BIOLOGICAL CONSEQUENCES OF COMPLEX DNA LESIONS INDUCED BY *BIS*-ELECTROPHILES

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Susith Samira Kumara Wickramaratne

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Natalia Y. Tretyakova and Dr. Mark D. Distefano, Co-Advisers

August 2014

ACKNOWLEDGEMENTS

As I pen down these words at the verge of completing my doctoral dissertation, flashes of memories of many people, who have helped me through years in graduate school, come to my mind. First and foremost, I would like to convey my heartfelt gratitude to my advisers, Dr. Natalia Tretyakova and Dr. Mark Distefano for their guidance and support throughout the last five years. I have learnt a lot of science as well as the way of life in academia. I am especially thankful to them for the opportunities to work collaboratively on numerous multidisciplinary projects and to get involved in professional development prospects. I also would like to thank Dr. Colin Campbell for his mentorship in biology, when I worked in his laboratory over the years. He was not only a collaborator, but also a wonderful person to spend time with. A special word of gratitude goes to my dissertation committee, Dr. Valerie Pierre, Dr. Michael Bowser and Dr. Daniel Harki for their valuable time and positive feedback.

I am grateful to Tretyakova and Distefano group members for their mentorship, contributions and helpful discussions. Dr. Uthpala Seneviratne, with whom I started my graduate research at the dawn of everyday during my first winter, and Dr. Delshanee Kotandeniya, with whom I spent long days at the synthesis hood and cell culture lab, were great mentors. Emily Boldry, Charles Buehler, James Lloyd and Shaofei Ji who worked with me in my projects never complained for long days we had together in the lab. I am also thankful to Dr. Srikanth Kotapati, Charuta Palsuledesai, Veronica Diaz-Rodriguez, Dewakar Sangaraju, Dr. Bhaskar Malayappan, Xun Ming, Arnie Groehler, Chris Seiler, Dr. Melissa Goggin, Erin Michaelson-Richie, Dr. Jung-Eun Yeo, Dr. Teshome Gherezghiher, Dr. Shantoshkumar Khatwani, Dr. Yeng-Chih Wang, Jeffrey Vervacke, Yi Zhang and all current and past group members for all their contributions in many ways throughout these five years. Especially the amazing friendship we shared and the moral support during hard times kept me going.

Most of my projects were of collaborative nature, and the work would not have been possible without our gracious collaborators. My sincere thank you to Dr. Orlando Scharer, Shivam Mukherjee, Dr. Sheila David, Douglas Banda, Dr. Peter Guengerich, Dr. Chuan He, Dr. Leona Samson, Dr. Michael Stone, Dr. Ewa Kowal, Dr. John Essigmann Dr. Robert Vince and Dr. Ashis Basu for performing experiments and/or generous donation of enzymes for my thesis projects and for numerous other projects. I am very much thankful to Brock Matter for his help and expertise in method development and troubleshooting as well as for the training in analytical instrumentation. I would also like to acknowledge Dr. Peter Villalta for his help with mass spectrometry instrumentation, especially in DNA-protein cross-linking projects. Many thanks to Robert Carlson for his help with preparing figures for my thesis, publications and presentations. I am also thankful to members of Hecht, Harki, Murphy, Wagner, Walseth, Peterson, Finzel, Romero, Stepanov, Portoghese, Abul-Hajj and Remmel laboratories for kindly allowing me to use instrumentation and sharing chemicals from time to time.

Last but not least, I would like to thank my family and friends, especially my parents, my in-laws, my sister and my brother-in-law for their unconditional love, continuous support and encouragement, without whom I could not have accomplished most things in my life. I am so much grateful to my beloved wife, Thakshila, for her unwavering love and endless care. Without her inspiration and dedication, I could not have completed this degree. Finally, I would like to pay my sincere gratitude to each and every one, who directly or indirectly contributed to my doctoral dissertation and to my graduate studies.

DEDICATION

To my parents,
especially to my father,
who had faith in me that I will succeed in life,
who gave me the freedom to explore life of my own
since my young days.

To my beloved wife,
Thakshila,
for her unwavering love and endless care,
for her dazzling smile
that made my long and hard days refreshing.

ABSTRACT

Genomic DNA is continuously damaged by various endogenous and exogenous physicochemical agents. Specifically, *bis*-electrophiles produced endogenously as a result of normal cellular metabolism, environmental carcinogens, and chemotherapeutic agents can alkylate multiple sites on biomolecules to form a wide array of damaged DNA called DNA lesions (DNA adducts). Among these are small DNA monoadducts and exocyclic lesions as well as helix distorting DNA-DNA cross-links and bulky DNA-protein cross-links. The resulting lesions, if not repaired, can elicit adverse biological effects including cytotoxicity and mutations due to their ability to interfere with key cellular DNA transactions such as DNA repair, replication and transcription. Unless recognized by cellular repair mechanisms to restore normal DNA, DNA adducts can be replicated in an error-prone manner. Despite the eminent threat of DNA damage by *bis*-electrophiles to cell viability and genomic integrity, the biological consequences of the resulting complex DNA adducts are not fully understood. In the present work, we investigated the mechanisms of repair and replication of complex DNA lesions formed by reactive metabolites of a known human carcinogen, 1,3-butadiene (BD) and antitumor nitrogen mustards.

BD is an industrial chemical and environmental pollutant that has a widespread occurrence in the urban environment and high potential for human exposure. BD is metabolized by cytochrome P450 enzymes in cells to generate highly reactive epoxides, 3,4-epoxy-1-butene (EB), 3,4-epoxy-1,2-butanediol (EBD) and 1,2,3,4-diepoxybutane (DEB) that can form covalent DNA adducts. In the first part of this thesis, we investigated the repair mechanisms responsible for the removal of three potentially mutagenic BD-induced 2'-deoxyadenosine lesions recently discovered in our laboratory: N^6-(2-hydroxy-3-buten-1-yl)-2'-deoxyadenosine (N^6-HB-dA), N^6,N^6-(2,3-dihydroxy-butan-1,4-diyl)-2'-deoxyadenosine (N^6,N^6-DHB-dA) and 1,N^6-(2-hydroxy-3-hydroxy-

methylpropan-1,3-diyl)-2'-deoxyadenosine (1,N^6-HMHP-dA). Synthetic oligodeoxynucleotides containing site- and stereospecific BD-dA adducts were prepared by a post-oligomerization strategy. *In vitro* repair assays using human fibrosarcoma nuclear extracts revealed that all three BD-dA adducts were repaired by base excision repair (BER) pathway.

Covalent entrapment of cellular proteins on genomic DNA mediated by various endogenous and exogenous physicochemical species results in the formation of highly heterogeneous and unusually bulky DNA adducts called DNA-protein cross-links (DPCs). The structural complexity and diversity of DPCs and the difficulty of generating site-specific DNA-protein conjugates represent a special challenge, hindering the previous studies of their biological outcomes. In the second part of the thesis, we developed novel synthetic methods to generate hydrolytically stable model DPCs and examined their effects on DNA replication.

In our first approach, we employed post-synthetic reductive amination between 7-(2-oxoethyl)-7-deazaguanine containing DNA and amino side chains of lysine/arginine residues on proteins to generate hydrolytically stable DPCs to the C7 position of deazaguanine in DNA. The resulting model lesions are structural analogs of the major DPCs formed in cells mediated by *bis*-electrophiles such as chlorooxirane and antitumor nitrogen mustards. In our second approach, copper-catalyzed [3+2] Huisgen cycloaddition between alkyne containing DNA and polypeptides engineered to contain an azide functionality was used to generate DPCs to the C5 of thymidine on DNA.

The resulting site-specifically modified DNA was subjected to *in vitro* replication bypass assays using recombinant human translesion synthesis (TLS) polymerases, hPol η, κ and ι. These experiments have revealed that large proteins and peptides cross-linked to either 7-deaza-G or C5-T completely blocked replication, while a decapeptide cross-link was bypassed with varying efficiencies, suggesting that the protein component of DPCs undergoes proteolytic processing prior to *in vivo* replication. Steady-state kinetic studies provided evidence for the highly error-prone lesion bypass of peptides cross-linked to C5-T and high fidelity bypass of the corresponding 7-deaza-G conjugates. Although hPol

η showed a higher catalytic efficiency in bypassing these lesions, hPol κ was more accurate in lesion bypass. Further, HPLC-ESI$^-$-HRMS and MS/MS quantitation and sequencing of replication products of decapeptide cross-linked to C5-T revealed a large number of deletion and substitution mutations. Taken together, our data suggest that efficiency and fidelity of replication bypass of DNA polypeptide cross-links are dependent on the lesion size, the cross-linking site within DNA, and the identity of the polymerase.

TABLE OF CONTENTS

Chapter *Page*

ACKNOWLEDGEMENTS……...i

DEDICATION……..iii

ABSTRACT……...iv

LIST OF TABLES……..xvi

LIST OF SCHEMES……xviii

LIST OF CHARTS……xxi

LIST OF FIGURES……...xxii

LIST OF ABBREVIATIONS……..xxvi

1 LITERATURE REVIEW 1

1.1 DNA Damage and Human Diseases 1

1.1.1 DNA Modification and Its Consequences in Cells 1

1.1.2 Causative Agents and Types of DNA Damage..... 1

1.1.3 Role of DNA Damage in Human Disease 8

1.2 1,3-Butadiene-Mediated Mutagenesis and Carcinogenesis 10

1.2.1 1,3-Butadiene: A Potent Environmental Carcinogen..... 10

1.2.2 Metabolic Activation of 1,3-Butadiene and DNA Adduction 11

1.2.3 Adverse Biological Effects of 1,3-Butadiene 14

1.2.4 DNA Lesions of 1,3-Butadiene..... 18

1.3 Repair of 1,3-Butadiene-Induced DNA Adducts 24

1.3.1 Cellular DNA Repair Mechanisms 24

1.3.2 *In Vivo* and *In Vitro* Persistence and Repair Studies 31

1.4 Replication of 1,3-Butadiene-Mediated DNA Adducts 33

1.4.1 DNA Adducts and Translesion Synthesis..... 33

1.4.2 *In Vivo* Replication and Mutagenesis Studies of 1,3-Butadiene DNA Adducts 36

1.4.3 *In Vitro* Replication Bypass of 1,3-Butadiene-Induced DNA Adducts..... 40

1.5 DNA-Protein Cross-Linking by *Bis*-Electrophiles 44

1.5.1 DNA-Protein Cross-Links: A Ubiquitous Class of DNA Adducts..... 44

1.5.2 Types of DNA-Protein Cross-Links 46

1.5.3 DNA-Protein Cross-Links Induced By Physicochemical Agents 47

1.5.4 DNA-Protein Cross-Linking by Antitumor Drugs 52

1.5.5 Biological Consequences of DNA-Protein Cross-Links..... 59

1.6 Repair Studies of DNA-Protein Cross-Links 62

1.7 Replication Bypass of DNA-Protein Cross-Links..... 66

1.8 Synthetic methodlogies to prepare Site-Specific DNA-Protein Cross-Links 70

1.9 Thesis Goals 73

2 BASE EXCISION REPAIR OF EXOCYCLIC 2'-DEOXYADENOSINE ADDUCTS OF 1,3-BUTADIENE 75

2.1 Introduction 75

2.2 Materials and Methods 81

2.2.1 Materials 81

2.2.2 Preparation of Human Fibrosarcoma Nuclear Extracts 82

2.2.3 Synthesis of Site-Specifically Modified DNA Substrates 83

2.2.4 Preparation of Radiolabeled Double-Stranded Oligodeoxynucleotides 84

2.2.5 Base Excision Repair Assays with Human Fibrosarcoma Nuclear Extracts 84

2.2.6 Base Excision Repair Assays Using Recombinant Enzymes 85

2.2.7 Monitoring Base Excision Repair by Gel Electrophoresis 86

2.2.8 Analysis of Base Excision Repair by Liquid Chromatography-Tandem Mass Spectrometry 87

2.3 Results 89

2.3.1 Synthesis, Purification and Characterization of Adducted DNA Oligodeoxy-nucleotides 89

2.3.2 Base Excision Repair of 1,3-Butadiene-Induced 2'-Deoxyadenosine Adducts Using Human Fibrosarcoma Nuclear Extracts 95

2.3.3 Analysis of Base Excision Repair by Liquid Chromatography-Tandem Mass Spectrometry 103

2.3.4 Activity of Recombinant BER Glycosylases against BD-dA Adducts 105

2.4 Discussion 108

3 SYNTHESIS OF SEQUENCE-SPECIFIC DNA-PROTEIN CONJUGATES CROSS-LINKED TO 7-DEAZAGUANINE *VIA* A REDUCTIVE AMNATION STRATEGY 111

3.1 Introduction 111

3.2 Materials and Methods 116

3.2.1 Materials 116

3.2.2 Preparation of Radiolabeled DNA Duplexes 116

3.2.3 Reductive Amination to Generate DNA-Protein Cross-Links 117

3.2.4 Gel Electrophoretic Analysis of DNA-Protein Cross-Links Generated by Reductive Amination 117

3.2.5 Sample Processing for Mass Spectrometry Analysis 118

3.2.6 Characterization of DNA-Protein Cross-Links by Mass Spectrometry 119

3.2.7 Synthesis and Characterization of 7-Deaza-7-(2-(*N*-acetyllysine)ethan-1-yl)-2'-deoxyguanosine and 7-Deaza-7-(2-(*N*-acetylarginine)ethan-1-yl)-2'-deoxyguanosine Conjugates 120

3.2.8 Synthesis and Characterization of Nucleoside-Peptide Conjugates 121

3.3 Results..... 123

3.3.1 Experimental Strategy for the Generation of Hydrolytically Stable Model DNA-Protein Cross-Links 123

3.3.2 Characterization of DNA-Protein Cross-Link Formation by Denaturing Gel Electrophoresis..... 126

3.3.3 Effects of Reaction Conditions on DNA-Protein Cross-Link Yields 129

3.3.4 Influence of Protein Identity on DNA-Protein Cross-Link Formation 131

3.3.5 Mass Spectrometry Characterization of DNA-Protein Cross-Links..... 135

3.3.6 Synthesis and Structural Characterization of 7-Deaza-7-(2-(*N*-acetyl lysine)ethan-1-yl)-2'-deoxyguanosine and 7-Deaza-7-(2-(*N*-acetylarginine) ethan-1-yl)-2'-deoxyguanosine Conjugates 141

3.3.7 Synthesis and Characterization of Nucleoside-Peptide Conjugates 143

3.4 Discussion 145

4 REPLICATION BYPASS OF MODEL DNA-PROTEIN CROSS-LINKS CONJUGATED TO THE 7-DEAZAGUANINE OF DNA 151

4.1 Introduction 151

4.2 Materials and Methods 153

4.2.1 Materials 153

4.2.2 Synthesis and Characterization of Oligodeoxynucleotides..... 154

4.2.3 Synthesis and Characterization of DNA-Protein and DNA-Peptide Cross-Links 154

4.2.4 Preparation of Primer-Template Duplexes 157
4.2.5 Primer Extension Assays 157
4.2.6 Single Nucleotide Incorporation Assays 158
4.2.7 Steady-State Kinetic Analyses 160
4.3 Results 161
4.3.1 Synthesis and Characterization of DNA-Protein and DNA-Peptide Cross-Links 161
4.3.2 Replication Bypass of DNA-Protein Cross-Links 162
4.3.3 Primer Extension Past DNA-Peptide Cross-Links 166
4.3.4 Single Nucleotide Incorporation Opposite a 10-mer Peptide Cross-Link 169
4.3.5 Steady-State Kinetic Analyses of Nucleotide Incorporation Opposite the 10-mer Peptide Cross-Link 171
4.4 Discussion 174
5 SYNTHESIS OF SITE-SPECIFIC DNA-PROTEIN CROSS-LINKS CONJUGATED TO THE C5 OF THYMIDINE AND THEIR EFFECTS ON DNA REPLICATION 180
5.1 Introduction 180
5.2 Materials and Methods 182
5.2.1 Materials 182
5.2.2 Oligodeoxyonucleotide Synthesis 183
5.2.3 Preparation of Radiolabeled Oligodeoxynucleotides 184

5.2.4 Preparation of 6×His-eGFP-N_3 184

5.2.5 Preparation of Azide-Containing Peptides..... 185

5.2.6 Copper-Catalyzed Cycloaddition Reaction between 6×His-eGFP-N_3 and Alkyne-Containing DNA 187

5.2.7 Cross-Linking Reaction between Azide-Functionalized Peptides and Alkyne-Containing DNA 188

5.2.8 Gel Electrophoresis Purification of DNA-Protein Cross-Links..... 189

5.2.9 Mass Spectrometry Analysis of DNA-Protein Cross-Links 189

5.2.10 Mass Spectrometry Characterization of Synthetic DNA-Peptide Cross-Links.....191

5.2.11 Polymerase Bypass Assay..... 192

5.3 Results..... 194

5.3.1 Site-Specific DNA–Protein Cross-Linking Using Alkyne−Azide Cycloaddition Reaction..... 194

5.3.2 Mass Spectrometry Characterization of DNA–Protein and DNA–Peptide Conjugates..... 204

5.3.3 Polymerase Bypass of Synthetic DNA–Protein and DNA–Peptide Conjugates..... 208

5.4 Discussion 215

6 ERROR-PRONE TRANSLESION SYNTHESIS PAST DNA-PEPTIDE CROSS-LINKS CONJUGATED TO THE MAJOR GROOVE OF DNA *VIA* C5 OF THYMIDINE 219

6.1 Introduction 219

6.2 Materials and Methods 222

6.2.1 Materials 222

6.2.2 Synthesis and Characterization of Oligodeoxynucleotides 222

6.2.3 Synthesis and Characterization of DNA-Peptide Conjugates 223

6.2.4 Preparation of Primer-Template Duplexes 224

6.2.5 Single Nucleotide Incorporation Assays 224

6.2.6 Steady-State Kinetic Analyses 225

6.2.7 Sequencing and Quantitation of Primer Extension Products by Liquid Chromatography-Tandem Mass Spectrometry 226

6.3 Results 229

6.3.1 Synthesis of Primer-Template Duplexes Containing Site-Specific DNA-Peptide Conjugates 229

6.3.2 Single Nucleotide Incorporation Assays 232

6.3.3 Steady-State Kinetic Analyses 235

6.3.4 Sequencing and Quantitation of Primer Extension Products by Liquid Chromatography-Tandem Mass Spectrometry 238

6.4 Discussion 250

7 SUMMARY AND CONCLUSIONS 255

8 FUTURE DIRECTIONS 266

8.1 Identification of DNA Glycosylases Involved in Repair of N^6-2'-Deoxy-adenosine Lesions of 1,3-Butadiene 266

8.2 *In Vitro* and *In Vivo* Investigation of Repair of 1,3-Butadiene-induced N^6-Adenine Adducts 272

8.3 Examination of *In Vivo* Mutagenesis of 2'-Deoxy-adenosine Lesions of 1,3-Butadiene 275

8.4 Synthesis of Site- and Stereo-Specific Oligo-deoxynucleotides Containing 1-(Aden-N^6-yl)-4-(guan-7-yl)-butan-2,3-diol for Structural and Biological Studies 279

8.5 Further Investigation of *In Vitro* Replication of Model DNA-Protein Cross-Links 285

8.6 Identification of Repair Mechanisms Responsible for Removal of DNA-Protein Cross-Links 290

8.7 Investigation of Mutagenicity of DNA-Protein Cross-Links in Cells 293

8.8 Structure Elucidation of DNA-Protein Cross-Links 295

9 BIBLIOGRAPHY 296

LIST OF TABLES

Table *Page*

Table 1-1 Known chemical carcinogens and their carcinogenic properties. 6

Table 1-2 *Hprt* mutation spectra of 3,4-epoxy-1-butene (EB) and 1,2,3,4-diepoxybutane (DEB) in human TK6 lymphoblast cells after 24 h exposure. 15

Table 1-3 Summary of *in vivo* mutagenicity of 1,3-butadiene-induced DNA lesions. ... 37

Table 1-4 *In vitro* replication studies of 1,3-butadiene-induced DNA adducts. 43

Table 2-1 Capillary HPLC-ESI$^-$-MS characterization of synthetic DNA oligomers. 94

Table 2-2 First order rate constants observed for the repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts and known BER substrates by human fibrosarcoma nuclear extracts. 99

Table 3-1 Proteins and peptides used to generate DNA-protein cross-links using the reductive amination reactions with deaza-DHP-dG containing DNA. 133

Table 3-2 Sites of reductive amination-mediated cross-linking between deaza-DHP-dG containing DNA and recombinant AlkB protein as identified by nano HPLC-ESI$^+$-MS/MS of tryptic digests. 137

Table 3-3 Sites of reductive amination-mediated cross-linking between deaza-DHP-dG containing DNA and proteins as identified by nano HPLC-ESI$^+$-MS/MS of tryptic digests. 138

Table 4-1 Steady-state kinetic parameters for single nucleotide insertion opposite the positive control (dG) and the 10mer peptide cross-linked to *C7* of deaza-guanine (10-mer peptide) by human TLS polymerases, hPol η and κ. 173

Table 6-1 Steady-state kinetic parameters for single nucleotide insertion opposite the positive control (dT) and the 10-mer peptide cross-linked to *C5* of thymidine (dT-peptide) by human TLS polymerases, hPol κ and η. 236

Table 6-2 Summary of extension products formed by hPol κ and η as identified by liquid chromatography-tandem mass spectrometry. 242

Table 8-1 Mutation analysis of ss-pMS2 plasmids containing 6×His-eGFP cross-link replicated in HEK293T cells. 294

LIST OF SCHEMES

Scheme *Page*

Scheme 1-1 Metabolic activation of 1,3-butadiene and formation of DNA adducts. 13

Scheme 1-2 Cellular repair mechanisms: direct repair, base excision repair and nucleotide excision repair. 25

Scheme 1-3 Cellular repair mechanisms: homologous recombination, non-homologous end-joining and mismatch repair. 30

Scheme 1-4 Polymerase switching model for translesion synthesis. 34

Scheme 1-5 Mechanisms of DNA-protein cross-linking by alkyl sulfonates (A) and 2-chloroethylnitrosoureas (B)..... 54

Scheme 1-6 Mechanisms of DNA-protein cross-linking by nitrogen mustards (A) and platinum drugs (B)..... 57

Scheme 1-7 Synthetic methodologies available in the literature to prepare site-specific DNA-protein cross-links. 72

Scheme 2-1 Proposed mechanism for the formation of 1,3-butadiene-induced 2'-deoxyadenosine adducts. 77

