (1992) Isolation and characterization of human cDNA clones encoding a high mobility group box protein that recognizes structural distortions to DNA caused by binding of the anticancer agent cisplatin. *Proc. Natl. Acad. Sci. U. S. A 89*(6), 2307-2311.

(176) Yaneva, J., Leuba, S. H., vanHolde, K., and Zlatanova, J. (1997) The major chromatin protein histone H1 binds preferentially to *cis*-platinum-damaged DNA. *Proceedings of the National Academy of Sciences of the United States of America 94*(25), 13448-13451.

(177) Michaelson-Richie, E. D., Ming, X., Codreanu, S. G., Loeber, R. L., Liebler, D. C., Campbell, C., and Tretyakova, N. Y. (2011) Mechlorethamine-induced DNA-protein cross-Linking in human fibrosarcoma (HT1080) cells. *J. Proteome. Res. (In press)*.

(178) Rasheed, S., Nelson-Rees, W. A., Toth, E. M., Arnstein, P., and Gardner, M. B. (1974) Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer 33*(4), 1027-1033.

(179) Malmstrom, J., Larsen, K., Malmstrom, L., Tufvesson, E., Parker, K., Marchese, J., Williamson, B., Patterson, D., Martin, S., Juhasz, P., Westergren-Thorsson, G., and Marko-Varga, G. (2003) Nanocapillary liquid chromatography interfaced to tandem matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry: Mapping the nuclear proteome of human fibroblasts. *Electrophoresis 24*(21), 3806-3814.

(180) Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O., and Lippard, S. J. (1999) Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature 399*(6737), 708-712.

(181) Baik, M. H., Friesner, R. A., and Lippard, S. J. (2002) Theoretical study on the stability of *N*-glycosyl bonds: why does N7-platination not promote depurination? *J. Am. Chem. Soc. 124*(16), 4495-4503.

(182) Reardon, J. T., and Sancar, A. (2005) Nucleotide excision repair. *Prog. Nucleic Acid Res. Mol. Biol. 79*, 183-235.

(183) Chaney, S. G., and Sancar, A. (1996) DNA repair: enzymatic mechanisms and relevance to drug response. *J. Natl. Cancer Inst. 88*(19), 1346-1360.

(184) Mello, J. A., Acharya, S., Fishel, R., and Essigmann, J. M. (1996) The mismatch-repair protein hMSH2 binds selectively to DNA adducts of the anticancer drug cisplatin. *Chem. Biol. 3*(7), 579-589.

(185) Yamada, M., O'Regan, E., Brown, R., and Karran, P. (1997) Selective recognition of a cisplatin-DNA adduct by human mismatch repair proteins. *Nucleic Acids Res. 25*(3), 491-496.

(186) Turchi, J. J., and Henkels, K. (1996) Human Ku autoantigen binds cisplatin-damaged DNA but fails to stimulate human DNA-activated protein kinase. *J. Biol. Chem. 271*(23), 13861-13867.

(187) Riesbeck, K. (1999) Cisplatin at clinically relevant concentrations enhances interleukin-2 synthesis by human primary blood lymphocytes. *Anti-Cancer Drugs 10*(2), 219-227.

(188) Cadet, J., Douki, T., and Ravanat, J. L. (2010) Oxidatively generated base damage to cellular DNA. *Free Radic. Biol. Med. 49*(1), 9-21.

(189) Valko, M., Rhodes, C. J., Moncol, J., Izakovic, M., and Mazur, M. (2006) Free radicals, metals and antioxidants in oxidative stress-induced cancer. *Chem. Biol. Interact. 160*(1), 1-40.

(190) Evans, M. D., Dizdaroglu, M., and Cooke, M. S. (2004) Oxidative DNA damage and disease: induction, repair and significance. *Mutat. Res. 567*(1), 1-61.

(191) Bennett, M. R. (2001) Reactive oxygen species and death: oxidative DNA damage in atherosclerosis. *Circ. Res. 88*(7), 648-650.