Scheme 2-2 Structures and sequences of DNA lesions used in repair studies……...80

Scheme 2-3 Synthesis of site- and stereospecific oligodeoxynucleotides containing 1,3-butadiene-induced 2'-deoxyadenosine adducts using a post-oligomerization approach. 91

Scheme 3-1 Structures of DNA-protein cross-links induced by *bis*-electrophiles. ……..113

Scheme 3-2 Synthesis of DNA-protein cross-links by a post-synthetic reductive amination strategy. 124

Scheme 3-3 Structural similarity between the model DNA-protein cross-links prepared by reductive amination and DNA-protein cross-links formed *in vivo*. 125

Scheme 4-1 Synthesis of DNA-protein and DNA-peptide cross-links by post-synthetic reductive amination. 156

Scheme 4-2 Sequences of DNA oligomers employed in *in vitro* replication experiments……...159

Scheme 5-1 Generation of site-specific DNA-protein conjugates by copper-catalyzed [3+2] Huisgen cycloaddition. 195

Scheme 5-2 Synthesis of site-specific DNA-peptide cross-links by copper-catalyzed azide-alkyne cycloaddition reaction. 196

Scheme 5-3 Sequences of DNA oligomers used for conjugation reactions with proteins and peptides (A) and DNA substrates employed in standing start (B) and running start primer extension experiments (C). 198

Scheme 6-1 Synthesis of DNA-peptide cross-links by copper-catalyzed [3+2] Huisgen cycloaddition. 230

Scheme 6-2 Sequences of DNA oligomers employed in *in vitro* replication studies. ……...233

Scheme 6-3 Streptavidin capture in combination with capillary HPLC-ESI$^-$-MS/MS method for sequencing and quantitation of primer extension products. .. 240

Scheme 8-1 Mass spectrometry-based proteomics approach to identifying the base excision repair glycosylases responsible for the repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts. 269

Scheme 8-2 Synthesis of 1,3-butadiene-induced 2'-deoxyadenosine adducts on solid support. 270

Scheme 8-3 *In vitro* repair experiments to examine the involvement of nucleotide excision repair in the removal of 2'-deoxyadenosine lesions of 1,3-butadiene. 274

Scheme 8-4 *In vivo* mutagenesis studies of 1,3-butadiene-induce 2'-deoxyadenosine adducts. 278

Scheme 8-5 Formation of hydrolytically stable 1-(adenosin-N^6-yl)-4-(guan-7-yl)-butan-2,3-diol adduct. 281

Scheme 8-6 Synthesis of oligodeoxynucleotides containing 1-(adenosin-N^6-yl)-4-(guan-7-yl)-butan-2,3-diol. 283

Scheme 8-7 Streptavidin capture in combination with capillary HPLC-ESI$^-$-MS/MS methodology for sequencing and quantitation of primer extension products. 286

Scheme 8-8 Preparation of double-stranded pGL3 plasmids containing site-specific DNA-protein cross-links (A) and host cell reactivation assay to study repair (B). 292

LIST OF CHARTS

Chart *Page*

Chart 1-1 Structures of some known chemical carcinogens. 5

Chart 1-2 Stereoisomers of 1,2,3,4-diepoxybutane and their relative toxicities. 17

Chart 1-3 Structures of DNA monoadducts formed by epoxide metabolites of 1,3-butadiene. 19

Chart 1-4 Structures of DNA-DNA cross-links induced by 1,2,3,4-diepoxybutane. 21

Chart 1-5 Structures of 1,2,3,4-diepoxybutane-induced exocyclic DNA adducts. 23

Chart 1-6 Chemical structures of DNA-protein cross-links. 45

Chart 1-7 Chemical structures of *bis*-alkylating antitumor drugs known to induce DNA-protein cross-links. 53

LIST OF FIGURES

Figure *Page*

Figure 1-1 DNA damage and repair. Types of DNA damage caused by common DNA damaging agents and mechanisms responsible for the repair of damaged DNA. 3

Figure 2-1 Mass spectrometry-based characterization of 1,3-butadiene-induced 2'-deoxyadenosine adducts. 92

Figure 2-2 Concentration dependent repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts by human fibrosarcoma nuclear extracts..... 96

Figure 2-3 Time dependent repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts by human fibrosarcoma nuclear extracts. 98

Figure 2-4 Repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts and known BER substrates by human fibrosarcoma nuclear extracts. 100

Figure 2-5 Comparison of repair rates of 1,3-butadiene-induced 2'-deoxyadenosine adducts against known BER substrates in human fibrosarcoma nuclear extracts..... 102

Figure 2-6 Liquid chromatography-tandem mass spectrometry analysis of repair products of 1,3-butadiene-induced 2'-deoxyadenosine adducts by human fibrosarcoma nuclear extracts..... 104

Figure 2-7 Base excision repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts by human and bacterial BER enzymes. 106

Figure 2-8 Base excision repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts by recombinant glycosylases. 107

Figure 3-1 Denaturing PAGE analysis of DNA-protein cross-links generated by post-synthetic reductive amination. 127

Figure 3-2 SDS-PAGE analysis of reductive amination-mediated DNA-protein cross-linking. 128

Figure 3-3 Effects of reaction conditions on the yields of DNA-protein cross-links. .. 130

Figure 3-4 Influence of protein identity on DPC yield for reductive amination. 134

Figure 3-5 Mass spectrometry chracaterization of AlkB-DNA cross-links. 136

Figure 3-6 Crystal structure of AlkB protein bound to double stranded DNA (PDB ID: 3BI3) showing the amino acids participating in DNA-protein cross-linking. 140

Figure 3-7 HPLC-ESI$^+$-MS/MS characterization of 7-deaza-7-(2-(*N*-acetyllysine)ethan-1-yl)-2'-deoxyguanosine (A) and 7-deaza-7-(2-(*N*-acetylarginine)ethan-1-yl)-2'-deoxyguanosine (B). 142

Figure 3-8 NanoLC-nanospary-MS^2 (A) and MS^3 (B) spectra of the peptide, Angiotensin I (DRVYIHPFHL) cross-linked to 7-deaza-7-(ethan-1-yl)-2'-deoxyguanosine. 144

Figure 4-1 Standing start assays for replication bypass of DNA-protein cross-links by hPol η, κ and ι. 163

Figure 4-2 Running start assays for replication bypass of DNA-protein cross-links by hPol η, κ and ι. 165

Figure 4-3 Standing start assays for replication bypass of DNA-peptide cross-links by hPol η and κ. 167

Figure 4-4 Running start assays for replication bypass of DNA-peptide cross-links by hPol η and κ. 168

Figure 4-5 Single nucleotide insertion opposite unmodified guanine (dG) and the 10mer peptide cross-linked to *C7* of 7-deazaguanine (10-mer pep) by TLS polymerase hPol η and κ. 170

Figure 5-1 SDS-PAGE analysis of site-specific DNA-protein cross-links generated by Cu-catalyzed azide-alkyne cycloaddition. 199

Figure 5-2 SDS-PAGE analysis of site-specific DNA-protein cross-links generated by Cu-catalyzed azide-alkyne cycloaddition. 201

Figure 5-3 Generation of site-specific DNA-peptide cross-links by Cu-catalyzed azide-alkyne cycloaddition. 203

Figure 5-4 Mass spectrometry characterization of DNA-protein conjugates. 205

Figure 5-5 Mass spectrometry characterization of DNA-peptide cross-links. 207

Figure 5-6 Purification of DNA-protein cross-links by gel electrophoresis–electroelution. 209

Figure 5-7 Replication bypass of DNA-peptide and DNA-protein conjugates of increased size adduct by human lesion bypass polymerases under standing start conditions. 212

Figure 5-8 Replication bypass of DNA-peptide and DNA-protein conjugates of increased size adduct by human lesion bypass polymerases under running start conditions. 214

Figure 6-1 Denaturing PAGE analysis of DNA-peptide cross-links generated by click chemistry.. 231

Figure 6-2 Single nucleotide insertion opposite DNA-peptide cross-links by human TLS polymerases. 234

Figure 6-3 Collision-induced dissociation mass spectra of the major extension products of *in vitro* polymerase bypass by hPol κ. 243

Figure 6-4 HPLC-ESI$^-$-FTMS analysis of primer extension by hPol κ. 244

Figure 6-5 Representative MS/MS spectra of extension products observed following *in vitro* replication past 10-mer peptide cross-linked to *C5* of thymidine by hPol κ. 245

Figure 6-6 HPLC-ESI$^-$-FTMS analysis of primer extension by hPol η. 247

Figure 6-7 Representative MS/MS spectra of extension products observed following *in vitro* replication past unmodified template by hPol η. 248

Figure 6-8 Representative MS/MS spectra of extension products observed following *in vitro* replication past 10-mer peptide cross-linked to *C5*-T by hPol η. 249

Figure 8-1 Characterization of 18-mer oligodeoxynucleotides containing site- and stereospecific 1,3-butadiene lesions after cleavage from the solid support. 271

Figure 8-2 Capillary HPLC-ESI$^-$-MS spectrum of a 11-mer, 5'-C GGA C***X***A GAA G-3', containing 1-(adenosin-N^6-yl)-4-(guan-7-yl)-butan-2,3-diol (A) and capillary HPLC-ESI$^+$-MS (B) and MS/MS spectra (C) of 1-(adenin-N^6-yl)-4-(guan-7-yl)-butan-2,3-diol. 284

Figure 8-3 Standing start assay for replication bypass of DNA-peptide cross-links by human replicative polymerase δ. 289

LIST OF ABBREVIATIONS

α-hOGG1	α isoform of human 8-oxoguanine DNA glycosylase
γ-HOPdG	3-(2'-deoxyrobo-1'-syl)-5,6,7,8-tetrahydro-8-hydroxy-pyrimido[1,2*a*] purin-10(3*H*)-one
[6-4]PP	6-4 pyrimidone photoproducts
1,N^6-α-HMHP-dA	1,N^6-(1-hydroxymethyl-2-hydroxy-1,3-propanodiyl)-2'-deoxyadenosine
1,N^6-γ-HMHP-dA	1,N^6-(2-hydroxy-3-hydroxymethyl-1,3-propanodiyl)-2'-deoxyadenosine
5F-dC	5-fluoro-2'-deoxycytosine
6-Cl-HOBt	1-hydroxy-6-chloro-benzotriazole
8-oxo-G	8-oxoguanine
AAG	alkyladenine DNA glycosylase
CAN	Acetonitrile
AGT	O^6-alkylguanine DNA alkyltransferase
ALS	amyotrophic lateral sclerosis
AP site	apurinic/apyrimidinic site
APE1	human apurinic/apyrimidinic endonuclease 1
APNG	alkyl-*N*-purine DNA glycosylase
APTX	aprataxin
As	arsenic

AT	ataxia talengiectasia
azaC	5-azacytidine
BCNU	1,3-bis(2-chloroethyl)-1-nitrosourea
BD	1,3-butadiene
BED	1-butene-3,4-diol
BER	base excision repair
bis-N^2G-BD	1,4-*bis*-(2'-deoxyguanosin-N^2-yl)-2,3-butanediol
bis-N^6A-BD	1,4-*bis*-(2'-deoxyadenosin-N^6-yl)-2,3-butanediol
bis-*N7*G-BD	1,4-*bis*-(guan-7-yl)butan-2,3-diol
BRCA2	breast cancer associated 2
C8-alkyne-dT	5-(octa-1,7-diynyl)-uracil
Carboplatin	*cis*-diammine-[1,1-cyclobutanedicarboxylato]platinum(II)
CD	circular dichroism
CHO	Chinese hamster ovary
CID	collision-induced dissociation
Cisplatin	*cis*-diamminedichloroplatinum(II)
CLP	cross-linked protein
CPD	cyclobutane pyrimidine dimers
Cr	chromium
CS	Cockayne's syndrome
CSB	cockayne syndrome A (CSA) and B
CT DNA	calf thymus DNA
dA	2'-deoxyadenosine
deaza-DHP-dG	7-deaza-7-(2,3-dihydroxypropan-1-yl)-2'-deoxyguanosine

DEB	1,2,3,4-diepoxybutane
dG	guanine
DIC	*N,N'*-diisopropylcarbodiimide
DIPEA	*N,N*-diisopropylethylamine
DMF	dimethylformamide
DNA	deoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase
Dnmt	DNA methyltransferase
DPCs	DNA-protein cross-links
Ds	double-stranded
DSB	double strand breaks
EB	3,4-epoxybut-1-ene
EBD	3,4-epoxybutan-1,2-diol
eGFP	enhanced green fluorescent prrotein
EH	epoxide hydrolase
Endo III	endonuclease III
ESI	electrospray ionization
EXO1	exonuclease 1
F	misinsertion frequency
FA	formaldehyde
FAMdT	5-[*N*-((fluoresceinyl)-aminohexyl)-3-acrylimido]-2'-deoxyuridine
FEN1	flap endonuclease 1
FTMS	fourier transform mass spectrometry

GG-NER	global genomic repair
Gh	guanidinohydantoin
hAPE-1	human apurinic/apyrimidinic endonuclease
HCTU	2-(6-chloro-*1H*-benzotriazole-1-yl)-1,1,3,3-tetramethyl-aminium hexafluorophosphate
HDnmt	HhaI DNA methyltransferase
HeLa	human cervical carcinoma
HNO_2	nitrous acid
hOGG1	human 8-oxoguanine DNA glycosylase
HPLC	high performance liquid chromatography
hPol κ	human polymerase kappa
HR	homologous recombination
HR23B	Rad23 homolog B complex
HT1080	human fibrosarcoma
IR	Infrared
KF^-	Klenow fragment of *E. coli* DNA polymerase I
LC-MS	Liquid chromatography-mass spectrometry
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LF	little finger
LIG1	DNA ligase 1
LIG3	DNA ligase 3
MGMT	O^6-methylguanine DNA methyltransferase
MMR	mismatch repair
MPG/Mpg	methylpurine glycosylase

MRN	Mre11-Rad50-Nbs1
MS	mass spectrometry
MTase	methyltransferase
MutLα	Mutator Lα
MutSα	Mutator Sα
MWCO	molecular weight cut off
MX	methoxyamine
*N1*A-*N7*G-BD	1-(aden-1-yl)-4-(guan-7-yl)butan-2,3-diol
N1-HB-dI	*N1*-(1-hydroxybut-3-en-2-yl)-2'-deoxyinosine
*N1*HX-*N7*G-BD	1-(hypoxanth-1-yl)-4-(guan-7-yl)butan-2,3-diol
N1-Me-A	*N1*-methylguanine
*N3*A-*N7*G-BD	1-(aden-3-yl)-4-(guan-7-yl)butan-2,3-diol
N3-HB-dU	*N3*-(2-hydroxybut-3-en-1-yl)-2'-deoxyuridine
N3-Me-A	*N3*-methyladenine
N^6,N^6-DHB-dA	N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine
N^6A-*N7*G-BD	1-(aden-N^6-yl)-4-(guan-7-yl)butan-2,3-diol
N^6-dA-$(OH)_2$butyl-GSH	*S*-[4-(N^6-deoxyadenosinyl)-2,3-dihydroxybutyl]glutathione
N^6-HB-dA	N^6-(2-hydroxy- but-3-en-1-yl)-2'-deoxyadenine
N^6-HB-dA II	N^6-(1-hydroxy-3-buten-2-yl)-2'-deoxyadenosine
N^6-THB-dA	N^6-(2,3,4-trihydroxybut-1-yl)-2'-deoxyadenine
*N7*A-*N7*G-BD	1-(aden-7-yl)-4-(guan-7-yl)butan-2,3-diol
N7-HB-G	*N7*-(2-hydroxybut-3-en-1-yl)guanine
NER	nucleotide excision repair

NHEJ	non-homologous end-joining
Ni	nickel
NO	nitric oxide
O^6-Me-G	O^6-methylguanine
Ogg	oxoguanine glycosylase
Oxa	oxanine
Oxaliplatin	([(1*R*,2*R*)-cyclohexane-1,2-diamine](ethanedioato-*O*,*O'*)platinum(II))
P4-1	(7-hydroxy-3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-6-hydroxymethyl-5,6,7,8-tetrahydropyrimido[1,2-*a*]purin-10(10*H*)-one
P6	7,8-dihydroxy-3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-3,5,6,7,8,9-hexahydro-1,3-diazepino [1,2-*a*]purin-11(11*H*)one
PAD	polymerase-associated domain
PAHs	polycyclic aromatic hydrocarbons
PARP1	poly ADP-ribose polymerase 1
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PDE I	phosphodiesterase I
PDE II	phosphodiesterase II
PFTase	protein farnesyltransferase
PMSF	phenylmethylsulfonyl fluoride
Pol β	DNA polymerase β
RNAPII	RNA polymerase II
RNS	reactive nitrogen species

ROS	reactive oxygen species
RPA	replication protein A
RP-HPLC	reversed phase high performance liquid chromatography
RT HIV-1	HIV-1 reverse transcriptase
SO_2	sulfur dioxide
Sp	spiroiminodihydantoin
SPO II	meiotic recombination protein
Ss	single-stranded
SSBs	single-strand breaks
STO	single-turnover
T4-pdg	T4 pyrimidine dimer glycosylase/apurinic/apyrimidinic site lyase
T4-PNK	T4 polynucleotide kinase
TBTA	tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine
TCEP	tris(2-carboxyethyl)phosphine
TC-NER	transcription-coupled repair
TDP1/Tdp1	tyrosyl DNA phosphodiesterase 1
Tdp2	tyrosyl-DNA pohosphodiesterase 2
TFA	trifluoroacetic acid
TFA	trifluoroacetic acid
TFIIH	transcription factor IIH
THBG	(2,3,4-trihydroxybut-1-yl)guanine
TLS	translesion synthesis
TOPO I	topoisomerase I

TOPO II	topoisomerase II
U	uracil
UDG	uracil DNA glycosylase
UV	Ultraviolet
WBCs	white blood cells
WT	wild-type
XP	xeroderma pigmentosum
XPC	xeroderma pigmentosum complementation group C
XRCC1	X-ray repair cross-complementing protein 1

1 LITERATURE REVIEW

1.1 DNA DAMAGE AND HUMAN DISEASES

1.1.1 DNA Modification and Its Consequences in Cells

The integrity and stability of genomic DNA are essential for life since DNA is the repository of hereditary information in cells.[1-3] DNA is not chemically inert, but possesses multiple reactive sites.[1-3] Genomic DNA is continuously modified by various endogenous and exogenous sources to form structurally altered nucleobases (DNA adducts).[1,4,5] In addition to damage caused by various chemical and physical agents,[1,3,6,7] DNA replication, the process by which genetic code is inherited to new generations of cells, can itself introduce errors compromising the cell viability.[1,3,7] DNA adducts can be recognized by cellular repair mechanisms to restore normal DNA.[8,9] Alternatively, they can be replicated *via* cellular replication machinery or DNA damage tolerance pathways.[8,9] Some DNA adducts can escape cellular repair mechanisms and persist in cells, and replication of such modified DNA can be error-prone, ultimately leading to mutations.[1,8,9]

1.1.2 Causative Agents and Types of DNA Damage

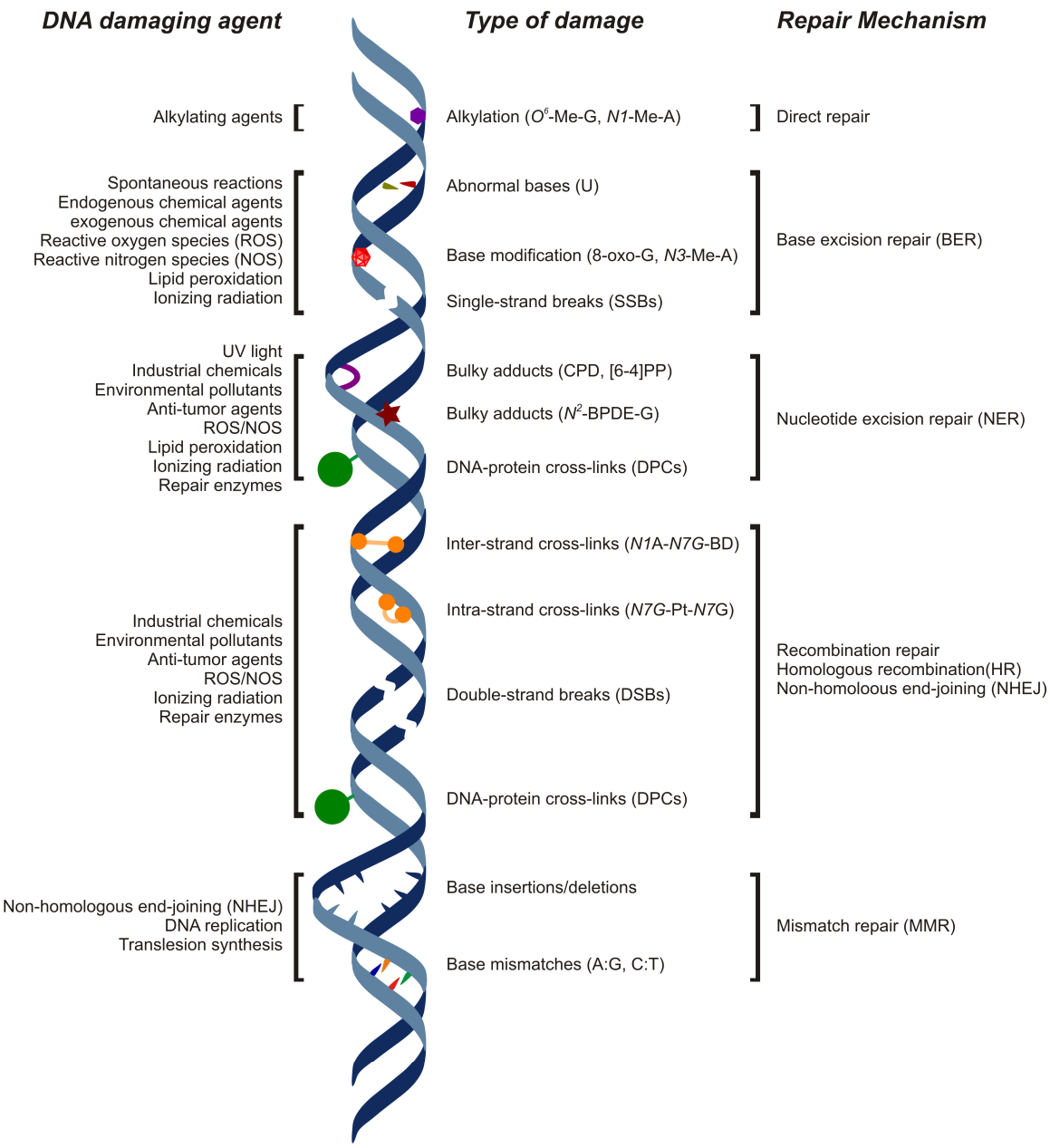
There are many causative agents that induce DNA adducts including endogenous chemicals produced inside the body by metabolic pathways, exogenous chemicals from

environmental and food sources, ionizing radiation and UV light (**Figure 1-1**).[8,10-14] DNA damaging agents can be broadly categorized into three classes: spontaneous chemical reactions occurring in cells, endogenous species generated from cellular metabolism and exogenous physicochemical agents.[8,14]

Spontaneous hydrolysis and deamination lead to DNA damage in cells.[15-17] Acid-catalyzed hydrolysis of *N*-glycosidic bond between the nucleobase and deoxyribose sugar generates apurinic/apyrimidinic (AP) or abasic sites.[15] AP sites are prone to β-elimination, resulting in single-strand DNA breaks (SSBs).[16] Deamination reactions of exocyclic amine bearing nucleosides can generate abnormal DNA bases.[15,17] For example, cytosine gets deaminated to form uracil.[15] Deamination of adenine and guanine forms hypoxanthine and xanthine, although at lower rates compared to cytosine deamination.[15]

DNA modification also occurs through endogenous reactive chemical species naturally formed in cells such as reactive oxygen species (ROS),[18-20] reactive nitrogen species (RNS)[21] and electrophilic species[15] generated during cellular metabolism. ROS like O_2^-, H_2O_2 and $OH^\bullet$ are known to generate a large number of oxidative lesions, single- and double-strand breaks (DSBs) and DNA-protein cross-links (DPCs).[18,19] Nitric oxide (NO) and its byproducts can also induce oxidative DNA damage.[21] Nitrosated amines and radicals generated by lipid peroxidation mediate alkylation of heteroatoms in DNA bases.[15]

Figure 1-1 DNA damage and repair. Types of DNA damage caused by common DNA damaging agents and mechanisms responsible for the repair of damaged DNA.



Genomic DNA undergoes constant attack from exogenous damaging agents. Ultraviolet (UV) light produce cyclobutane pyrimidine dimers (CPD) and 6-4 pyrimidone photoproducts ([6-4]PP).[22] Ionizing radiation like X-rays and γ-rays, directly or indirectly *via* ROS generation, can cause multiple types of damage including single-strand breaks (SSBs), double-strand breaks (DSBs), base modifications, DNA-DNA cross-links and DNA-protein cross-links (DPCs).[4]

Hundreds of chemical agents are also known to form DNA adducts. These include industrial and household chemicals, environmental pollutants, agrochemicals, chemicals from dietary sources, and chemotherapeutics (**Chart 1-1** and **Table 1-1**).[23,24] These chemical agents can alkylate multiple sites on DNA forming a variety of DNA monoadducts, exocyclic lesions, DNA-DNA cross-links, DSBs and DPCs. Highly toxic DNA adducts, such as DNA-DNA cross-links and DPCs, formed by some exogenous chemicals trigger apoptosis or programmed cell death.[25,26] This is the mechanism of action of most anticancer drugs including cisplatin,[27] mitomycin C,[28] 1-haloalkyl-1-nitrosoureas and nitrogen mustards.[29] Therefore, better understanding of mechanisms of DNA damage can be beneficial in preventing adverse effects of damaged DNA as well as in the development of novel therapeutics against lethal disease conditions.

Chart 1-1 Structures of some known chemical carcinogens.

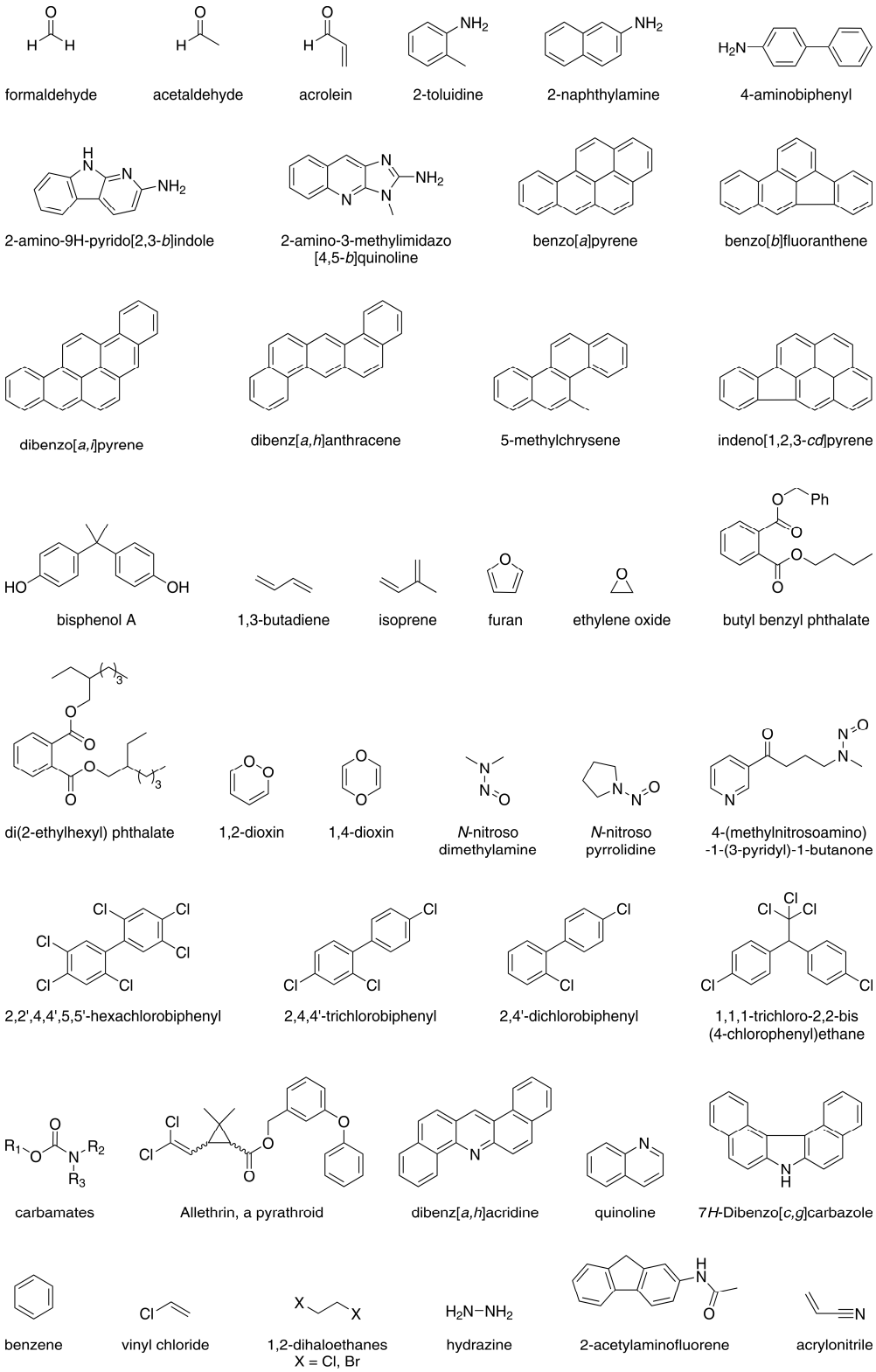


Table 1-1 Known chemical carcinogens and their carcinogenic properties.[23,24]

Carcinogen class	*Carcinogenicity*			*Examples*
	Cocarcinogen	*Mutagen*	*Promoter*	
Aldehydes		✓		formaldehyde, acetaldehyde, acrolein
Aromatic amines		✓		2-toluidine, 2-naphthylamine, 4-aminobiphenyl
N-heterocyclic amines		✓		2-amino-*9H*-pyrido[2,3-*b*]indole, 2-amino-3-methylimidazo[4,5-*b*]quinoline
Compounds of mineral origin	✓	✓		Asbestos
Polycyclic aromatic hydrocarbons (PAHs) and derivatives	✓	✓	✓	benzo[*a*]pyrene, benzofluoranthenes, dibenzopyrenes, dibenz[*a*,*h*]anthracene, 5-methylchrysene, indeno[1,2,3-*cd*]pyrene
Diphenylmethanes and bisphenols		✓	✓	bisphenol A
Miscellaneous organic compounds	✓	✓		1,3-butadiene, isoprene, furan, ethylene oxide
Phthalates		✓	✓	di(2-ethylhexyl) phthalate, butyl benzyl phthalate
Dioxins	✓	✓	✓	1,2-dioxin, 1,4-dioxin
Hormonal residues			✓	prolactin, estrogens, androgens
Metals and metalloids	✓	✓		nickel, chromium, cadmium, lead, arsenic
N-nitroso compounds		✓		*N*-nitrosodiethylamine, *N*-nitrosopyrrolidine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone

Table 1-1 (*continued from page 6*)

Carcinogen class	*Carcinogenicity*			*Examples*
	Cocarcinogen	*Mutagen*	*Promoter*	
Polychorobiphenyls (PCBs)	✓	✓	✓	2,2',4,4',5,5'-hexachlorobiphenyl, 2,4,4'-trichlorobiphenyl, 2,4'-dichlorobiphenyl
Pesticides		✓		organochlorines, carbamates, pyrethroids
Aza-arenes	✓			dibenzacridines, quinoline, 7*H*-dibenzo[*c*,*g*]carbazole
Anticancer drugs	✓			anti-estrogens such as tamoxifen, toremifen and fulvestrant, alkylating agents like nitrogen mustards, nitrosoureas, alkyl sulfonates and platinum drugs
Pharmaceuticals	✓			oral contraceptives, hormone replacement therapy
Air fine particles	✓			air carbonaceous particles
Benzene and related compounds	✓	✓		benzene
Organohalides		✓		vinyl chloride, 1,2-dichloroethane
Inorganic compounds	✓			hydrazine
Fluorene compounds	✓	✓	✓	2-acetylaminofluorene
Nitriles	✓			acrylonitrile

1.1.3 Role of DNA Damage in Human Disease

DNA adducts are implicated in various disease conditions including cancer,[30-33] cardiovascular diseases[34-36] and age-related neurodegeneration.[35-39] Mutations in genes associated with cell cycle checkpoints and chromosomal instability can lead to cancer. For example, mutations in oncogenes, like *RAS* and *MYC*, and in tumor suppressor genes, such as *TP53* and *CDKN2A*, can disturb the balance of cell proliferation and apoptosis forming malignant cells.[25,26] Chronic inflammation increases the risk of developing many human cancers.[40] The inflammatory response generates a complex spectrum of reactive species that can damage biomolecules including DNA.[41]

Genome maintenance is implicated in anti-aging.[14] It has been found that aging is accelerated in patients with segmental progeroid syndrome, whose genome maintenance is compromised.[14] Defects in several genes related to nucleotide excision repair (NER) enzymes (see section 1.3.1) cause rare hereditary diseases such as xeroderma pigmentosum, Cockayne's syndrome and trichothiodystrophy. Xeroderma pigmentosum (XP) patients, who have defects in genes *XPA–XPG*, show sun-induced pigmentation abnormalities, increased risk of skin cancer and internal cancers, and accelerated neurodegeneration.[14] Cockayne's syndrome (CS) results from the mutated *CSA* and *CSB* genes, while defects in *XPB* and *XPD* genes cause trichothiodystrophy.[14] These patients show, among other conditions, premature aging, growth and developmental arrest, and neurodysfunction.[14] These effects are attributed to increased mutations caused by the inability to repair DNA adducts.[14] Ataxia talengiectasia (AT) is characterized by

increased risk of leukemia and lymphoma, and neurologic abnormalities. AT is caused by a mutated *ATM* gene that code for a protein kinase involved in DSB repair.[42] Amyotrophic lateral sclerosis (ALS) is related to defects in *APE* gene, while Ataxia is related to defective *APTX* and *TDP1* genes.[37] Proteins transcribed from these genes are involved in base excision repair (BER) and SSB repair (see section 1.3.1).[37]

Significant damage to DNA was observed in cell culture models of heart disease and samples obtained from patients with cardiovascular disease.[43] Further, recent experimental and clinical studies suggest that oxidative DNA damage is enhanced in cardiovascular disease.[44] Taken together, these experimental and clinical evidences strongly suggest that DNA damage play a direct or indirect role in many human diseases. Although the formation of DNA adducts by endogenous sources and genetic factors are beyond our control, such adverse effects induced by exogenous agents are preventable to a certain extent.

1.2 1,3-BUTADIENE-MEDIATED MUTAGENESIS AND CARCINOGENESIS

1.2.1 1,3-Butadiene: A Potent Environmental Carcinogen

1,3-Butadiene (BD) is an important industrial chemical widely used as a raw material in synthetic rubber and plastic industry,[45] and manufacture of certain fungicides.[46] It is a colorless, highly volatile gas, which is a byproduct of ethylene manufacturing process.[45,47] Global production of BD was estimated to be ~9 million metric tons.[45,47] Occupational BD exposure levels, in 8-hour time-weighted average concentration units, can be as high as 10 ppm, while average daily concentrations in ambient air near petrochemical facilities have been reported to exceed 100 ppb.[46] BD is also an environmental pollutant found in automobile exhaust[47,48] with average concentration in urban air estimated as 1–10 ppb.[47] Further, high quantities of BD are present in cigarette smoke: 20–75 μg/cigarette and 205–360 μg/cigarette in mainstream and sidestream smoke, respectively.[25,49,50] In retrospect, these amounts are 10^2–10^6-fold higher than those of more extensively studied polycyclic aromatic hydrocarbons (PAHs, 0.6–70 ng) and nitrosamines (0.1–80 ng).[25] This widespread occurrence of BD in the environment in turn contributes to a high potential for human exposure.

The adverse biological effects of BD have been well documented. Chronic exposure to BD has been implicated in various health conditions including respiratory disorders, cardiovascular disease, and various types of cancers in both humans[51-60] and laboratory animals.[61,62] According to toxicological risk analysis, BD has the highest

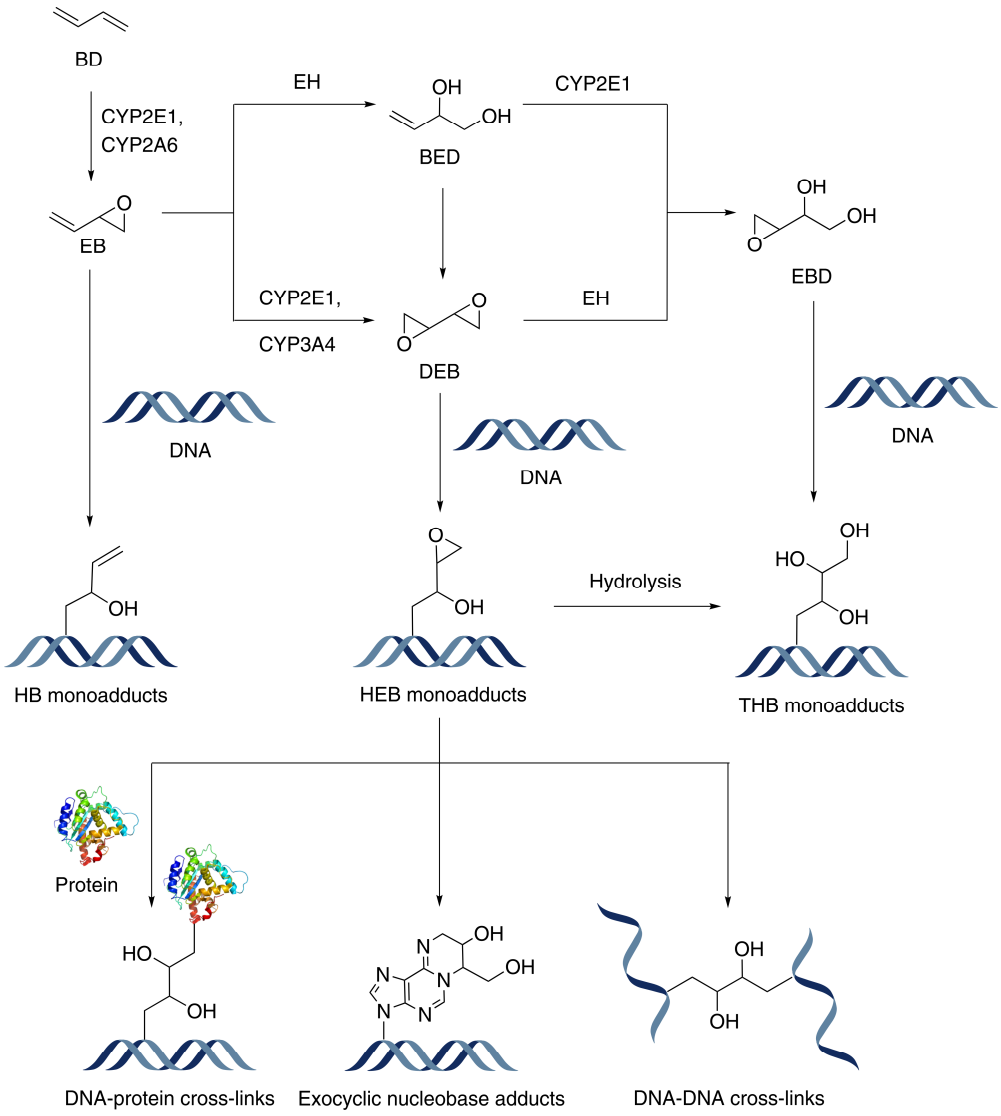
cancer risk index among all tobacco constituents.[63] Further, based on epidemiological and toxicological studies in humans and inhalation studies in laboratory animals, BD has been classified as a human carcinogen.[46,64,65] Nevertheless, the exact molecular mechanisms of BD-induced mutagenesis and carcinogenesis are yet to be fully understood.

1.2.2 Metabolic Activation of 1,3-Butadiene and DNA Adduction

BD required metabolic activation inside the body to cause any undesirable biological effects. Cytochrome P450 monooxygenases, CYP2E1 and CYP2A6, catalyze the epoxidation of a double bond of BD to form 3,4-epoxy-1-butene (EB),[66,67] which is either hydrolyzed by epoxide hydrolase (EH) to 1-buten-3,4-diol (BED)[68] or further oxidized by CYP2E1 and CYP3A4 to 1,2,3,4-diepoxybutane (DEB) (**Scheme 1-1**).[67,69] BED can be oxidized to 3,4-epoxybutan-1,2-diol (EBD) by CYP2E1, while DEB can be hydrolyzed to EBD.[70] These epoxide metabolites can be detoxified by glutathione conjugation followed by mercapturic acid pathway, and excreted from the body as the corresponding *N*-acetylcysteine conjugates.[71,72] If not deactivated, they can alkylate nucleophilic sites on cellular biomolecules to form covalent adducts.[73-77] For example, *bis*-alkylation of DNA by DEB can produce exocyclic nucleobase adducts,[78,79] DNA-DNA cross-links[80-84] and DNA-protein cross-links.[85,86] These DNA lesions can have varying toxic and mutagenic effects depending on their size, shape, hydrogen-bonding ability, and the extent of helix distortion. The generation of multiple reactive species upon BD metabolism, ability of those electrophiles to react at numerous sites

within DNA and formation of various stereo- and regioisomeric DNA lesions contribute to the BD-induced biological effects.

Scheme 1-1 Metabolic activation of 1,3-butadiene and formation of DNA adducts.*



* BD: 1,3-butadiene, BED: 1-buten-3,4-diol, CYP2E1, CYP2A6, CYP3A4: cytochrome P450 enzymes, DEB: 1,2,3,4-diepoxybutane, EB: 3,4-epoxy-1-butene, EBD: 3,4-epoxybutan-1,2-diol, EH: epoxide hydrolase, HB: 2-hydroxy-3-buten-1-yl, HEB: 2-hydroxy-3,4-epoxybutan-1-yl, THB: 2,3,4-trihydroxybutan-1-yl.

1.2.3 Adverse Biological Effects of 1,3-Butadiene

Acute and chronic exposure to BD can induce various adverse effects such as toxicity, mutagenicity, and carcinogenicity. Epidemiological studies suggest that long-term occupational exposure to BD causes chromosomal abnormalities and is associated with an increased incidence of various types of malignancies including respiratory, gastrointestinal and lymphato-hematopoietic cancers, and cardiovascular conditions such as rheumatic and arteriosclerotic heart diseases.[51-60,87] Inhalation exposure of laboratory animals to BD induces tumors at various tissues such as heart, lung, brain, liver, ovary, testis and pancreas, in addition to adverse effects on cardiovascular and respiratory systems.[60-62,87-89]

BD shows tissue specific differences in mutation spectra. For example, when B6C3F1 *laci* transgenic mice were exposed to BD by inhalation, both bone marrow and spleen showed AT→GC transitions and AT→TA transversions, while GC→AT transitions were only observed in the spleen.[90] In addition, EB induced GC→AT, AT→TA and deletion mutations in the lung of mice.[90] It is also known that DEB induces AT→TA transversions and partial deletions at *hprt* in human TK6 lymphoblasts (**Table 1-2**).[90] The less mutagenic EB induces substitutions at G:C and A:T base pairs, while EBD causes G→T mutations.[90,91] A→T transversions and frameshift mutations were also observed in workers occupationally exposed to BD.[92,93]

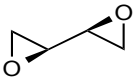
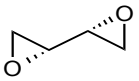
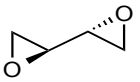
Table 1-2 *Hprt* mutation spectra of 3,4-epoxy-1-butene (EB) and 1,2,3,4-diepoxybutane (DEB) in human TK6 lymphoblast cells after 24 h exposure.[90]

Mutational class	***No. (%) of mutations***		
	Control	***400 μM EB***	***4 μM DEB***
GC → AT	10 (23)	*15 (30)*[†]	3 (6)
GC → CG	1 (2)	2 (4)	2 (4)
GC → TA	3 (7)	2 (4)	2 (4)
AT → GC	3 (7)	6 (12)	1 (2)
AT → CG	3 (7)	2 (4)	1 (2)
AT → TA	2 (5)	*12 (24)*[†]	*9 (18)*[†]
Partial deletions	1 (2)	0 (<2)	*7 (14)*[†]
Other alterations	22 (41)	12 (24)	40 (59)

[†] significant increase compared to control.