(192) Canella, K. A., Diwan, B. A., Gorelick, P. L., Donovan, P. J., Sipowicz, M. A., Kasprzak, K. S., Weghorst, C. M., Snyderwine, E. G., Davis, C. D., Keefer, L. K., Kyrtopoulos, S. A., Hecht, S. S., Wang, M., Anderson, L. M., and Rice, J. M. (1996) Liver tumorigenesis by Helicobacter hepaticus: considerations of mechanism. *In Vivo 10*(3), 285-292.

(193) Malins, D. C., Holmes, E. H., Polissar, N. L., and Gunselman, S. J. (1993) The etiology of breast cancer. Characteristic alteration in hydroxyl radical-induced DNA base lesions during oncogenesis with potential for evaluating incidence risk. *Cancer 71*(10), 3036-3043.

(194) Greenberg, M. M. (2004) *In vitro* and *in vivo* effects of oxidative damage to deoxyguanosine. *Biochem. Soc. Trans. 32*(Pt 1), 46-50.

(195) Paz-Elizur, T., Krupsky, M., Blumenstein, S., Elinger, D., Schechtman, E., and Livneh, Z. (2003) DNA repair activity for oxidative damage and risk of lung cancer. *J. Natl. Cancer Inst. 95*(17), 1312-1319.

(196) Helbock, H. J., Beckman, K. B., Shigenaga, M. K., Walter, P. B., Woodall, A. A., Yeo, H. C., and Ames, B. N. (1998) DNA oxidation matters: the HPLC-electrochemical detection assay of 8-oxo-deoxyguanosine and 8-oxo-guanine. *Proc. Natl. Acad. Sci. U. S. A 95*(1), 288-293.

(197) Halliwell, B., Gutteridge, J. M., and Cross, C. E. (1992) Free radicals, antioxidants, and human disease: where are we now? *J. Lab Clin. Med. 119*(6), 598-620.

(198) Malins, D. C. (1996) Free radicals and breast cancer. *Environ. Health Perspect. 104*(11), 1140.

(199) Cadet, J., and Treoule, R. (1978) Comparative study of oxidation of nucleic acid components by hydroxyl radicals, singlet oxygen and superoxide anion radicals. *Photochem. Photobiol. 28*(4-5), 661-667.

(200) Douki, T., and Cadet, J. (1996) Peroxynitrite mediated oxidation of purine bases of nucleosides and isolated DNA. *Free Radic. Res. 24*(5), 369-380.

(201) Luo, W., Muller, J. G., Rachlin, E. M., and Burrows, C. J. (2000) Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-oxo-7,8-dihydroguanosine. *Org. Lett. 2*(5), 613-616.

(202) Ravanat, J. L., and Cadet, J. (1995) Reaction of singlet oxygen with 2'-deoxyguanosine and DNA. Isolation and characterization of the main oxidation products. *Chem. Res. Toxicol. 8*(3), 379-388.

(203) Niles, J. C., Wishnok, J. S., and Tannenbaum, S. R. (2001) Spiroiminodihydantoin is the major product of the 8-oxo-7,8-dihydroguanosine reaction with peroxynitrite in the presence of thiols and guanosine photooxidation by methylene blue. *Org. Lett. 3*(7), 963-966.

(204) Luo, W., Muller, J. G., Rachlin, E. M., and Burrows, C. J. (2001) Characterization of hydantoin products from one-electron oxidation of 8-oxo-7,8-dihydroguanosine in a nucleoside model. *Chem. Res. Toxicol. 14*(7), 927-938.

(205) Cadet, J., Berger, M., Buchko, G. W., Joshi, P. C., Raoul, S., and Ravanat, J. L. (1994) 2,2-Diamino-4-[(3,5-di-*O*-acetyl-2-deoxy-*beta*-D-*erythro*-pentofuranosyl)amino]-5-(*2H*)-oxazolone - A novel and predominant radical oxidation-product of 3',5'-di-*O*-acetyl-2'-deoxyguanosine. *Journal of the American Chemical Society 116*(16), 7403-7404.

(206) Saito, I., Takayama, M., Nakamura, T., Sugiyama, H., Komeda, Y., and Iwasaki, M. (1995) The most electron-donating sites in duplex DNA: guanine-guanine stacking rule. *Nucleic Acids Symp. Ser.*(34), 191-192.