According to experimental data on genotoxicity and mutagenicity of BD metabolites, DEB is 50- and 100-fold more genotoxic and mutagenic than EB and EBD, respectively.[94-96] It induces sister chromatid exchanges and chromosomal aberrations.[90,97,98] All three stereoisomers of DEB (*S,S*, *R,R* and *meso*) are formed *in vivo*,[69] and they have varying levels of toxicities and mutagenicity (**Chart 1-2**). Among these, (*S,S*)-DEB is the most genotoxic followed by (*R,R*)- and *meso*-DEB.[99-101] This variability in toxicity is attributed to different DNA adducts formed. (*S,S*)-isomer produces the largest number of 1,3-interstrand DNA cross-links, which are highly toxic due to their ability to block DNA repair, replication and transcription, whereas *meso*-DEB forms both 1,3-interstrand and 1,2-intrastrand cross-links (**Chart 1-2**).[83,102]

Chart 1-2 Stereoisomers of 1,2,3,4-diepoxybutane and their relative toxicities.[83,102]

Stereoisomer	***Percentage of DNA-DNA cross-links***			***Relative toxicity***
	1,3-interstrand	***1,2-interstrand***	***1,2-intrastrand***	
 *S*,*S*-DEB	96	0	4	most toxic
 *R*,*R*-DEB	68	13	19	moderate
 *meso*-DEB	49	0	51	least toxic

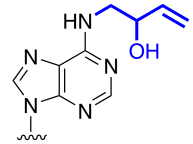
1.2.4 DNA Lesions of 1,3-Butadiene

Upon activation *in vivo*, BD metabolites can form a range of DNA adducts, protein adducts[103-107] and urinary metabolites.[71,72,108] These can be used as biomarkers of exposure to BD. More importantly, covalent DNA adducts can be extremely useful in understanding the structures and mechanisms responsible for the adverse biological effects of BD.

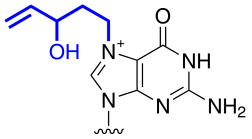
DNA Monoadducts

A significant number of DNA lesions caused by epoxide metabolites of BD has been identified.[109] Most of the observed adducts involve nucleophilic attack by BD metabolites at the *N7* position of guanine, *N3* of thymidine, and the *N1*, *N3* and N^6 of adenine (**Chart 1-3**).[73,74,110] Amongst the DNA monoadducts observed are *N3* and *N7*-(2,3,4-trihydroxybutan-1-yl)-2'-deoxyguanine (THB-dG), N^6-(2,3,4-trihydroxybutan-1-yl)-2'-deoxyadenine (N^6-THB-dA), *N7*-(2-hydroxy-3-buten-1-yl)-2'-deoxyguanine (*N7*-HB-dG), *N3*-(2-hydroxy-3-buten-1-yl)-2'-deoxyuridine (*N3*-HB-dU), N^6-(2-hydroxy--3-buten-1-yl)-2'-deoxyadenine (N^6-HB-dA) and their regio- and stereo-isomers (**Chart 1-3**).[73-77]

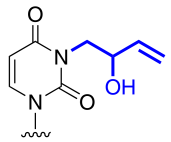
Chart 1-3 Structures of DNA monoadducts formed by epoxide metabolites of 1,3-butadiene.



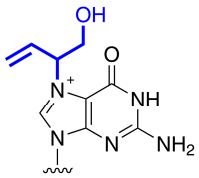
N^6-(2-hydroxybut-3-en-1-yl) 2'-deoxyadenine (***N*^{6}*-HB-dA***)



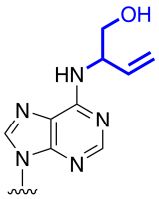
N7-(2-hydroxybut-3-en-1-yl) guanine (***N7-HB-dG***)



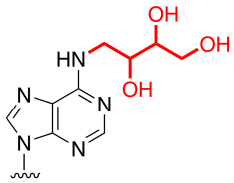
N3-(2-hydroxybut-3-en-1-yl) 2'-deoxyuridine (***N3-HB-dU***)



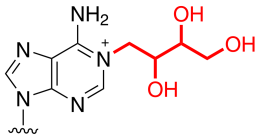
N7-(1-hydroxybut-3-en-2-yl) guanine (***N7-HB-dG II***)



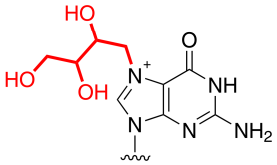
N^6-(1-hydroxybut-3-en-2-yl) 2'-deoxyadenine (***N*^{6}*-HB-dA II***)



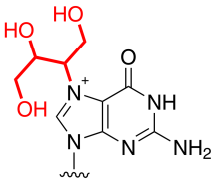
N^6-(2,3,4-trihydroxybut-1-yl) 2'-deoxyadenine (***N*^{6}*-THB-dA***)



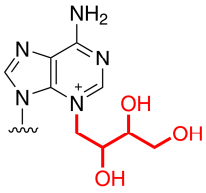
N1-(2,3,4-trihydroxybut-1-yl) adenine (***N1-THB-dA***)



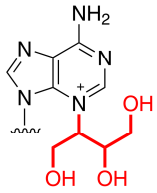
N7-(2,3,4-trihydroxybut-1-yl) guanine (***N7-THB-dG***)



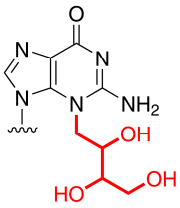
N7-(1,3,4-trihydroxybut-2-yl) guanine (***N7-THB-dG II***)



N3-(2,3,4-trihydroxybut-1-yl) adenine (***N3-THB-dA***)



N3-(1,3,4-trihydroxybut-2-yl) guanine (***N3-THB-dA II***)

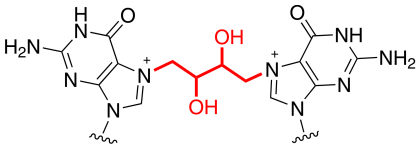


N3-(2,3,4-trihydroxybut-1-yl) 2'-deoxyguanine (***N3-THB-dG***)

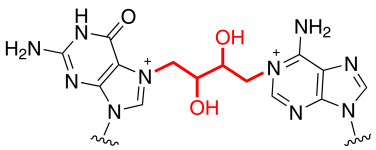
DNA Cross-Links of 1,2,3,4-Diepoxybutane

The higher genotoxicity of DEB compared to other BD metabolites is attributed to its ability to form bifunctional lesions. Both interstrand and intrastrand DNA-DNA cross-links contribute heavily to the mutational spectra of the BD diepoxide. Although 1,4-*bis*-(guan-7-yl)butan-2,3-diol (*bis*-*N7*G-BD) is the predominant DNA-DNA cross-link found *in vivo* and *in vitro*,[80,81] many regioisomeric G-A cross-links are also observed (**Chart 1-4**). 1-(Aden-1-yl)-4-(guan-7-yl)butan-2,3-diol (*N1*A-*N7*G-BD) and its deaminated product, 1-(hypoxanth-1-yl)-4-(guan-7-yl)butan-2,3-diol (*N1*HX-*N7*G-BD), 1-(aden-3-yl)-4-(guan-7-yl)butan-2,3-diol (*N3*A-*N7*G-BD), 1-(aden-N^6-yl)-4-(guan-7-yl) butan-2,3-diol (N^6A-*N7*G-BD) and 1-(aden-7-yl)-4-(guan-7-yl)butan-2,3-diol (*N7*A-*N7*G-BD) are some examples (**Chart 1-4**).[82-84] Further, mass spectrometry-based proteomics have revealed that DEB can form DNA-protein cross-links with numerous proteins having various cellular functions including DNA repair, replication, transcription and translation.[85] These super bulky DNA-protein adducts may have significant toxic effects because of their ability to block DNA replication, transcription and repair (see section 1.5.5).

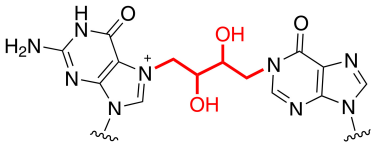
Chart 1-4 Structures of DNA-DNA cross-links induced by 1,2,3,4-diepoxybutane.



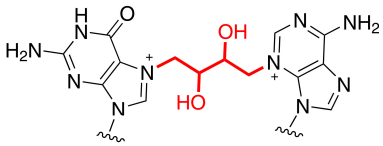
1,4-*bis*-(guan-7-yl)butan-2,3-diol
(***bis-N7G-BD***)



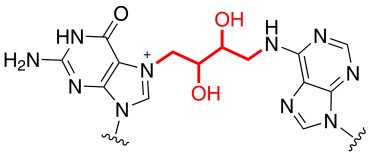
1-(aden-1-yl)-4-(guan-7-yl)butan-2,3-diol
(***N1A-N7G-BD***)



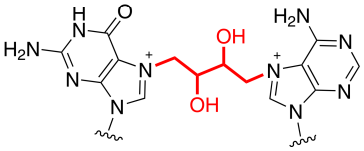
1-(hypoxanth-1-yl)-4-(guan-7-yl)butan-2,3-diol
(***N1HX-N7G-BD***)



1-(aden-3-yl)-4-(guan-7-yl)butan-2,3-diol
(***N3A-N7G-BD***)



1-(aden-N^6-yl)-4-(guan-7-yl)butan-2,3-diol
(***N^6A-N7G-BD***)

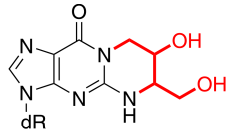


1-(aden-7-yl)-4-(guan-7-yl)butan-2,3-diol
(***N7A-N7G-BD***)

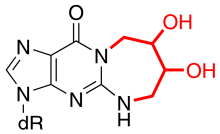
DNA Exocycles

As discussed above, DEB has the ability to alkylate two nucleophilic sites within the same nucleobase to form exocyclic lesions (**Chart 1-5**). DEB forms 1,N^2 exocyclic lesions with guanine, (7-hydroxy-3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-6-hydroxy-methyl-5,6,7,8-tetrahydropyrimido[1,2-*a*]purin-10(10*H*)-one (P4-1) and 7,8-dihydroxy-3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-3,5,6,7,8,9-hexahydro-1,3-diazepino[1,2-*a*]purin-11(11*H*)one (P6).[79] However, these guanine exocycles have not been observed *in vivo*. If formed in cells, they can potentially be mutagenic due to their hydrolytic stability and disruption of normal Watson-Crick base pairing. In addition, Tretyakova and coworkers have observed that adenine forms 1,N^6-(1-hydroxymethyl-2-hydroxypropan-1,3-diyl)-2'-deoxyadenosine (1,N^6-α-HMHP-dA), 1,N^6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl) -2'-deoxyadenosine (1,N^6-γ-HMHP-dA), and N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine (N^6,N^6-DHB-dA) (**Chart 1-5**).[78] Both 1,N^6-HMHP-dA regioisomers have been detected in tissues of rodents exposed to BD,[111] and *in vitro* replication studies have provided evidence for the potential mutagenicity of 1,N^6-HMHP-dA.[112]

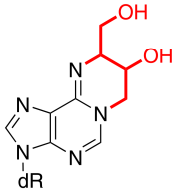
Chart 1-5 Structures of 1,2,3,4-diepoxybutane-induced exocyclic DNA adducts.



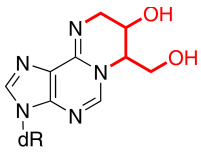
(7-hydroxy-3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-6-hydroxymethyl-5,6,7,8-tetrahydropyrimido[1,2-*a*]purin-10(10*H*)-one (***P4-1***)



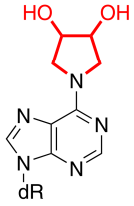
7,8-dihydroxy-3-(2-deoxy-β-D-*erythro*-pentofuranosyl)-3,5,6,7,8,9-hexahydro-1,3-diazepino[1,2-*a*]purin-11(11*H*)one (***P6***)



1,N^6-(1-hydroxymethyl-2-hydroxy propan-1,3-diyl)-2'-deoxyadenosine (***1,N^6-α-HMHP-dA***)



1,N^6-(2-hydroxy-3-hydroxymethyl propan-1,3-diyl)-2'-deoxyadenosine (***1,N^6-γ-HMHP-dA***)



N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine (***N^6,N^6-DHB-dA***)

1.3 REPAIR OF 1,3-BUTADIENE-INDUCED DNA ADDUCTS

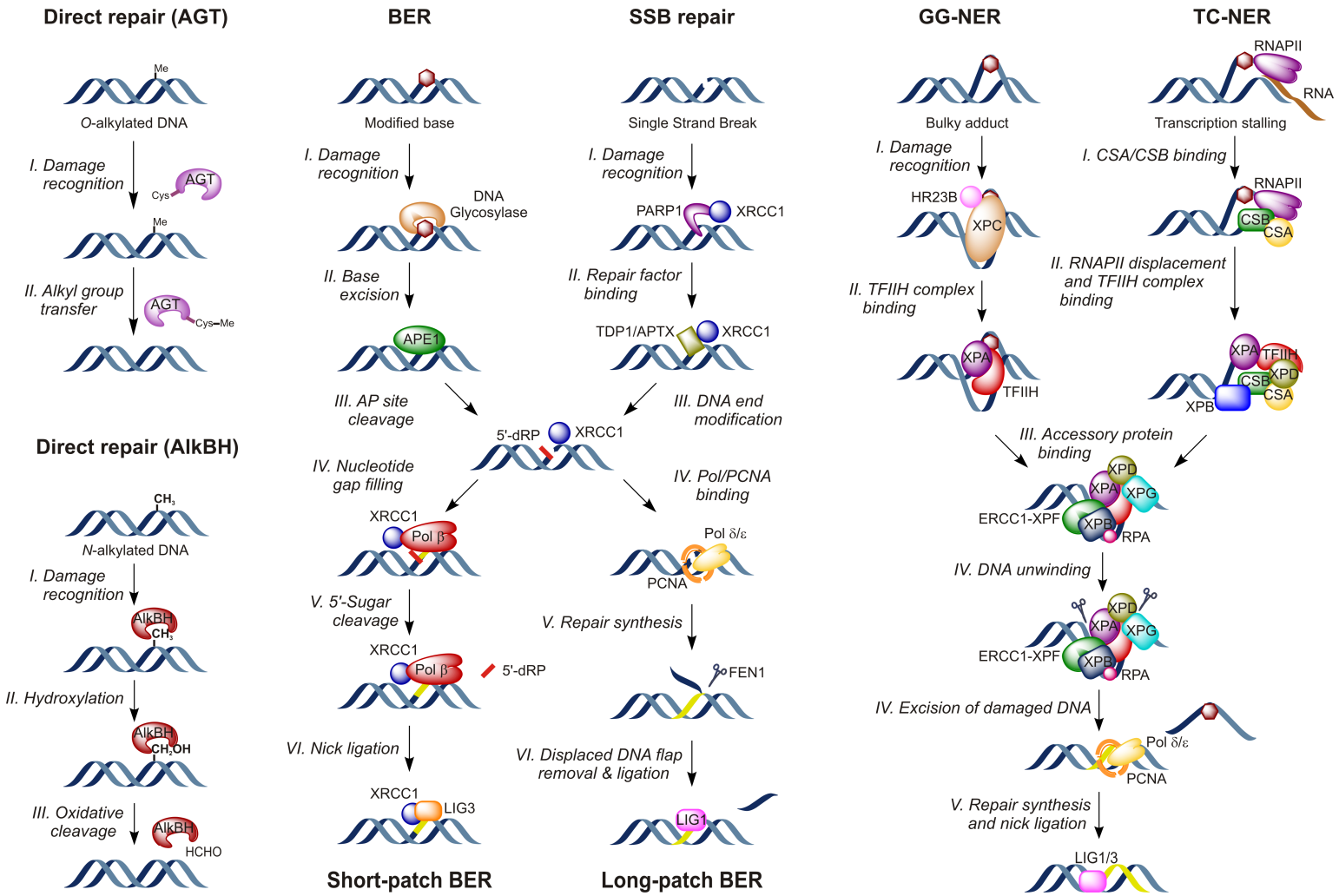
1.3.1 Cellular DNA Repair Mechanisms

All living cells have hundreds of genes dedicated to DNA repair. These cellular repair mechanisms are in constant watch to restore normal DNA and to mitigate the effects of DNA damage.[1] Multiple repair pathways have evolved to recognize damaged DNA in eukaryotic systems (**Schemes 1-2** and **1-3**), including direct repair,[113,114] base excision repair (BER),[13,113-115] nucleotide excision repair (NER),[13,113-118] mismatch repair (MMR),[7,113,114] homologous recombination (HR),[13,114,119] and non-homologous end-joining (NHEJ).[13,114,117,120] One or more factors including size and shape of the DNA lesion, helix distorting nature and their thermodynamic instability can play a role to which repair pathways are involved in identifying the damaged DNA.[1,121]

Direct Repair

Direct repair is the reversal of an alkylated nucleobase to a normal base without cleaving the phosphate backbone (**Scheme 1-2**).[113,114] For example, O^6-alkylguanine DNA alkyltransferase (AGT, also known as O^6-methylguanine DNA methyltransferase, MGMT) transfers the alkyl group from the O^6-alkylguanine onto Cys^{145} in its active site.[113,114] *E. coli* AlkB and its homologues can repair alkylated bases like *N1*-Me-dA and *N3*-Me-dC in an oxygen, ketoglutarate and Fe(II)-dependent oxidative mechanism.[113,114]

Scheme 1-2 Cellular repair mechanisms: direct repair, base excision repair and nucleotide excision repair.



Base Excision Repair

BER is the principal mechanism responsible for the removal of alkylated or incorrect DNA bases and oxidative lesions, which can be recognized by specific enzymes called DNA glycosylases (**Scheme 1-2**).[6,114,115] Type I/monofunctional glycosylases only possess the glycosylase activity, while type II/bifunctional glycosylases have both glycosylase and 3'-endonuclease activity.[6] Uracil DNA glycosylases (UDGs), methyl purine glycosylase (MPG) and adenine-specific mismatch DNA glycosylase (MYH) are human type I/monofunctional enzymes, while oxoguanine glycosylase (OGG1), NTH1, NEIL1 and NEIL2 are examples of human type II glycosylases.[6] Single-strand breaks (SSBs) and apurinic/apyrimidinic (AP) sites resulting from spontaneous hydrolysis are also processed by BER.[8,117] In the case of base modifications, DNA glycosylases hydrolyze the *N*-glycosidic bond *via* a base-flipping mechanism to excise the damaged base.[6,8,13,113-115] The resulting AP site is cleaved by AP endonuclease 1 (APE1) in the case of type I glycosylases or by 3'-endonuclease activity of type II enzymes to create a single-strand break.[6,8,13,113-115] In the case of SSBs, poly ADP-ribose polymerase 1 (PARP1) accumulation at the strand break promotes the binding of X-ray repair cross-complementing protein 1 (XRCC1).[8,117] XRCC1 recruits either of the repair factors, tyrosyl DNA phosphodiesterase 1 (TDP1) or Aprataxin (APTX), to modify the DNA ends for ligation.[117] In short-patch BER, DNA polymerase β (Pol β) inserts a nucleotide to fill the gap and cleaves the 5'-deoxyribose phosphate.[6,8,13,114,115] DNA ligase 3 (LIG3) then seals the remaining nick.[6,8,13,114,115] Long-patch BER recruits Pol δ/ε and

proliferating cell nuclear antigen (PCNA) to carry out repair synthesis by adding 2–10 nucleotides, flap endonuclease 1 (FEN1) to remove the displaced DNA flap, and DNA ligase 1 (LIG1) to seal the nick.[6,8,13,114,115]

Nucleotide Excision Repair

A wide range of bulky, helix-distorting lesions such as UV-induced photolesions and intrastrand cross-links are removed by NER.[114-116] NER consists of two sub-pathways that differ in the damage recognition step: global genomic repair (GG-NER) and transcription-coupled repair (TC-NER) (**Scheme 1-2**). GG-NER repairs helix-distorting DNA lesions located throughout the genome, while TC-NER eliminates lesions located on the coding strand of actively transcribed genes.[114-117] In GG-NER, xeroderma pigmentosum complementation group C and Rad23 homolog B complex (XPC-HR23B) or XPE binds to the damage site.[8,114-118] TC-NER is triggered, when RNA polymerase II (RNAPII) is stalled upon encountering the DNA damage. Cockayne syndrome A (CSA) and B (CSB) displaces the RNAPII to initiate TC-NER.[8,114-117] Subsequent stages of both pathways are similar. Transcription factor IIH (TFIIH) is recruited at the damage site,[8,114-118] while XPA confirms the presence of a damage by probing for abnormal backbone structure.[8,116,118] DNA dependent ATPase, XPB and XPD helicase of TFIIH complex locally unwind DNA,[13,114-118] and ssDNA binding replication protein A (RPA) binds to the undamaged strand, stabilizing the open intermediate.[8,116,118] XPG cleaves 3' to the damage site, while ERCC1-XPF performs the 5' incision to remove a ~24–32 nucleotide stretch containing the damage.[8,13,113-118] Repair synthesis is then carried out by

Pol δ/ε with PCNA using the undamaged strand as a template, and LIG seals the nick.[13,113-118]

Homologous Recombination

HR is a high fidelity, template dependent repair mechanism involved in repairing DNA double-strand breaks (DSBs), DNA gaps, and interstrand DNA-DNA cross-links in proliferating cells (**Scheme 1-3**).[13,114,119] 5'→3' Exonuclease activity of Mre11-Rad50-Nbs1 (MRN) complex creates 3'-overhangs at the break site.[8,13,114,119] RPA binds the single-stranded overhangs to remove disruptive secondary structures, and along with Rad52 facilitate the assembly of Rad51 nucleoprotein filament that includes mediator proteins such as Rad54, breast cancer associated 2 (BRCA2) and Rad51 paralogs.[8,13,114,119] Rad54 ATPase-mediated homology search follows the strand invasion by Rad51 recombinase.[8,13,114,119] In this process, homologous strand from the sister chromatid is inserted into the damaged chromatid, which acts as an error-free template for DNA synthesis.[8,13,114,119] Pol η replicates from the 3'-end of the invading strand, and subsequent ligation by LIG1 forms a four-way junction called a Holliday junction. The Holliday junction is resolved into crossover or non-crossover products by resolvases such as BLM DNA helicase and type 1 topoisomerase TOPOIII α complex.[119]

Non-Homologous End-Joining

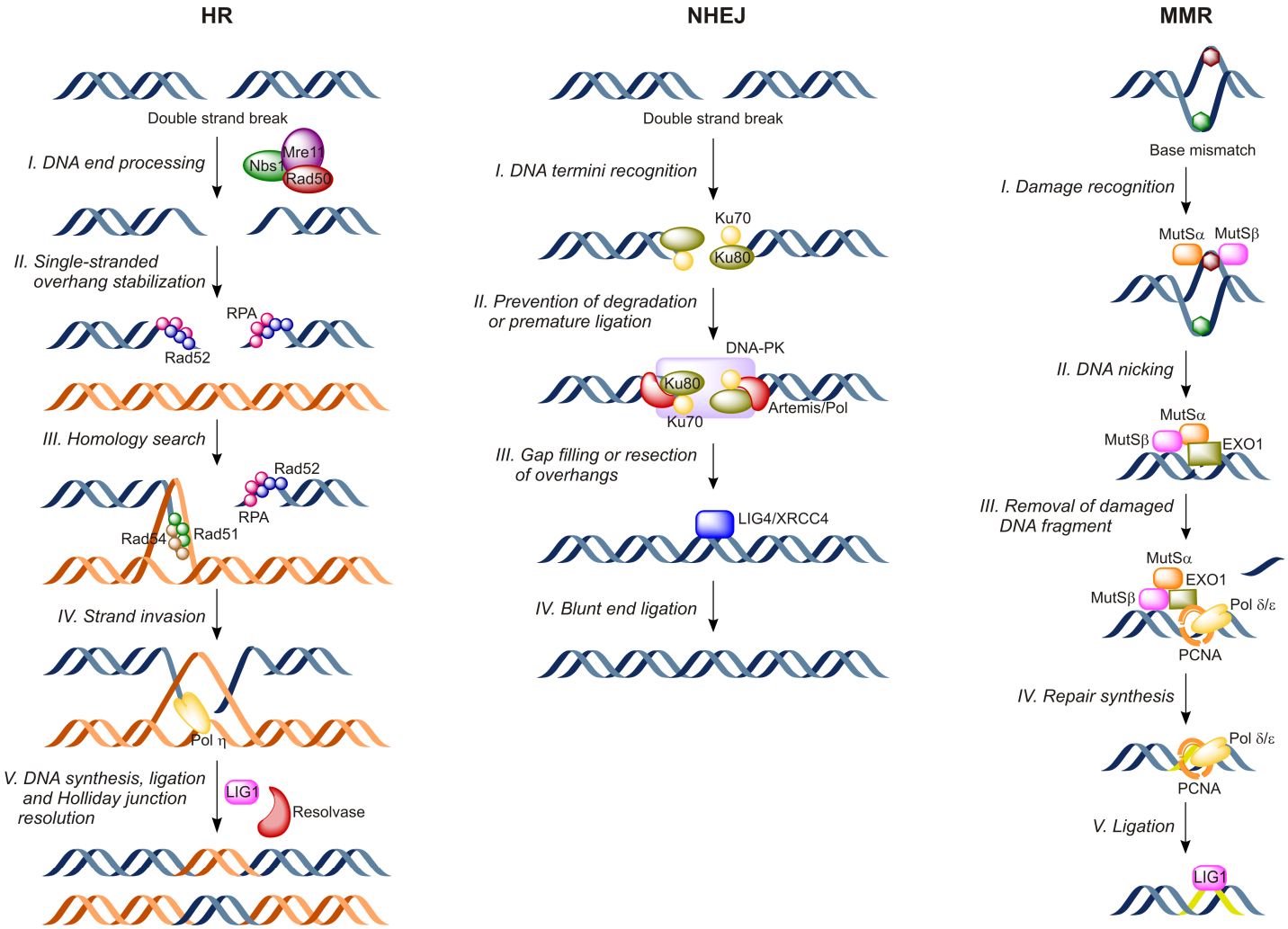
NHEJ directly ligates two ends of DSBs without any sequence homology information (**Scheme 1-3**), and sometimes results in nucleotide loss, affecting genome integrity.[8,114,120] DNA termini recognition and binding of Ku70/Ku80 heterodimer attracts

the catalytic subunit of DNA-dependent protein kinase (DNA-PK) to Ku encircled DNA duplex.[8,13,114,117,120] DNA-PK protects degradation and premature ligation of the DNA termini, and recruits either polymerases for gap filling or Artemis for resection of single-stranded overhangs.[120] Blunt ends are then ligated by LIG4/XRCC4.[8,13,114,117,120]

Mismatch Repair

MMR pathway is mainly responsible for eliminating misinsertions that have escaped proofreading mechanism during replication and fixes insertion/deletion loops resulting from polymerase slippage events (**Scheme 1-3**).[7,114] It also repairs base mismatches caused by deamination, oxidation, and alkylation.[114] Mutator Sα (MutSα) complex recognizes the mismatch and recruits the MutLα heterodimer to the damaged site.[7,8,113,114] Diffusion of MutSα-MutLα leads to nicking of the unmethylated daughter strand either upstream or downstream of the mismatch followed by the removal of a fragment of DNA by exonuclease 1 (EXO1).[7,8,113,114] Pol δ/ε with PCNA fills the gap by repair synthesis and subsequent nick sealing by LIG1.[7,8,113,114]

Scheme 1-3 Cellular repair mechanisms: homologous recombination, non-homologous end-joining and mismatch repair.



1.3.2 In Vivo and In Vitro Persistence and Repair Studies

Information on repair and persistence of BD-DNA adducts are crucial for the assessment of lesions contributing to the adverse effects of BD *in vivo*. Wickliffe *et al.* investigated the role of GG-NER on repairing BD-induced DNA adducts.[122,123] *Xpc-null* mice exposed to BD or its epoxide metabolite, EB showed ~2–3-fold higher *Hprt* mutation frequencies as compared to wild-type mice.[122,123] In another study, human bronchial epithelial cells treated with an organic extract of BD soot revealed overexpression of α isoform of human 8-oxoguanine DNA glycosylase (α-hOGG1) and human apurinic/apyrimidinic endonuclease (hAPE-1).[124] These studies provide initial evidence for the repair of BD-DNA adducts by NER and BER pathways, although they suffer from the lack of information about the structures of actual DNA adducts.

A limited number of studies is available on repair of specific BD-induced DNA adducts. Goggin *et al.* quantified 1,N^6-HMHP-dA, *bis*-*N7*G-BD and *N7*G-*N1*A-BD in wild-type, *Mpg*-deficient and *Xpa*-deficient mice exposed to BD using isotope dilution liquid chromatography-tandem mass spectrometry.[111] The half-life ($t_{½}$) of *bis*-*N7*G-BD was 2.3–5.7 days, while $t_{½}$ of *N7*G-*N1*A-BD and 1,N^6-HMHP-dA were estimated to be 36–42 days.[111] No differences in number of exocyclic adducts or either of the cross-links were observed between the wild-type and *Mpg*-null or *Xpa*-null mice.[111] Methyl purine glycosylase (MPG), also known as alkyl-*N*-purine DNA glycosylase (APNG) or alkyladenine DNA glycosylase (AAG), is a BER enzyme (see section 1.3.1) known to recognize alkylated purines including 1,N^6-etheno-dA, which is structurally analogous to

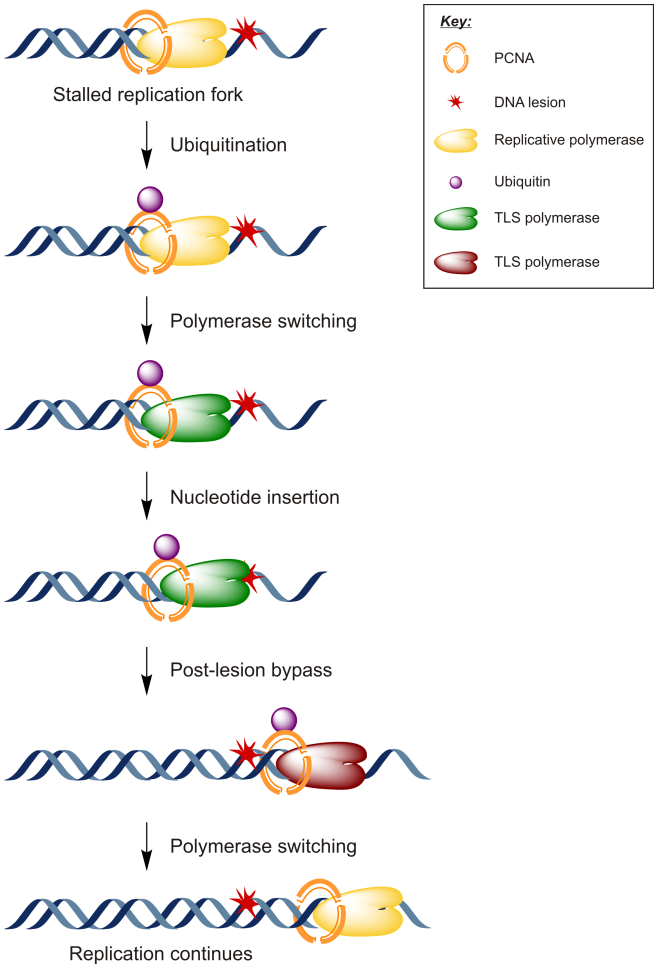
1,N^6-HMHP-dA.[125,126] As discussed in section 1.3.1, XPA is a protein involved in lesion recognition/verification in NER pathway.[116,118,127] *In vitro* repair assays using site-specifically adducted 50-mer DNA oligomers containing 1,4-*bis*-(2'-deoxyguanosin-N^2-yl)-2,3-butanediol (*bis*-N^2G-BD) have revealed that this intrastrand cross-link is not recognized by *E. coli uvrABC* complex.[128] These results provide evidence against the involvement of MPG-mediated BER or NER mechanisms for repair of these BD-DNA adducts. Nonetheless, the potential roles of other BER enzymes in BD adduct repair have not been investigated, and NER studies have not been conducted for the majority of the adducts.

1.4 REPLICATION OF 1,3-BUTADIENE-MEDIATED DNA ADDUCTS

1.4.1 DNA Adducts and Translesion Synthesis

Structural modifications of DNA nucleobases can alter their interactions with other biomolecules involved in cellular homeostasis. In some cases, adducted DNA can evade cellular repair mechanisms leading to adduct accumulation in tissues.[129] Because of their altered size and shape as compared to native bases, these adducts may not be accommodated in the compact active site of the high fidelity human replicative polymerases like hPol δ and ε.[130,131] So upon encountering the adduct, the replication fork stalls.[131,132] During such instances, the replication machinery recruits a specialized group of low fidelity polymerases called translesion synthesis (TLS) polymerases to the blocked replication fork (**Scheme 1-4**).[130,133,134] Ubiquitination of sliding clamp (PCNA in humans) increases the affinity of TLS polymerases for the replication fork, and the replicative polymerase is switched with a TLS polymerase.[130,131,135] Nucleotide insertion opposite the adduct is followed by the extension past the lesion by the same or a different TLS polymerase (**Scheme 1-4**).[131] The identity of TLS polymerase employed is dependent on the specific lesion, while the fidelity is determined by the nature of the lesion, as well as the polymerase employed.[131] TLS polymerases have low processivity, leading to their dissociation from the replication fork, and the replicative polymerase takes over to complete DNA synthesis (**Scheme 1-4**).[130,131,136]

Scheme 1-4 Polymerase switching model for translesion synthesis.



Human Y family polymerases hPol η, κ, ι and Rev1, human B family polymerase ξ, and human A family polymerase ν are the main TLS polymerases known to be involved in replication bypass of a range of DNA adducts.[131,136-142] *E. coli* Pol IV and Pol V, and *S. solfataricus* DPO4 are amongst the most-studied prokaryotic TLS polymerases, while *S. cerevisiae* Pol η, Rev1 and ξ belong to eukaryotic TLS polymerases.[130,131,136,143,144] Processivity, fidelity, and efficiency of lesion bypass polymerases are extremely low. For example, the error rate for TLS polymerases is one misinsertion for every 10^1–10^4 bases, while that of replicative polymerases is one incorrect base for every 10^6–10^8 bases.[139,145] These features are directly related to their structures. TLS polymerases have significantly smaller thumb and finger domains as compared to replicative polymerases, resulting in a more open and solvent accessible active site.[131,142,144,146] These polymerases also have a unique wrist domain called as little finger (LF) or polymerase-associated domain (PAD) that extends from the finger domain.[129,131,139,144,146-148] This large and flexible active site enables for the accommodation of bulky DNA adducts, while PAD domain makes additional contacts with DNA and is implicated in lesion specificity.[129,131,139,144,146-148] The interactions between the DNA major groove and PAD, and the orientation of DNA and PAD relative to the active site can alter processivity and fidelity of a TLS polymerase.[139,146,148]

The accuracy of nucleotide incorporation by replicative polymerases is further enhanced by an induced fit mode of replication and proof reading activity. Upon regular Watson-Crick base pairing of the incoming nucleotide and the template base, a conformational change in the finger domain of the replicative polymerases allows the

new base pair to be accommodated in a closed active site.[144,146] However, induced fit conformation is not achieved when a DNA adduct is encountered, thus stalling the replication fork.[144,146] In contrast, TLS polymerases have an open and preformed active site and as a result tolerate the non-Watson-Crick base pairing along with distorted primer-template complexes.[143,144,146] In addition, bypass polymerases lack the intrinsic 3'→5' proof reading activity,[129,130,139,143,149,150] reducing the replication fidelity by 2–3 orders of magnitude.[136,139]

1.4.2 In Vivo Replication and Mutagenesis Studies of 1,3-Butadiene DNA Adducts

As discussed in sections 1.2.3 and 1.2.4, BD forms a vast array of DNA adducts and induces a range of mutations, including A→T transversions and frameshift mutations. However, the structures and mechanisms responsible for these mutagenic effects are not fully understood. *N7*-G adducts are the most abundant form of BD-DNA lesions observed *in vivo*.[80,111,151,152] However, due to the rapid hydrolysis of *N7*-G lesions and synthetic challenges in preparing hydrolytically stable *N7*-G adducted DNA, mutagenicity of these lesions has not been investigated. On the other hand, a limited number of studies has been reported on replication and mutagenicity of less abundant but hydrolytically stable BD-DNA adducts (**Table 1-3**).[70,91,128,153,154]

Table 1-3 Summary of *in vivo* mutagenicity of 1,3-butadiene-induced DNA lesions.[70,91,128,153,154]

DNA adduct (X)	***Mutagenicity***	***Total mutation yield***‡	***Mutations (%)***			
			X→A	X→C	X→G	X→T
R-N^6-HB-dA II	non-mutagenic	N/A	-	-	-	-
S-N^6-HB-dA II	non-mutagenic	N/A	-	-	-	-
R,R-N^6-THB-dA	very weak	0.1%	-	0	*100*	0
S,S-N^6-THB-dA	very weak	0.2%	-	*100*	0	0
R-N^2-HB-dG	strongly blocking	0.01%	*50*	*50*	-	0
S-N^2-HB-dG	strongly blocking	0.2%	32	23	-	*45*
R,R-N^2-THB-dG	strongly blocking	0.03%	30	30	-	*40*
S,S-N^2-THB-dG	strongly blocking	0.07%	*35*	*35*	-	30
N3-HB-dU	extremely high	96.9%	34	-	11	*55*
R-N1-HB-dI	high	59%	-	12	*82*	6
S-N1-HB-dI	extremely high	94.5%	-	10	*84*	6
R,R-bis-N^6A-BD	high	54%	-	17	*74*	9
S,S-bis-N^6A-BD	mutagenic	19.4%	-	4	*67*	29
R,R-bis-N^2G-BD§	strongly blocking	0.1%	27	7	-	*66*
S,S-bis-N^2G-BD§	strongly blocking	0.6%	*44*	28	-	28

‡ total mutation yield = no. of mutagenic colonies/total no. of colonies×100.
§ deletion mutations were observed.

In vivo replication studies of N^6-adenine and N^2-guanine adducts derived from EB and DEB have been conducted.[70,91] Synthetic 11-mer oligodeoxynucleotides containing site-specific *R* and *S*-N^6-(1-hydroxy-3-buten-2-yl)-2'-deoxyadenosine (N^6-HB-dA II), or *R,R* and *S,S*-N^6-(2,3,4-trihydroxybutan-1-yl)-2'-deoxyadenosine (N^6-THB-dA) within the *N-ras* codon 61 were ligated into single-stranded (ss) M13mp7L2 plasmids.[70] Repair-deficient (*uvrA*$^-$, *recA*$^-$) *E. coli* were transfected with the adducted plasmids, and resulting plaques were screened for mutations. No effect on plaque forming ability was observed, suggesting these adducts do not significantly block *in vivo* replication.[70] N^6-HB-dA II isomers were non-mutagenic, whereas stereoisomer-specific mutation spectra were observed for N^6-THB-dA adducts.[70] *R,R*-N^6-THB-dA induced A→G transitions, while S,S-isomer induced A→C transversions.[70] Similar study conducted using site-specific *R* and *S*-N^2-HB-dG, and *R,R* and *S,S*-N^2-THB-dG within *N-ras* 12 codon revealed that all four adducts cause G→A and G→C mutations, while all but *R*-N^2-HB-dG also induced G→T mutations, although mutagenic frequencies were low.[91] In contrast, replication of *N3*-(2-hydroxy-3-buten-1-yl)-2'-deoxyuridine (*N3*-HB-dU) and stereoisomeric *N1*-(1-hydroxy-3-buten-2-yl)-2'-deoxyinosine (*N1*-HB-dI) adducted ss-pMS2 vectors in COS-7 monkey fibroblasts was extremely error-prone.[153,154] *N3*-HB-dU induced all possible types of base substitutions, with an overall mutation yield of ~97%.[153] Similarly, *R*- and *S*-*N1*-HB-dI produced 59% and ~95% total mutations, respectively. The majority of mutations were A→G transitions.[154] Overall, these results provide evidence for highly mutagenic nature of BD-DNA adducts and suggest that factors such as site of adduction and carcinogen stereochemistry can affect the mutagenicity.

DNA-DNA cross-links are expected to hinder replication and to cause a significant amount of mutations. Indeed, 1,4-*bis*-(2'-deoxyadenosin-N^6-yl)-2,3-butanediol (*bis*-N^6A-BD) and 1,4-*bis*-(2'-deoxyguanosin-N^2-yl)-2,3-butanediol (*bis*-N^2G-BD) decreased plaque forming ability and increased mutagenic frequency.[128,154] Mutation spectra from diastereomeric *bis*-N^6A-BD containing ss-pMS2 plasmids transfected in mammalian cells revealed that these adducts cause primarily A→G mutations, with 54% and 19% total mutations observed for *R,R* and *S,S*-isomers, respectively.[154] Plaque forming ability of *bis*-N^2G-BD adducted ss-M13mp7L2 vectors replicated in repair-deficient *E. coli* was decreased by 4–5 orders of magnitude as comparrd to unmodified plasmids, suggesting these adducts strongly hinder replication.[128] Mutation spectra have revealed both base substitutions and deletions, and *S,S*-*bis*-N^2G-BD was more mutagenic of the two.[128] However, it should be noted that these BD-DNA cross-links have not been detected in animals exposed to BD or BD-treated DNA. The mutagenic potency of many BD-DNA adducts formed *in vivo* remain to be established.

1.4.3 In Vitro Replication Bypass of 1,3-Butadiene-Induced DNA Adducts

Lloyd and coworkers have investigated the *in vitro* replication of several BD-DNA monoadducts and intrastrand cross-links (**Table 1-4**). Synthetic DNA strands containing site-specific *R* and *S*-N^6-HB-dA II, and *R,R* and *S,S*-N^6-THB-dA annealed to (-2), (-5) or (+5)-primers were efficiently and fully extended by *E. coli* repair polymerase, Pol I, TLS polymerase, Pol II, and the replicative polymerase, Pol III.[70] Single nucleotide insertion assays using Klenow fragment of Pol I (KF^-) have revealed that correct base (dT) was preferentially incorporated, while mispair (dA) was inserted with much less efficiency.[70] Although N^6-THB-dA was predicted to mispair with dG from molecular modeling and NMR studies,[155] no such transition mutation was observed.[70] Bypass studies conducted using *R* and *S*-N^2-HB-dG, and *R,R* and *S,S*-N^2-THB-dG have revealed that *E. coli* DNA polymerases were completely blocked by all four N^2-dG adducts under both running start and standing start conditions.[91] Further, *N3*-HB-dU adducted templates presented a complete block to KF^- and hPol ε, while calf thymus Pol δ was able to bypass the adduct, but with extremely low efficiency.[153] These results suggest the possibility for translesion synthesis of N^2-dG and *N3*-HB-dU adducts, if formed in cells. Fernandes *et al.* have tested the ability of TLS polymerases to replicate *N3*-HB-dU adducted DNA.[156] In running start experiments, the adducted template significantly blocked hPol κ, ι, and yeast Pol ζ one nucleotide prior to the lesion.[156] hPol η fully extended the adducted primer, with preferential incorporation of the correct base, dG or mismatch, dA as revealed by single nucleotide insertion assays.[156] Further, both hPol η and Pol ζ were able

to extend primers with mismatched bases opposite the adduct in post-lesion synthesis assays with the extension of dA mismatch being the most efficient, and the hPol η and Pol ζ showed a synergy in primer extension.[156] On the other hand, DNA-DNA cross-links are expected to strongly block replication. As expected, *bis*-N^2G-BD completely blocked *E. coli* polymerases *in vitro*.[128]

Tretyakova and coworkers have investigated the replication bypass of three BD-dA lesions (**Table 1-4**), *R,S*-1,N^6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2'-deoxyadenosine (*R,S*-1,N^6-HMHP-dA), *R,R*-N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine (*R,R*-N^6,N^6-DHB-dA) and N^6-(2-hydroxy-3-buten-1-yl)-2'-deoxyadenine (*S*-N^6-HB-dA). These adducts were site- and stereo-specifically incorporated into 18-mer oligodeoxynucleotides, and replication bypass by human repair polymerase β, *S. solfataricus* TLS polymerase DPO4, and human TLS polymerases η, κ and ι was studied.[112,157] (*R,S*)-1,N^6-HMHP-dA completely blocked hpol β, while hpol η and κ fully extended the adducted primer with η being the most efficient.[112] DPO4 was inefficient in primer extension, while hpol ι catalyzed nucleotide insertion opposite the adduct, but showed no further extension.[112] However, subsequent addition of η or κ showed full extension, suggesting co-operativity of TLS polymerases in replication bypass.[112] While hPol η, κ and Dpo4 preferentially incorporated the correct nucleotide, dT, opposite the adduct, steady-state kinetic studies have revealed high misinsertion frequencies (*f*) for dA and dG.[112] hPol ι exclusively inserted the correct base. Further, HPLC-MS/MS analyses of extension products of hPol η, κ and Dpo4 revealed significant amounts of dA and dG incorporation opposite the lesion site as well as base deletions.[112] Among the three

enzymes, hpol η was the least mutagenic (19% of error-prone products), whereas κ showed 82% of erroneous products, hence the most mutagenic.[112] These results suggest that replication bypass of (*R,S*)-1,N^6-HMHP-dA is highly error-prone, causing A→T and A→C mutations. In contrast, polymerase bypass of (*R,R*)-N^6,N^6-DHB-dA was extremely inefficient, and hPol η and κ incorporated all four nucleotides opposite the modified base with almost similar frequencies.[157] Hence, *(R,R)*-N^6,N^6-DHB-dA can cause A→T, A→C and A→G mutations. HPLC-MS/MS data further confirmed the highly error-prone replication past this exocyclic lesion.[157] On the other hand, (*S*)-N^6-HB-dA was readily bypassed by all DNA polymerases examined including hPol β in a highly accurate fashion.[157] Taken together, these results provide support for the mutagenic potential of BD-dA adducts, and confirm that structural features of the modified base can influence their biological effects.