(207) Kino, K., Saito, I., and Sugiyama, H. (1998) Product analysis of GG-specific photooxidation of DNA *via* electron transfer: 2-aminoimidazolone as a major guanine oxidation product. *Journal of the American Chemical Society 120*(29), 7373-7374.

(208) Prat, F., Houk, K. N., and Foote, C. S. (1998) Effect of guanine stacking on the oxidation of 8-oxoguanine in B-DNA. *Journal of the American Chemical Society 120*(4), 845-846.

(209) Nunez, M. E., Rajski, S. R., and Barton, J. K. (2000) Damage to DNA by long-range charge transport. *Methods Enzymol. 319*, 165-188.

(210) Odom, D. T., and Barton, J. K. (2001) Long-range oxidative damage in DNA/RNA duplexes. *Biochemistry 40*(30), 8727-8737.

(211) Hirakawa, K., Yoshida, M., Oikawa, S., and Kawanishi, S. (2003) Base oxidation at 5 ' site of GG sequence in double-stranded DNA induced by UVA in the presence of xanthone analogues: Relationship between the DNA-damaging abilities of photosensitizers and their HOMO energies. *Photochemistry and Photobiology 77*(4), 349-355.

(212) Naseem, I., Ahmad, M., and Hadi, S. M. (1988) Effect of alkylated and intercalated DNA on the generation of superoxide anion by riboflavin. *Biosci. Rep. 8*(5), 485-492.

(213) Foote, C. S. (1991) Definition of type I and type II photosensitized oxidation. *Photochem. Photobiol. 54*(5), 659.

(214) Ito, K., Inoue, S., Yamamoto, K., and Kawanishi, S. (1993) 8-Hydroxydeoxyguanosine formation at the 5' site of 5'-GG-3' sequences in double-stranded DNA by UV-radiation with riboflavin. *Journal of Biological Chemistry 268*(18), 13221-13227.

(215) Tretyakova, N., Matter, B., Jones, R., and Shallop, A. (2002) Formation of benzo[a]pyrene diol epoxide-DNA adducts at specific guanines within *K-ras* and *p53* gene sequences: stable isotope-labeling mass spectrometry approach. *Biochemistry 41*(30), 9535-9544.

(216) Burrows, C. J., and Muller, J. G. (1998) Oxidative nucleobase modifications leading to strand scission. *Chem. Rev. 98*(3), 1109-1152.

(217) Shallop, A. J., Gaffney, B. L., and Jones, R. A. (2003) Use of 13C as an indirect tag in 15N specifically labeled nucleosides. Syntheses of [8-13C-1,7,NH2-15N3]adenosine, -guanosine, and their deoxy analogues. *J. Org. Chem. 68*(22), 8657-8661.

(218) Matter, B., Malejka-Giganti, D., Csallany, A. S., and Tretyakova, N. (2006) Quantitative analysis of the oxidative DNA lesion, 2,2-diamino-4-(2-deoxy-*beta*-D-*erythro*-pentofuranosyl)amino]-5(*2H*)-oxazolone (oxazolone), *in vitro* and *in vivo* by isotope dilution-capillary HPLC-ESI-MS/MS. *Nucleic Acids Res. 34*(19), 5449-5460.

(219) Tornaletti, S., and Pfeifer, G. P. (1995) Complete and tissue-independent methylation of CpG sites in the *p53* gene: implications for mutations in human cancers. *Oncogene 10*(8), 1493-1499.

(220) Rajesh, M., Wang, G., Jones, R., and Tretyakova, N. (2005) Stable isotope labeling-mass spectrometry analysis of methyl- and pyridyloxobutyl-guanine adducts of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in *p53*-derived DNA sequences. *Biochemistry 44*(6), 2197-2207.

(221) Sheskin, D. J. (2000) The between-subjects factorial analysis of variance. In *Handbook of parametric and Nonparametric Statistical Procedures* pp 705-755, CRC Press, New York.