Table 1-4 *In vitro* replication studies of 1,3-butadiene-induced DNA adducts.[70,91,112,128,153,156,157]

DNA Adduct	*Polymerase*	*Primer extension*		*Misinsertion*			
		Efficiency	*Fidelity*	*A*	*C*	*G*	*T*
R- and *S*-N^6-HB-dA II	Pol I, II and III	high	error-prone	✓			-
R,R- and *S,S*-N^6-THB-dA	Pol I, II and III	high	error-prone	✓			-
R- and *S*-N^2-HB-dG	Pol I, II and III	blocking	N/A		-		
R,R- and *S,S*-N^2-THB-dG	Pol I, II and III	blocking	N/A		-		
N3-HB-dU	KF^- and hPol ε	blocking	N/A			-	
	CT Pol δ**	low	-			-	
	hPol κ, ι and yeast Pol ζ	blocking	N/A			-	
	hPol η	moderate	error-prone	✓	✓	-	✓
R,R- and *S,S*-*bis*-N^2G-BD	Pol I, II and III	blocking	N/A		-		
R,S-1,N^6-HMHP-dA	hPol β	blocking	N/A				-
	hPol η and κ††	moderate	error-prone	✓		✓	-
	Dpo4††	low	error-prone	✓		✓	-
	hPol ι	insertion only	error-free				-
R,R-N^6,N^6-DHB-dA	hPol β	low	-				-
	hPol η and κ††	low	error-prone	✓	✓	✓	-
	hPol ι	insertion only	error-free				-
S-N^6-HB-dA	hPol β	moderate	error-free				-
	hPol η and κ	moderate	high	✓	✓	✓	-
	hPol ι	moderate	error-free				-

** CT – calf thymus

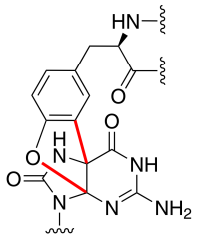
†† deletion mutations were observed.

1.5 DNA-PROTEIN CROSS-LINKING BY *BIS*-ELECTROPHILES

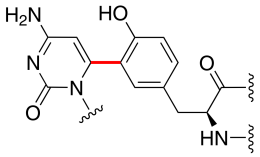
1.5.1 DNA-Protein Cross-Links: A Ubiquitous Class of DNA Adducts

DNA-protein cross-links (DPCs) are a class of structurally diverse, unusually bulky DNA lesions formed when cellular proteins become covalently trapped on DNA upon exposure to various endogenous and exogenous physicochemical agents.[4,5] Cellular DPCs are highly complex and heterogeneous due to the involvement of a wide range of proteins of varying size, physicochemical properties, cellular distribution and functions.[85,158-160] Participation of numerous sites on DNA and multiple amino acids on proteins in the cross-linking reactions further contributes to the heterogeneity of these super bulky lesions (**Chart 1-6**).[4,85,158,161-168] DPCs can compromise genetic stability and cell viability by interfering crucial cellular processes such as DNA repair, replication and transcription due to their enormous size and disruption of key DNA-protein interactions.[4,5] Despite implications of DPCs in cancer,[30-33] cardiovascular disease[34-36] and age-related neurodegeneration[35-39] and their eminent threat to cell viability, the structures and mechanisms responsible for such adverse effects are not well understood. Hence, elucidating the mechanisms of DPC formation, repair and replication is crucial to fully comprehend the biological consequences of these super bulky adducts in order to develop approaches to mitigate their lethal effects.

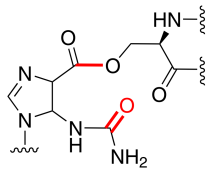
Chart 1-6 Chemical structures of DNA-protein cross-links.



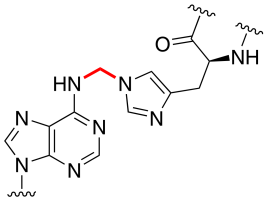
Oxidative conditions
Gua and Tyr



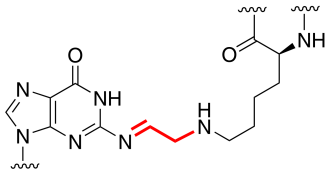
Reactive oxygen species
Cyt and Tyr



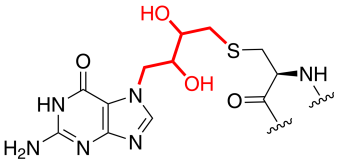
Reactive nitrogen species
Gua and Ser



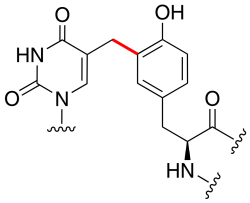
Formaldehyde
Ade and His



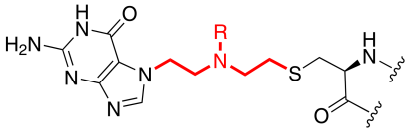
Acrolein
Gua and Lys



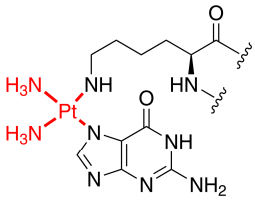
1,2,3,4-Diepoxybutane
Gua and Cys



Ionizing radiation
Thy and Tyr



Nitrogen mustards, *R = Me, Ph(CH$_2$)$_3$CO$_2$H*
Gua and Cys



Cisplatin
Gua and Lys

1.5.2 Types of DNA-Protein Cross-Links

DPCs are broadly categorized into four classes in literature.[5,169] Three classes (*A*, *B* and *C*) result from regular DNA-protein transactions, while the fourth class of DPCs (*D*), the most common and most abundant types of DPCs under physiological conditions, is generated upon exposure to various physical and chemical agents. Class *D* has received the most attention amongst researchers, since this class of DPC lesions is the culprit responsible for preventable adverse biological effects of DPCs in cells.

Class *A* DPCs occur when topoisomerase I (TOPO I) forms a covalent tyrosine phosphodiester bond at the 3' end of a DNA single strand break (SSB) during DNA unwinding,[5,169,170] and when the complex is persistent due to damaged DNA in close proximity or treatment with topoisomerase inhibitors.[170-173] These DPC lesions are repaired by tyrosyl-DNA phosphodiesterase 1 (Tdp1).[169,170]

Covalent attachment of topoisomerase II (TOPO II) and meiotic recombination protein (SPO II) to the 5' ends of double strand breaks (DSB) on DNA, possibly *via* a similar mechanism to that of class *A*, leads to the formation of class *B* DPCs.[174,175] TOPO II DPCs are repaired by tyrosyl-DNA phosphodiesterase 2 (Tdp2) or MRN-catalyzed, CtIP-dependent endonucleolytic pathway,[176] while SPOII adducts are endonucleolytically processed.[177]

An amide linkage between a protein and an oxidized form of DNA generates class *C* DPCs. For example, human DNA polymerase β (Pol β) and endonuclease III (Endo III), two enzymes involved in base excision repair (BER) can form an amide bond

with the 2'-deoxyribonolactone generated upon cleavage of the oxidized abasic site by human apurinic/apyrimidinic endonuclease 1 (APE 1).[178,179] An *in vivo* preventative mechanism for class *C* DPCs was revealed, where 2'-deoxyribonolactone is rapidly removed by long-patch BER.[180]

Class *D* DPCs are the most abundant and most interesting type of lesions, as they are formed upon exposure to a variety of physical agents and chemical species from both endogenous and exogenous sources (**Chart 1-6**). Endogenous species involved are reactive oxygen species (ROS),[166,167,181-186] reactive nitrogen species (RNS),[164,165,187,188] oxidants and free radical inducing agents,[181] and *bis*-electrophiles resulting from cellular metabolism.[189,190] Environmental pollutants and industrial chemicals,[85,86,191-210] heavy metals,[211-215] and chemotherapeutic agents[158,159,216-221] are examples of exogenous chemicals that mediate DPC formation. In addition, physical agents such as UV light and ionizing radiation can contribute to class *D* DPCs.[160,168,222,223]

1.5.3 DNA-Protein Cross-Links Induced By Physicochemical Agents

Reactive oxygen species (ROS) are mainly generated in cells as byproducts during cellular respiration in the mitochondria[224] and lipid peroxidation within cell membranes.[225] Exposure to ionizing radiation like infrared (IR) and *X*-rays can also generate ROS.[4] These species can cause one electron oxidation of DNA or proteins to initiate free radical reactions. Such electron deficient species can ultimately form DPCs by covalent attachment of proteins that interact with DNA. For example, Gajewski *et al.* reported ROS-mediated DPCs between pyrimidines on DNA and several amino acid

residues on histones both *in vivo* and *in vitro*.[182-185] In addition, theoretical studies have suggested that ROS can induce DPCs involving cytosine and tyrosine residues.[186] Further, Burrows and coworkers have reported that 8-oxoguanosine get cross-linked to lysine and tyrosine residues to mediate DPC formation (**Chart 1-6**).[166,167,181]

Some reactive nitrogen species (RNS) function as important intracellular messengers, while some can have deleterious effects on cells.[226] One of the major RNS formed *in vivo*, nitric oxide (NO) is known to induce DPCs *via* a mechanism initiated by oxidation of DNA bases or amino acid residues.[227] In addition, Oxanine, a NO-induced deaminated product of guanine forms amide linkage to Lys and Ser residues of DNA binding proteins (**Chart 1-6**).[164,165] Further, Makino and co-workers reported that cytosine diazoate, a product of the reactions of cytosine with nitrous acid (HNO_2) and nitric oxide can form lysine adducts, providing evidence for DNA-nucleoprotein cross-linking by HNO_2 and NO in cells.[187,188]

Endogenous electrophilic species can also contribute to DNA-protein cross-linking in cells. Glyoxal and methylglyoxal are endogenous aldehydes.[4,228,229] Glyoxal is an α-oxoaldehyde biosynthesized during lipid peroxidation, ascorbate autoxidation, oxidative degradation of glucose and degradation of glycated proteins.[230] Methylglyoxal, an α-ketoaldehyde, is a physiological metabolite formed during glycolytic bypass, acetone metabolism, and amino acid catabolism.[231] These highly reactive aldehydes are known to cross-link guanine in DNA and lysine and arginine residues in proteins.[189,190,232]

Murata-Kamiya *et al.* reported that methylglyoxal mediate DNA-KF^- cross-link formation.[229]

Amongst the environmental pollutants and industrial chemicals that induce DNA-protein cross-linking, formaldehyde (FA) is the most extensively studied. FA is a highly reactive colorless gas classified as a human carcinogen. It is employed in the manufacture of thousands of household, medicinal, and industrial products, leading to a high potential for human exposure,[194,233,234] and also found in cigarette smoke, automobile exhaust and photochemical smog.[235] FA is also formed endogenously through various metabolic pathways including amino acid and methanol metabolism, lipid peroxidation and P450-dependent demethylation.[236] While FA is known to induce several types of DNA adducts, the main type of lesions formed is attributed to DPCs.[191-194,234] It is well characterized that FA reacts with *N*-terminal amino group and side chains of Lys, Arg, His and Cys[237,238] forming methylols, Schiff bases, and methylene bridges.[237] It can also modify exocyclic amines and endocyclic imines of all four nucleobases in DNA (**Chart 1-6**).[162,163,238] Based on the existing evidence for DPC formation by FA,[239-243] DNA-histone cross-links are considered the main contributors to its carcinogenicity, and all five major types of histones were found to be involved in DPC formation.[244]

Acrolein, another carcinogenic chemical used in biocide and polymer industry and also present in cigarette smoke and automobile exhaust,[195,245] is known to mediate DNA-protein cross-linking in cells.[246] Acrolein is formed endogenously during lipid peroxidation.[247] This α,β-unsaturated aldehyde can mediate cross-linking *via* Michael

addition or Schiff base formation between exocyclic amines of purines in DNA and Cys, His and Lys residues of proteins (**Chart 1-6**).[195-198]

1,2,3,4-Diepoxybutane (DEB), the ultimate carcinogenic metabolite of BD (see section 1.2), is yet another *bis*-electrophile known to facilitate DPC formation.[85,86,199-201,248] Evidence for DEB mediated cross-linking of proteins to DNA *ex vivo* and *in vivo* was initially obtained using biophysical methods.[199,200] Subsequent *in vitro* studies employed recombinant proteins and mass spectrometry as a tool to show that Cys residues in proteins and *N7* of guanine in DNA are involved in DEB-mediated DPC formation (**Chart 1-6**).[201,248,249] Over three dozens of proteins with various cellular functions formed covalent cross-links to DNA in DEB-treated cells as revealed by mass spectrometry-based proteomics.[85] Other known environmental carcinogens such as *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine, sulfur dioxide (SO_2) and dihaloalkanes are also found to induce DPCs that elicit cytotoxic and mutagenic effects.[202-210,250]

Numerous metallic compounds commonly found as environmental and occupational pollutants are another main class of DPC-inducing agents.[4,169,251,252] A study conducted by Kuykendall *et al.* has shown that DPCs are a potential biomarker of exposure to transition metal cations.[253] Metal ion-mediated DPC formation can occur *via* either of two mechanisms: a metal chelating mechanism[213] or an oxidative mechanism involving free radicals.[254,255] Arsenic (As) is an environmental pollutant present in the form of arsenite or arsenate in soil, water and air.[32] *In vitro* studies using multiple human cell lines have revealed DPC formation in cells upon treatment of arsenite.[211,212] Chromium (Cr) compounds, industrial chemicals resulting from stainless steel and

pressure-treated wood manufacturing[32] and water pollutants,[253] were found to mediate DPC generation through a chelatable form of Cr.[213,256] They are thought to first chelate Cys, His and Glu residues within proteins, followed by chelation with adenine, cytosine and guanine bases within DNA, forming bulky DPC adducts.[213] In contrast, nickel (Ni), a component of alloys used to manufacture various consumer products such as coins and jewelry,[257] is found to form DPCs *via* a ROS-mediated pathway.[214,215]

UV light and ionizing radiation are two physical agents known to generate a range of DNA adducts including DPCs *in vivo*.[4,160,258-260] DPC formation by ionizing radiation has been studied using biophysical and mass spectrometry-based methods.[160,168,222,223] The amounts of DPCs generated when cells exposed to IR is 3–5-fold higher than DNA double-strand breaks and DNA-DNA cross-links.[160] Radiation-induced DPCs result from the radicals produced in cells upon irradiation, while radiation can form such radicals by direct oxidation of DNA and proteins or generating ROS.[4] Cress *et al.* found that ionizing radiation induces DPC in dose dependent manner in Chinese hamster ovary (CHO) cells.[223] Murray and coworkers has identified 29 proteins in CHO cells and human fibroblasts that become cross-linked to chromosomal DNA following exposure to ionizing radiation,[160] while Gueranger *et al.* found that UVA/6-thioguanine treated CCRF-CEM leukemia cells cross-link proteins involved in DNA repair.[259] Further, radiation-induced DPCs are known to involve tyrosine residues within proteins and thymine in DNA (**Chart 1-6**).[168]

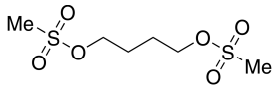
1.5.4 DNA-Protein Cross-Linking by Antitumor Drugs

Most antitumor drugs currently used in clinic elicit their therapeutic effects through their ability to modify genomic DNA, which ultimately leads to cell death.[261] The vast majority of these drugs are *bis*-alkylating agents, electrophilic species that have two reactive groups, and can form covalent bonds with two sites on biomolecules.[262,263] Antitumor drugs give rise to a variety of lesions including DNA monoadducts, single- and double-strand breaks, intra- and interstrand DNA-DNA cross-links, and DPCs.[263-266] Historically, the anti-cancer activity of these drugs has been attributed to DNA-DNA cross-linking.[267] Interstrand DNA-DNA cross-links can prevent DNA strand separation and can obstruct DNA replication inducing apoptosis.[262,267,268] However, these drugs also form unusually bulky DPC lesions.[158,159,216-221] Alkyl sulfonates, 2-chloroethyl-nitrosoureas, nitrogen mustards, and platinum-based drugs are important classes of antineoplastic agents known to induce DPCs (**Chart 1-7**).

Busulfan (1,4-butanediol dimethanesulfonate) is an antitumor alkyl sulfonate currently used in clinic (**Chart 1-7**A), and is prescribed for leukemia.[269] It is known to form a large number of DPCs (**Scheme 1-5**A),[270] although the therapeutic benefits were attributed to intrastrand DNA-DNA cross-links between *N7* of guanine and *N7* of adenine within 5'-GA-3' sequence, and between *N7* of two guanine bases within 5'-GG-3' sequence.[271]

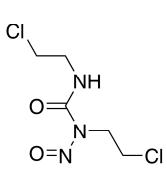
Chart 1-7 Chemical structures of *bis*-alkylating antitumor drugs known to induce DNA-protein cross-links. (A) alkyl sulfonates, (B) 2-chloroethylnitrosoureas, (C) nitrogen mustards and (D) platinum-based drugs.

A.

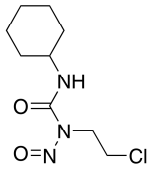


1,4-butanediol dimethanesulfonate
Busulfan

B.

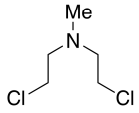


1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU)
Carmustine

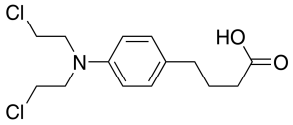


N-(2-chloroethyl)-*N'*-cyclohexyl-*N*-nitrosourea (CCNU)
Lomustine

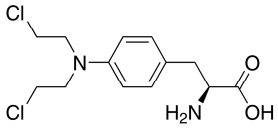
C.



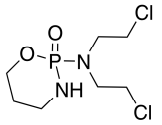
bis-(2-chloroethyl)methylamine
Mechlorethamine



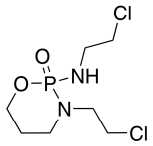
4-[bis(2-chlorethyl)amino]benzenebutanoic acid
Chlorambucil



4-[bis(chloroethyl)amino]phenylalanine
Melphalan

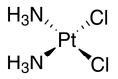


N,*N*-bis(2-chloroethyl)-1,3,2-oxaza phosphinan-2-amine 2-oxide
Cyclophosphamide

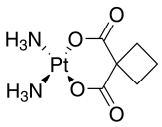


N-3-bis(2-chloroethyl)-1,3,2-oxazaphosphinan-2-amide 2-oxide
Ifosfamide

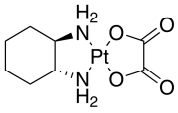
D.



cis-diamminedichloroplatinum(II)
Cisplatin

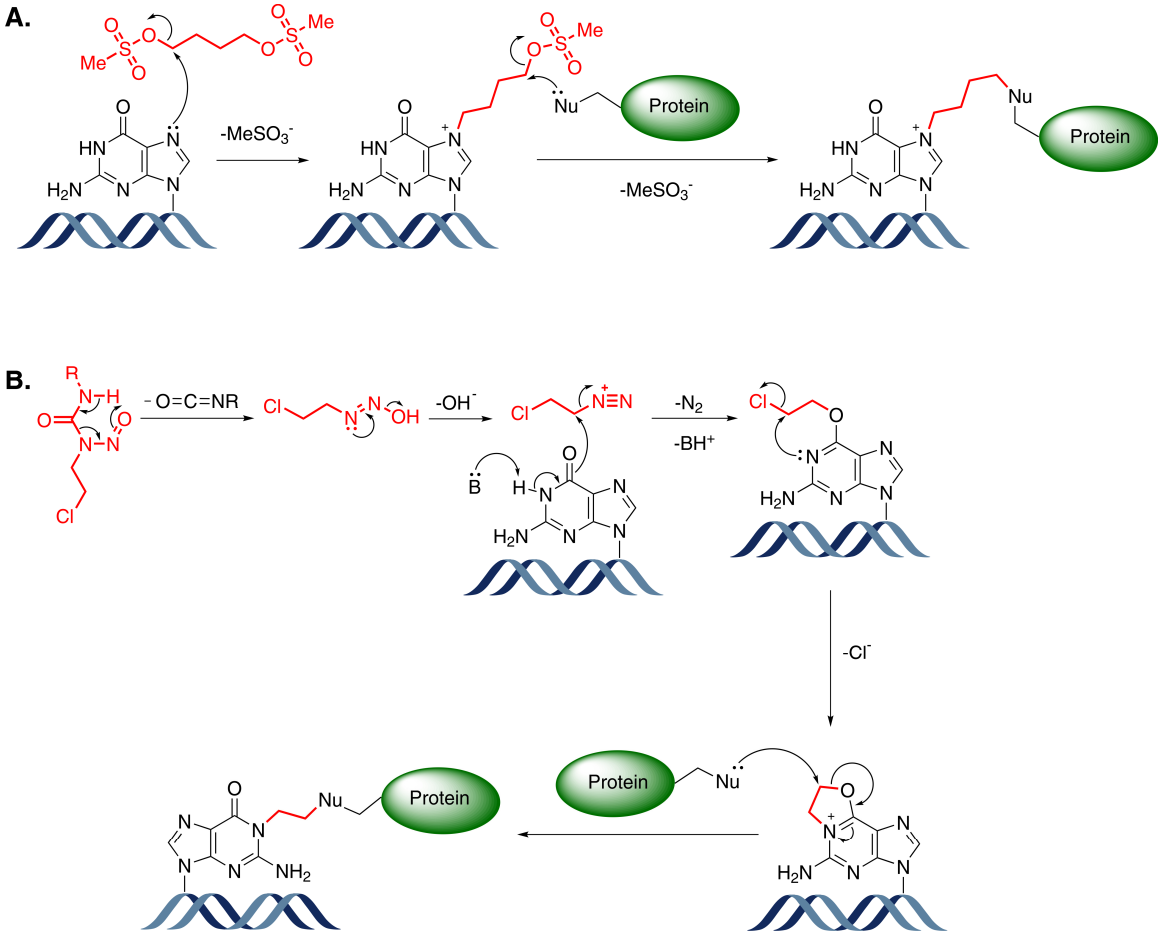


cis-diammine-[1,1-cyclobutanedicarboxylato] platinum(II)
Carboplatin



([(1*R*,2*R*)-cyclohexane-1,2-diamine] (ethanedioato-O,O')platinum(II)
Oxaliplatin

Scheme 1-5 Mechanisms of DNA-protein cross-linking by alkyl sulfonates (A) and 2-chloroethylnitrosoureas (B).[272,273]



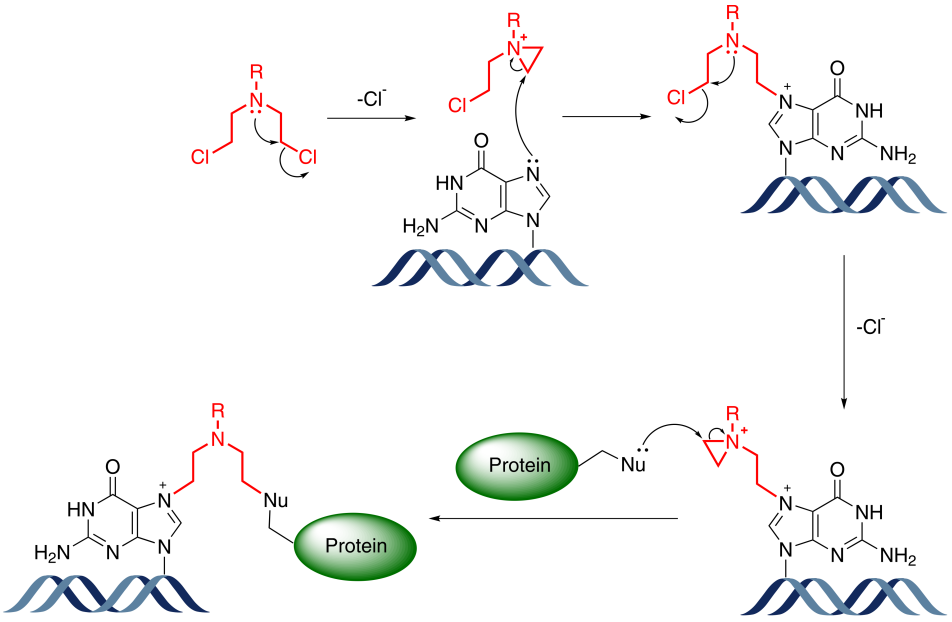
1,3-*Bis*(2-chloroethyl)-1-nitrosourea (BCNU) is one of the 2-chloroethyl-nitrosoureas used in clinic for brain tumors (**Chart 1-7**B).[274] Its therapeutic effect is due to the ability to form interstrand DNA-DNA cross-links between *N1* of guanine and *N3* of cytosine[275] that will inhibit the biosynthesis of DNA and RNA similar to other alkylating agents.[274] Seidenfeld *et al.* and Ewig *et al.* have reported that BCNU forms DPCs (**Scheme 1-5**B), in addition to DNA single strand breaks and interstrand DNA-DNA cross-links in human adenocarcinoma and mouse leukemia cell lines.[216,276]

Nitrogen mustards belong to a diverse class of chemotherapeutic agents, whose DNA-protein cross-linking ability is well documented (**Chart 1-7**C). These are *bis*-(2-chloroethyl)amines that structurally differ by the substitution at the third valence of the nitrogen atom.[262,277] Because of the high reactivity and off-target toxicity of the first nitrogen mustard drug, mechlorethamine, the *N*-methyl group was replaced with a variety of alkyl groups to reduce the reactivity and increase selectivity towards malignant cells.[262,277] This class of drugs has been used in clinic since 1940s for the treatment of a variety of cancers including leukemia, lymphoma, carcinoma and myeloma.[277,278] Their therapeutic effects are attributed to inter- and intrastrand cross-links involving primarily the *N7* of guanine, but also *N1*, *N3*, N^6 and *N7* of adenine.[279-284] Kohn *et al.* and Bonner and colleagues provided the initial evidence for DPC formation mediated by nitrogen mustards.[217,218,281,285] Tretyakova and coworkers have employed mass spectrometry-based methods to identify cellular proteins that get cross-linked to DNA in the presence of nitrogen mustards and to characterize structures of DPCs mediated by nitrogen mustards.[158,159,161] Structural elucidation was conducted using O^6-alkylguanine DNA

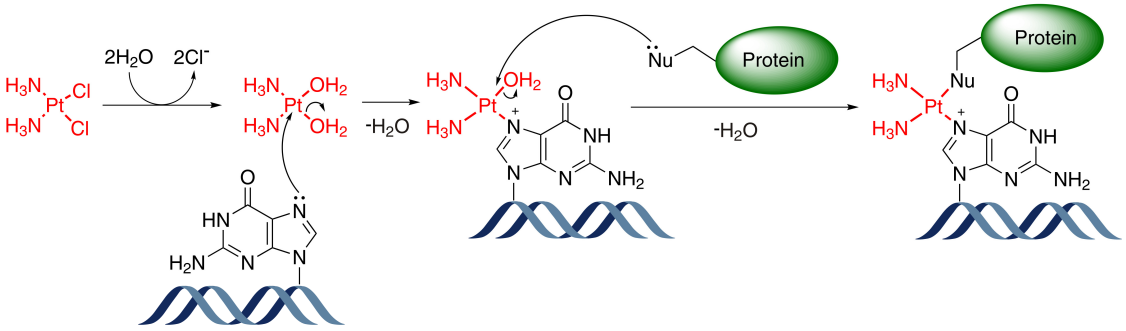
alkyltransferase (AGT) as a model protein. It was found that the Cys residues within the protein and *N7* of guanine in DNA were involved in cross-linking reactions mediated by mechlorethamine and chlorambucil (**Scheme 1-6**A).[161] Further, 15, 53 and 38 proteins, with varying cellular and biological functions including DNA repair, replication and transcription, were identified in mechlorethamine-treated Chinese hamster ovary (CHO), human cervical carcinoma (HeLa) and human fibrosarcoma (HT1080) cells, respectively.[158,159]

Scheme 1-6 Mechanisms of DNA-protein cross-linking by nitrogen mustards (A) and platinum drugs (B).[272,273]

A.



B.



Platinum-based drugs used in cancer treatment are coordination complexes of platinum (**Chart 1-7**D). Cisplatin (*cis*-diamminedichloroplatinum(II)) was the first generation antineoplastic in this class introduced to clinic in 1970s for testicular and ovarian cancers.[286] Neurotoxicity and nephrotoxicity[286,287] observed with cisplatin have led to the development of second generation platinum drugs: carboplatin (*cis*-diammine-[1,1-cyclobutanedicarboxylato]platinum(II)) and oxaliplatin ([(1*R*,2*R*)-cyclohexane-1,2-diamine](ethanedioato-*O*,*O'*)platinum(II)).[286,288] These compounds show biological activity against a wide range of cancers, especially solid tumors such as sarcoma, carcinoma, lymphoma, and brain, lung, colorectal, head and neck cancers.[286,288,289] The biological activity of platinum drugs is also attributed to their ability to form DNA monoadducts and cross-links. For example, cisplatin is known to form 1,2-intrastrand cross-links at GpG and ApG sites, 1,3-intrastrand cross-links at GpXpG sites, GG interstrand cross-links and adenine, cytosine and guanine monoadducts.[286,290] The *N3* position of cytosine, *N1* and *N7* of adenine, and *N7* of guanine were the main alkylating sites within DNA,[290] and the two amine groups remained bound to Pt.[286] Nonetheless, cisplatin has been shown to form DPCs.[221,291,292] For example, cisplatin formed DPCs with the Klenow fragment of DNA polymerase I (KF^-), histone H1, and NF-κB,[219] and DNA-protein cross-linking correlated with cell death.[293] Tandem mass spectrometry-based peptide mapping of AGT protein treated with cisplatin in the presence of a synthetic DNA has revealed that DPC occurs at *N7* of guanine within DNA, and Arg, Cys, Glu, His and Lys residues within the protein (**Scheme 1-6**B).[294] An affinity capture methodology coupled with mass spectrometry-based proteomics using HeLa nuclear

extracts has identified 131 proteins cross-linking to DNA *via* cisplatin *in vitro*.[294] Further, bottom-up mass spectrometry-based proteomics employing HT1080 cells incubated with cisplatin have discovered 256 proteins participating in cisplatin-mediated DNA-protein cross-linking.[294] The majority of proteins identified in these proteomics studies were localized in the nucleus, while their molecular and biological functions involved important cellular transactions such as DNA replication and damage response, RNA transcription and processing, and protein translation and transport.[294] Taken together, these evidence warrant the importance of investigating the biological consequences of antitumor drug-induced DNA-protein cross-links.

1.5.5 Biological Consequences of DNA-Protein Cross-Links

The unusually bulky nature of DPC lesions and their ability to block DNA protein interactions can interfere with cell viability and genetic integrity.[4,5] The high stability and long-term persistence of these lesions, except for the hydrolytically labile *N7* adducts, can further enhance these effects. For example, DPCs formed under oxidative conditions and mediated by metal ions are stable for days to weeks,[253,295,296] although some aldehyde-induced DPCs are hydrolytically unstable and last only a few hours.[199,297,298] Cytotoxic effects of many DNA damaging agents such as industrial and environmental carcinogens,[39,234,299-301] antitumor drugs,[132,216,220,302] metals,[251] and ionizing radiation[259,260] are attributed to DPCs, and some of these effects result from inhibition of DNA transactions like repair,[259] replication[132,220,234,260,300,301] and transcription,[260] ultimately leading to apoptosis.[39,301]

The genotoxic and mutagenic effects of DPC-inducing agents are also well documented.[32,95,191,299,300,303] Tretyakova *et al.* recently reported direct evidence to support DPC-induced toxicity and mutagenicity in human cells using protein monoepoxides engineered to selectively form DPCs.[304] Not only large proteins, but also smaller peptides cross-linked to DNA elicit mutagenic effects,[198,305] and the ability to cause mutations is dependent on the site of cross-linking within DNA.[198] Cytotoxicity and carcinogenicity observed in mice deficient in Fanconi Anemia DNA repair pathway is linked to DPCs formed by endogenous aldehydes.[306,307] Further, studies employing human subjects occupationally exposed to DPC-inducing agents such as transition metals have implicated DPCs in the observed genotoxicity.[308-310] For example, an increased incidence of chromosomal aberrations in lymphocytes was observed in workers occupationally exposed to Ni.[309,310]

Because of their unusual size, DPCs may affect various biological functions *via* mechanisms different to those of smaller, conventional DNA lesions. As mentioned before, they can block DNA repair, replication and transcription. DPCs have been implicated in many diseases including cancer,[30-33] cardiovascular disease[34-36] and age-related neurodegeneration.[35-39] DPCs have been observed in rat marrow cells exposed to FA, and it is hypothesized that DPC formation is the molecular mechanism responsible for FA-induced leukemia.[30] A dose-dependent increase in DPCs with exposure to DPC-inducing agents like SO_2, Cr^{VI} and Ni^{II} was observed in white blood cells (WBCs) and several organs such as lung, liver and heart of rodents.[250,251] Studies employing several strains of mice reported that the amounts of DPCs in numerous organs, including brain

and heart, increased significantly with age.[36,311] Exposure to metals such as Ni, Cr and As is implicated in various cancers including respiratory, adrenal and liver cancers in both rodents and humans.[213,312-314]

Despite their abundance in tissues, the biological effects of DPCs are poorly understood. The vast majority of DPC-inducing species can also produce other types of DNA adducts.[262]Hence, it is very difficult to isolate the deleterious effects of DPCs from those of other DNA adducts to fully comprehend biological consequences of DPCs. In this regard, novel methodologies are required to induce DPCs selectively in cells as well as synthetic approaches to generate site-specific DPCs for studies of their repair and replication.

1.6 REPAIR STUDIES OF DNA-PROTEIN CROSS-LINKS

The unusually bulky nature of DPCs presents a unique challenge to cellular repair machinery in order to restore normal DNA, because the cross-linked protein (CLP) hinders the access of repair proteins to the lesion site. Also, the stability and persistence of DPCs play a key role in the mechanisms involved in removal of these bulky adducts in cells. The hydrolytically labile DPCs, such as those generated by aldehydes, can be rapidly eliminated from cells within a matter of hours.[199,297,298] However, certain DPCs are hydrolytically stable and long-lived.[253,295,296] There is evidence that DPCs induced by transition metals and oxidative stress can persist through several cycles of DNA replication,[260,315,316] and are only partially repaired,[317] thus resulting in downstream effects *in vivo* due to permanent damage to genomic DNA.[4] Hence, such stable DPCs require the participation of active cellular repair. Several factors including the chemical nature of the DPC lesions and specific cellular system can dictate which damage tolerance mechanisms will be involved in recognizing a particular type of DPC lesions.[4,5] Direct reversal, nucleotide excision repair (NER) and homologous recombination (HR) are potential mechanisms by which these bulky adducts can be removed *in vivo* (**Schemes 1-2** and **1-3**).

Direct reversal by chelation is possible, if DNA and protein is bound through complexation with a metal like Cr.[4,318] FA-induced DPCs were found to be removed by hydrolysis.[4,297] DPCs-induced by platinum compounds are known to release the CLP by "platination migration".[319-321] This phenomenon was observed *in vitro* with DNA-AGT cross-links induced by cisplatin.[294] The AGT protein was released by the migration of the

Pt-S bond of a Cys residue within CLP to *N7*-G on DNA forming a DNA-DNA cross-link.[294] This can be a possible mechanism by which Pt-induced DPCs are removed from cells.

NER is known to have a broad substrate specificity and to remove bulky, helix distorting DNA adducts,[115,322,323] hence can be potentially involved in DPC repair. Lloyd and coworkers reported that bacterial *UvrABC* nuclease incised site-specific DPCs between T4 pyrimidine dimer glycosylase/apurinic/apyrimidinic site lyase (T4-pdg) and an apurinic/apyrimidinic (AP) site within duplex DNA *in vitro*.[324,325] Nakano *et al.* showed that *in vitro* excision of proteins cross-linked to DNA *via* an oxanine moiety by *UvrABC* is size-dependent, and NER only repairs DPCs smaller than 11–14 kDa.[326] In contrast, upper size limit of DPCs for mammalian NER was found to be 8–10 kDa.[327] However, FA-induced chromosomal DPCs were not repaired by mammalian NER proteins,[327] which contrasts a previous report that similar DPCs were removed by bacterial NER system.[326] In addition, *XPF*-deficient CHO cells showed a marked decrease in IR-induced DPC repair under hypoxic conditions as compared to wild-type cells, suggesting the involvement of NER.[328] In contrast, *XPD*- and *XPB*-deficient cells were not IR-sensitive under hypoxia, suggesting that NER is not involved in IR-induced DPC lesions, rather recombination repair is.[329]

Several groups have proposed that DPC repair involves partial proteasomal degradation of the DPCs followed by NER. Lloyd and colleagues reported that tetra- and dodecylpeptides cross-linked to either an AP site or N^2-G of DNA *via* a trimethylene linker were excised more efficiently by *UvrABC* complex as compared to T4-pdg

cross-linked to an AP site in DNA.[325] Reardon *et al.* confirmed these observations using the same substrates in both mammalian and bacterial systems.[330] Quievryn and Zhitkovich found that proteasome inhibitors hinder the repair of FA-induced DPCs in human HF/SV fibroblasts and NER-deficient *XPA* cells, while NER-deficient *XPF* cells showed higher sensitivity to FA, suggesting the involvement of XPF protein in the removal of DPCs.[297] Baker *et al.* have investigated the repair of site-specifically cross-linked HhaI DNA methyltransferase (HDnmt) to *C6* of 5-fluorocytosine containing oligodeoxynucleotides or plasmids.[331] *In vitro* assays using mammalian cell free extracts have shown that protease-digested, but not the full length HDnmt-DPC was excised.[331] Further, their *in vivo* assays demonstrated that wild-type cells repaired the site-specific DPCs, while both *Xpg*-deficient cells and 26S proteasome inhibitor-treated cells were ineffective in repair.[331] Based on these observations, Reardon *et al.* have proposed a replication coupled NER model for DPC repair:[169] when the replication fork encounters DPCs, fork regression triggers the recruitment of proteolytic enzymes to degrade the CLP to a smaller peptide.[169] Simultaneously, repair factors necessary to excise the resulting peptide cross-links are recruited at the lesion site to initiate NER.[169]

Independent reports on hypersensitivity of bacterial *uvrA* and *recA* mutants to FA treatment, and hypersensitivity of *recA*, but not the *uvrA* mutant, to 5-azacytidine (azaC) suggest that multiple repair pathways are involved in DPC repair.[326,332,333] These observations suggest that both FA and azaC-induced DPCs are repaired by HR, but only FA-induced DPCs are recognized by NER. Confirming this hypothesis, Nakano *et al.* observed that HR-deficient, but not HR-proficient bacterial and mammalian cells, were

hypersensitive to DPC-inducing agents.[326,327] Further, bacterial cells repaired double strand breaks resulting from replication folk stalling upon encounter of DPCs *via* *RecBDC*-dependent HR.[326] Accumulation of *RAD51*, a pivotal protein in HR, in mammalian cells treated with DPC-inducing agents further suggests that HR is responsible for the removal of the majority of DPCs *in vivo*.[327]

1.7 REPLICATION BYPASS OF DNA-PROTEIN CROSS-LINKS

Hydrolytically stable DPCs,[253,295,296] if not repaired, can interfere with DNA replication.[4,5] These super bulky lesions are likely to block high fidelity replicative polymerases.[132] Cellular DNA damage response mechanisms will then initiate translesion synthesis (TLS)[130,133,134] or HR.[5] During TLS, low fidelity bypass polymerases can potentially accommodate these bulky DPC lesions in their large and flexible active sites to replicate past the adducted DNA (**Scheme 1-4**). As mentioned in section 1.4.1, the replication bypass by TLS polymerases can be inefficient and error-prone.[149,150] Structural features such as site of cross-linking on DNA and protein, size of CLP, and identity of CLP as well as the polymerase can influence the ability of DPCs to be bypassed, and also the fidelity of replication.[132,164,219,326,327,334-338] However, only a handful of studies have directly examined the effects of DPCs on DNA replication.

In vitro replication studies conducted using histone H1 cross-linked to DNA *via* cisplatin or transplatin have revealed that histone DPCs completely blocked the Klenow fragment of *E. coli* DNA polymerase I (KF^-) and HIV-1 reverse transcriptase (RT HIV-1).[219,334] The main termination site was one nucleotide prior to the DPCs.[219,334] Ide and coworkers transformed wild-type and *uvrA E. coli* strains with pGL3-CMV plasmids containing FA-induced histone H1 DPCs.[326] With both wild-type and *uvrA* cells, the transformation efficiency of plasmids containing histone H1 DPCs was <10% compared to that of undamaged plasmids.[326] However, the transformation efficiency of plasmids containing partially digested histone H1 increased up to 58% with wild-type cells.[326] These results imply that histone DPCs hinder *in vivo* replication. Kuo *et al.*

investigated the *in vivo* replication of pBR322 plasmids in *E. coli* in the presence of azaC.[132] AzaC induces DPCs by cross-linking DNA cytosine-*C5* methyltransferase (MTase).[132] Two-dimensional agarose gel electrophoresis has revealed the accumulation of bubble and Y molecules in wild-type plasmids, but not in *Eco*RII methylation site mutant.[132] These results suggest that azaC-induced DPCs block DNA replication. Another study reported that UV-induced DPCs containing T4-pdg blocks plasmid replication in cells.[339] T4-pdg was trapped at UV-induced cyclobutane pyrimidine dimers on double-stranded pMS2 plasmids, and wild-type and *ΔuvrD* cells were transformed.[339] The relative transformation efficiency of DPC containing plasmids was ~5%, while that of UV-irradiated plasmids was ~50% in *ΔuvrD* strain suggesting replication inhibition by UV-induced DPCs.[339]

A few studies have also been reported in literature on *in vitro* replication of DNA-peptide cross-links. It is known that DNA-peptide cross-links are formed *in vivo*,[305,340] while natural and synthetic peptides have been cross-linked to DNA *in vitro*.[197,341-343] Further, the hypothesis that DPCs get proteolytically degraded to a peptide in cells prior to their repair,[297,331] hence the possibility of these lesion to encounter cellular replication machinery also warrants the efforts to investigate the replication of DNA-peptide cross-links. Lloyd and coworkers conducted replication bypass studies with DNA templates containing tetra- and dodecylpeptides cross-linked to N^2 position of guanine *via* a trimethylene linker, *i.e.*, acrolein-induced DNA-peptide cross-links.[335] DinB family polymerases, hPol κ and its *E. coli* orthologue, Pol IV were able to catalyze replication past these DNA-peptide adducts with varying efficiencies.[335] Both enzymes exclusively

incorporated the correct nucleotide, dC opposite the peptide cross-links, suggesting error-free replication.[335] In contrast, *E. coli* replicative polymerase, Pol III or damage inducible polymerases, Pol II and Pol V could not bypass the peptide adducts.[335] Interestingly, A family polymerase ν was capable of bypassing both 4-mer and 12-mer peptides cross-linked to N^6 of adenine *via* a trimethylene linkage, while similar adducts at the N^2 of guanine completely blocked replication in both standing and running start experiments.[336] Replication bypass of N^6-dA-peptide cross-links by Pol ν was also highly accurate, and only the correct base, dT was inserted opposite the adduct.[336] These observations suggest that major groove DPC adducts, but not the corresponding minor groove adducts, can be bypassed by Pol ν. The authors proposed that small major groove DPC adducts have sufficient conformational flexibility to be accommodated within the active site of the polymerase without disturbing primer-template-enzyme interactions.[336] Further, it is reported that *E. coli* high-fidelity DNA repair polymerase, Pol I can also efficiently and accurately bypass a dodecylpeptide cross-linked to N^6 of adenine in DNA.[337] The extent of primer extension was comparable to that of the undamaged template.[337] In contrast, *E. coli* DNA damage-inducible polymerases, Pol II, Pol IV and Pol V replicated past the peptide adducts with very low efficiency, although the process was still highly accurate. *E. coli* Pol III and hPol κ were completely blocked by this adduct.[337] Guengerich and coworkers employed a glutathione-DNA adduct of DEB, *S*-[4-(N^6-deoxyadenosinyl)-2,3-dihydroxybutylglutathione (N^6-dA-$(OH)_2$butyl-GSH) to investigate the efficiency and fidelity of replication bypass.[338] The highest primer extension efficiency was observed with Pol T7 and hPol η, with full extension under their

experimental conditions.[338] Human TLS polymerases κ and ι, as well as the bacterial polymerases Pol T7 and DPO4, preferentially incorporated the correct base, dT opposite the adducted site, with misincorporation of dC at varying efficiencies.[338] HPLC-MS/MS analysis of replication products revealed that 40–55% of replication products of hPol κ and Pol T7 were substitution products.[338] Further, replication by hPol η was highly erroneous, and all four dNTPs were incorporated opposite the adducted DNA, suggesting the possibility for both transition and transversion mutations during bypass.[338]