(222) Matter, B., Murphy, K., Zhao, J., Lindgren, B., Jones, R., and Tretyakova, N. (2011) Sequence distribution of oxidative guanine adducts following exposure of *K-ras* gene derived DNA duplexes to three classes of reactive oxygen species . *Manuscript In Preparation*.

(223) Hussain, S. P., Hollstein, M. H., and Harris, C. C. (2000) *p53* tumor suppressor gene: at the crossroads of molecular carcinogenesis, molecular epidemiology, and human risk assessment. *Ann. N. Y. Acad. Sci. 919*, 79-85.

(224) Hainaut, P., and Pfeifer, G. P. (2001) Patterns of *p53* G-->T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke. *Carcinogenesis 22*(3), 367-374.

(225) Hernandez-Boussard, T., Rodriguez-Tome, P., Montesano, R., and Hainaut, P. (1999) IARC *p53* mutation database: a relational database to compile and analyze *p53* mutations in human tumors and cell lines. International Agency for Research on Cancer. *Hum. Mutat. 14*(1), 1-8.

(226) Pfeifer, G. P., Tang, M., and Denissenko, M. F. (2000) Mutation hotspots and DNA methylation. *Curr. Top. Microbiol. Immunol. 249*, 1-19.

(227) Fritsche, M., Haessler, C., and Brandner, G. (1993) Induction of nuclear accumulation of the tumor-suppressor protein *p53* by DNA-damaging agents. *Oncogene 8*(2), 307-318.

(228) Prieme, H., Loft, S., Klarlund, M., Gronbaek, K., Tonnesen, P., and Poulsen, H. E. (1998) Effect of smoking cessation on oxidative DNA modification estimated by 8-oxo-7,8-dihydro-2 '-deoxyguanosine excretion. *Carcinogenesis 19*(2), 347-351.

(229) Park, J. H., Gopishetty, S., Szewczuk, L. M., Troxel, A. B., Harvey, R. G., and Penning, T. M. (2005) Formation of 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxo-dGuo) by PAH o-quinones: Involvement of reactive oxygen species and copper(II)/copper(I) redox cycling. *Chemical Research in Toxicology 18*(6), 1026-1037.

(230) Yu, D. S., Berlin, J. A., Penning, T. M., and Field, J. (2002) Reactive oxygen species generated by PAH o-quinones cause change-in-function mutations in *p53*. *Chemical Research in Toxicology 15*(6), 832-842.

(231) Pryor, W. A., Prier, D. G., and Church, D. F. (1983) Electron-spin resonance study of mainstream and sidestream cigarette-smoke - nature of the free-radicals in gas-phase smoke and in cigarette tar. *Environmental Health Perspectives 47*(JAN), 345-355.

(232) Sowers, L. C., Shaw, B. R., and Sedwick, W. D. (1987) Base stacking and molecular polarizability - effect of a methyl-group in the 5-position of pyrimidines. *Biochemical and Biophysical Research Communications 148*(2), 790-794.

(233) Lee, Y. A., Yun, B. H., Kim, S. K., Margolin, Y., Dedon, P. C., Geacintov, N. E., and Shafirovich, V. (2007) Mechanisms of oxidation of guanine in DNA by carbonate radical anion, a decomposition product of nitrosoperoxycarbonate. *Chemistry. 13*(16), 4571-4581.

(234) Aris, S. M., and Farrell, N. P. (2009) Towards antitumor active *trans*-platinum compounds. *European Journal of Inorganic Chemistry*(10), 1293-1302.

(235) Ross, P. L., Huang, Y. L. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Molecular & Cellular Proteomics 3*(12), 1154-1169.

(236) Maddukuri, L., Eoff, R. L., Choi, J. Y., Rizzo, C. J., Guengerich, F. P., and Marnett, L. J. (2010) *In vitro* bypass of the major malondialdehyde- and base propenal-derived DNA adduct by human y-family DNA polymerases kappa, iota, and Rev1. *Biochemistry 49*(38), 8415-8424.

(237) Ang, W. H., Myint, M., and Lippard, S. J. (2010) Transcription inhibition by platinum-DNA cross-links in live mammalian cells. *Journal of the American Chemical Society 132*(21), 7429-7435.