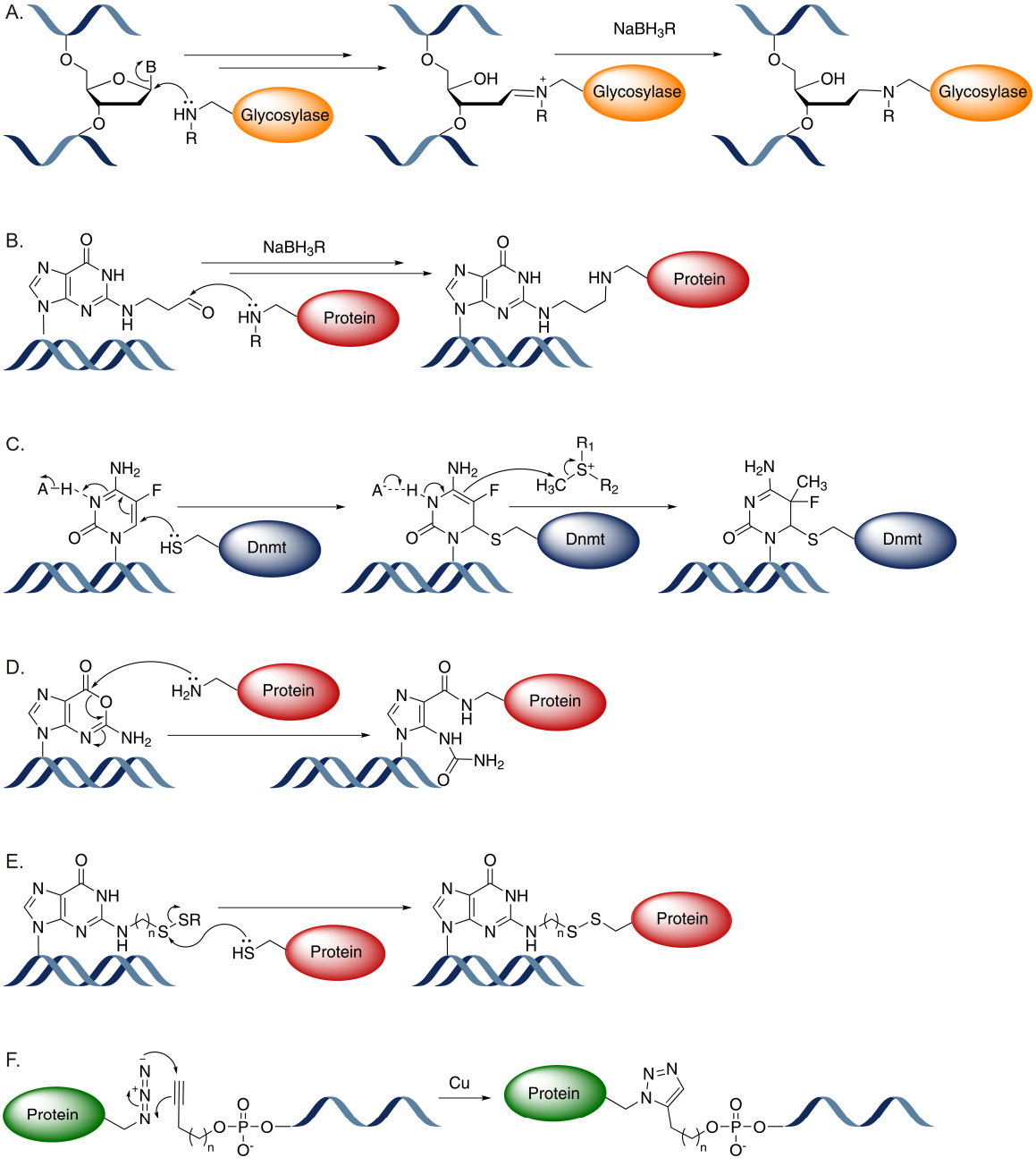
Amidst the paucity of systematic studies to evaluate the effects of DPCs on DNA replication, structures of DPCs employed in previous studies do not necessarily represent the major DPCs formed *in vivo*. For example, *N7* of guanine is the most common site of DNA that participates in DPC formation mediated by environmental carcinogens and antitumor drugs.[85,158,159,294] Yet, replication of *N7*-G cross-linked DPCs has not been previously investigated. Further, hardly any studies have been reported on *in vivo* replication of DPCs, and none have been conducted using eukaryotic systems.

1.8 SYNTHETIC METHODLOGIES TO PREPARE SITE-SPECIFIC DNA-PROTEIN CROSS-LINKS

The paucity of site-specific DNA substrates containing biologically relevant, hydrolytically stable DNA-protein conjugates is a major obstacle to fully comprehend the biological consequences of DPCs. Schiff base formation between an aldehyde-functionalized DNA and amine groups within proteins/peptides followed by reduction of the imine linkage to a stable amine has been the most widely used approach to generate DPCs. A semi-synthetic reductive amination methodology has been used to cross-link bifunctional glycosylases to DNA (**Scheme 1-7**A). A nitrogen nucleophile on the enzyme displaces the modified DNA base by forming a covalent linkage to C1'.[344] This intermediate rearranges to a Schiff base, prior to strand cleavage that can be irreversibly trapped using a reducing agent.[344] T4 pyrimidine dimer glycosylase/AP lyase (T4-pdg),[169,325,330] Endonuclease VIII (Nei),[345] formamidopyrimidine DNA glycosylase (Fpg),[346] 8-oxoguanine DNA glycosylase (Ogg and its bacterial ortholog, MutM)[347,348] have been trapped on abasic sites within DNA to generate site-specific DPCs. Lloyd, and Marnett and coworkers have used the reductive amination approach to prepare acrolein-induced DNA-polypeptide cross-links (**Scheme 1-7**B). The protein (T4-pdg, histones or EcoRI)[246,349] or Lys-rich tetra- and dodecyl-peptides[197,325,330,335-337] have been cross-linked to the ring open form of the γ–hydroxy-1,N^2-propano-dG or γ–hydroxy-1,N^6-propano-dA to form a Schiff base and subsequent reduction to yield DPCs conjugated to N^2 of guanine and N^6 of adenine *via* a trimethylene linkage. Another approach involved the trapping of DNA methyltransferase (Dnmt) on

oligodeoxynucleotide containing 5-fluoro-2'-deoxycytosine (5F-dC) (**Scheme 1-7**C).[331] A cysteine residue of the enzyme attacks the *C6* of 5F-dC forming a covalent intermediate, and the resulting enamine undergoes methylation at *C5*.[344] Fluoro substitution at *C5* prevents enzyme elimination resulting in a DPC.[344] Ide and coworkers have employed the spontaneous reaction of oxanine (Oxa), a NO-induced oxidative lesion of Gua, with amino groups of proteins to prepare site-specific DPCs containing a pyrimidine ring-open structure (**Scheme 1-7**D).[164,326,327] Disulfide cross-linking is another strategy used to generate site-specific DPCs (**Scheme 1-7**E). A disulfide tether-appended nucleobase has been incorporated site-specifically to DNA allowing the Cys residues within proteins to participate in cross-linking.[344] HIV reverse transcriptase (HIV-RT), a peptide derived from yeast transcription activator, GCN4, and Ogg have been cross-linked *via* alkyl linkers to N^2 of Gua,[350] N^6 of Ade[351] and *C4* of Cyt,[347] respectively. Distefano and coworkers have developed an alkyne-azide cycloaddition reaction (click reaction)-based methodology to prepare DPCs using alkyne containing DNA and azide-functionalized proteins (**Scheme 1-7**F). Green fluorescent protein (6×His-eGFP) and mCherry containing an azide were cross-linked to the 5'-ends of alkyne-functionalized oligodeoxynucleotides *via* the Cu-catalyzed click reaction[352] or the Cu-free variation.[353] However, these previous methodologies have several limitations such as poor reaction efficiency and low yields,[164] limited choices of protein reagents (e.g. specific DNA modifying proteins),[331,344] and insufficient site specificity with respect to the cross-linking site within the protein,[164,246,354] requiring novel synthetic methodologies to prepare site-specific, structurally defined DPCs.

Scheme 1-7 Synthetic methodologies available in the literature to prepare site-specific DNA-protein cross-links. (A) glycosylase trapping, (B) acrolein-mediated cross-linking, (C) DNA methyltransferase trapping, (D) oxanine-mediated cross-linking, (E) disulfide cross-linking, and (F) azide-alkyne cycloaddition.



1.9 THESIS GOALS

As described above (section 1.2), BD is a known human carcinogen and its epoxide metabolites can induce DNA damage in cells. However, the repair mechanisms responsible for the removal of BD-induced DNA lesions are not well understood. One of the goals of this thesis was to investigate the repair mechanisms responsible for the removal of three 2'-deoxyadenosine (dA) adducts of BD: a monoadduct formed by 3,4-epoxy-1-butene (EB), N^6-(2-hydroxy-3-buten-1-yl)-2'-deoxyadenosine (N^6-HB-dA) and two exocyclic lesions formed by 1,2,3,4-diepoxybutane (DEB), 1,N^6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2'-deoxyadenosine (1,N^6-HMHP-dA) and N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine (N^6,N^6-DHB-dA). We prepared site- and stereospecific BD-dA adducted synthetic oligodeoxynucleotides by a post-oligomerization methodology, and *in vitro* repair by human nuclear extracts and recombinant proteins was examined using gel electrophoresis and liquid chromatography-tandem mass spectrometry-based methodologies (Chapter 2).

DNA-protein cross-links (DPCs) are ubiquitous, structurally diverse DNA adducts formed in cells upon exposure to a variety of physical and chemical DNA damaging agents.[4,5] These super bulky, helix-distorting DNA adducts interfere with DNA metabolism, and elicit adverse biological effects *in vivo*.[4,5] The majority of *bis*-electrophile-induced DPCs formed *in vivo* involves the *N7* of guanine within DNA.[85,158,159,294] Moreover, it is hypothesized that cross-linked protein of DPCs is proteolytically processed to generate DNA-peptide conjugates prior to repair or replication.[297,330,331] Nonetheless, the biological consequences of DPCs are not fully

understood, partly due to the synthetic challenge in generating site-specific DPC substrates that are biologically relevant. Therefore, the second goal of this thesis was to develop synthetic strategies to generate hydrolytically stable, site-specific DPCs. In this regard, we used a post-synthetic reductive amination strategy to prepare structural mimics of nitrogen mustard-mediated DPCs (Chapter 3). Further, we employed copper-catalyzed [3+2] Huisgen cycloaddition reaction to generate site-specific DPCs to the *C5* of thymidine in DNA (Chapter 5).

The final goal of this thesis was to investigate the replication bypass of DPCs. We employed gel electrophoresis and HPLC-ESI$^-$-HRMS and MS/MS methods to evaluate the efficiency and fidelity of polymerase bypass of model DPC substrates by human translesion synthesis (TLS) polymerases. Primer extension assays were conducted using the model DPCs prepared in Chapter 3 and 5 to investigate the effects of DPCs of varying size on DNA replication, while steady state kinetic studies were performed to assess the fidelity and efficiency of replication bypass (Chapters 4, 5 and 6). Further, we employed an affinity capture-HPLC-ESI$^-$-HRMS and MS/MS-based methodology to quantify and sequence the replication products of DPCs (Chapter 6) to examine the fidelity of post-lesion synthesis and to detect any insertion/deletion mutations caused by human TLS polymerases upon replication of DPCs.

2 BASE EXCISION REPAIR OF EXOCYCLIC 2'-DEOXYADENOSINE ADDUCTS OF 1,3-BUTADIENE

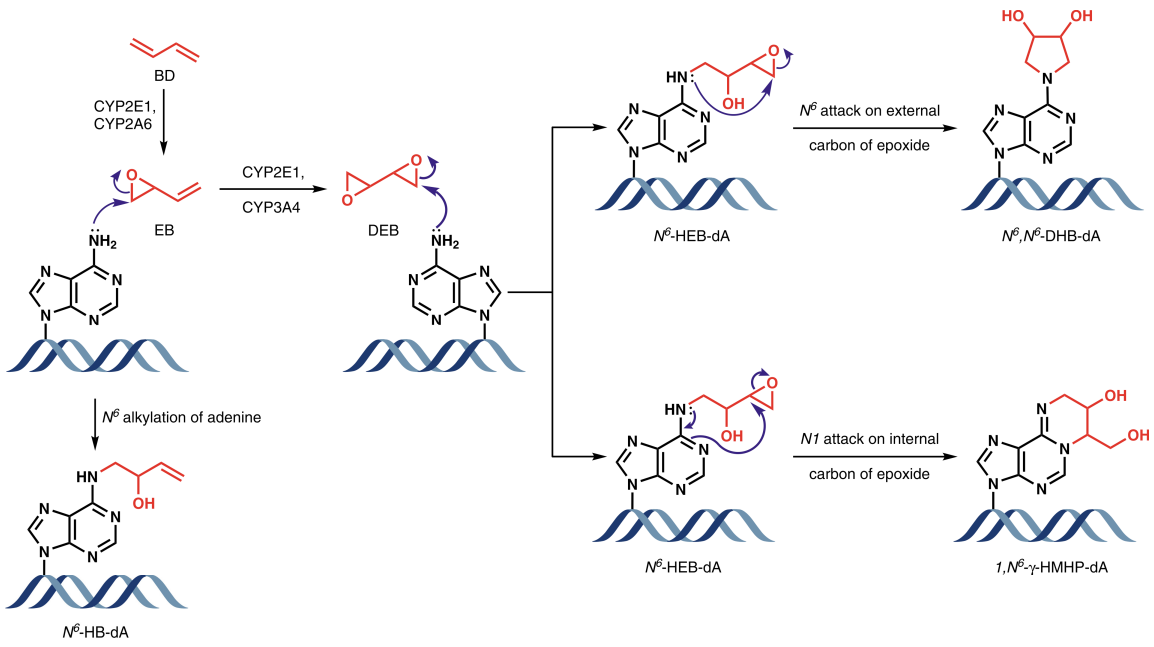
2.1 INTRODUCTION

1,3-butadiene (BD) is an important industrial and environmental chemical widely used in synthetic rubber and plastic industry and present in automobile exhaust, urban air, and cigarette smoke.[25,45,49] Based on the results of epidemiological and toxicological studies, BD is classified as a human carcinogen.[62,88,355,356] BD is metabolically activated to DNA-reactive epoxides, 3,4-epoxy-1-butene (EB), 3,4-epoxybutan-1,2-diol (EBD) and 1,2,3,4-diepoxybutane (DEB),[70,81,95,97] which form an array of DNA lesions including DNA monoadducts,[73-75] exocyclic lesions,[78,79] DNA-DNA cross-links,[80-82,84] and DNA-protein cross-links.[85,86]

Although the bulk of BD-DNA adducts are formed at the *N7* position of guanine,[357] BD lesions formed at adenine bases are of special interest because of the tendency of BD and its metabolites to cause a large number of A→T and A→G mutations.[70,90,358-360]. Multiple BD-adenine adducts have been identified.[73,75,82,84] Among these are N^6-(2-hydroxy-3-buten-1-yl)-2'-deoxyadenosine (N^6-HB-dA), 1,N^6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2'-deoxyadenosine (1,N^6-HMHP-dA) and N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine (N^6,N^6-DHB-dA) (**Scheme 2-1**).[75,78] N^6-HB-dA is formed upon alkylation of the N^6 position of adenine in DNA by 3,4-epoxy-1-

butene,[75] while 1,N^6-HMHP-dA and N^6,N^6-DHB-dA are induced by double alkylation of N^6-adenine by 1,2,3,4-diepoxybutane.[78]

Scheme 2-1 Proposed mechanism for the formation of 1,3-butadiene-induced 2'-deoxyadenosine adducts. 1,3-Butadiene (BD) is metabolized by cytochrome P450 enzymes to 3,4-epoxy-3-butene (EB) and 1,2,3,4-diepoxybutane (DEB). N^6 alkylation of 2'-deoxyadenosine (dA) in DNA by EB generates N^6-(2-hydroxy-3-buten-1-yl)-2'-deoxyadenosine (N^6-HB-dA), while N^6 alkylation of dA by DEB produces N^6-(2-hydroxy-3,4-epoxybutan-1-yl)-2'-deoxyadenosine (N^6-HEB-dA). N^6 attack on the external carbon of the oxirane ring on N^6-HEB-dA generates N^6,N^6-(2,3-dihydroxy-butan-1,4-diyl)-2'-deoxyadenosine (N^6,N^6-DHB-dA). Alternatively, *N1* attack on the internal carbon of the oxirane generates 1,N^6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2'-deoxy-adenosine (1,N^6-HMHP-dA).



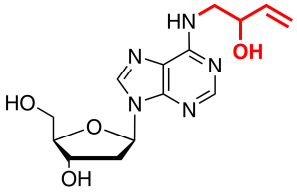
Primer extension studies with site- and stereospecific adducts have revealed that (*R,S*)-1,N^6-HMHP-dA completely blocked DNA replication by human polymerase β (hPol β), while human translesion synthesis (TLS) polymerases η and κ were able to bypass the adduct, introducing T, A, or G opposite the lesion and inducing deletion mutations.[112] Polymerase bypass of (*R,R*)-N^6,N^6-DHB-dA by both hPol β and TLS polymerases was extremely inefficient, and all four nucleotides were inserted with similar frequencies opposite the modified base.[157] In contrast, (*S*)-N^6-HB-dA was readily bypassed by all DNA polymerases examined (including hPol β) in an error-free fashion.[157] Taken together, these results provide evidence for the cytotoxic and mutagenic potential of 1,N^6-HMHP-dA and N^6,N^6-DHB-dA adducts. However, in order for BD-dA lesions to cause mutations, they have to persist long enough in cells to be replicated in an error-prone manner. Therefore, cellular repair of 1,N^6-HMHP-dA and N^6,N^6-DHB-dA adducts should be investigated as it can define the biological fate of these adducts *in vivo*.

In the present study, we investigated the ability of nuclear protein extracts from human cells and recombinant base excision repair (BER) enzymes to recognize and cleave site- and stereospecific BD-dA adducts: *S*-N^6-(2-hydroxy-3-buten-1-yl)-2'-deoxyadenosine (*S*-N^6-HB-dA), *R*,*S*-1,N^6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2'-deoxyadenosine (*R*,*S*-1,N^6-HMHP-dA) and *R*,*R*-N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine *(R*,*R*-N^6,N^6-DHB-dA) (**Scheme 2-2**A). We found that all three adducts were recognized and excised with efficiency comparable to other known BER substrates (1,N^6-ethenoadenine, guanidinohydantoin and 5-fluorouracil, **Scheme 2-2**B), although

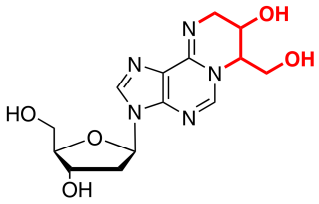
the specific BER enzyme(s) involved in recognition of BD-induced adducts remains to be identified.

Scheme 2-2 Structures and sequences of DNA lesions used in repair studies. (A) structures of novel BD-dA lesions studied in this work, (B) structures of known BER substrates used as positive controls in BER assays, (C) the sequences of 18-mer and 11-mer DNA duplexes employed in gel electrophoresis-based assays, and (D) the sequences of adducted 18-mer and 11-mer DNA duplexes and corresponding repair products employed in mass spectrometry-based assays.

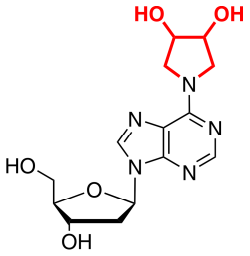
A



X = N^6-HB-dA

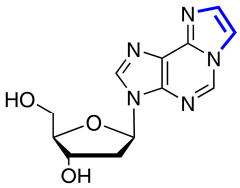


X = 1,N^6-HMHP-dA

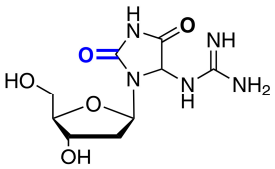


X = N^6,N^6-DHB-dA

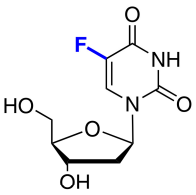
B



X = 1,N^6-ethano-dA



Y = Guanidinohydantoin



Z = 5-fluoro-dU

C

^{32}P-5'-TCA T***X***G AAT CCT TCC CCC-3'
3'-AGT A***T***C TTA GGA AGG GGG-5'

^{32}P-5'-TCA T***Y***G AAT CCT TCC CCC-3'
3'-AGT A***C***C TTA GGA AGG GGG-5'

^{32}P-5'-TCA T***Z***G AAT CCT TCC CCC-3'
3'-AGT A***A***C TTA GGA AGG GGG-5'

^{32}P-5'-CGG AC***X*** AGA AG-3'
3'-GCC TG***T*** TCT TC-5'

D

5'-TCA T***X***G AAT CCT TCC CCC-3'
3'-AGT A***T***C TTA GGA AGG GGG-5'

5'-CGG AC***X*** AGA AG-3'
3'-GCC TG***T*** TCT TC-5'

5'-TCA T-3' and 5'-CGG AC-3'

2.2 MATERIALS AND METHODS

2.2.1 Materials

Protected 2'-deoxyribonucleoside-3'-phosphoramidites (PAC-dA-CE, Ac-dC-CE, *p-i*Pr-PAC-dG-CE, dT-CE), 8-oxo-dG-CE, 1,N^6-etheno-dA-CE and 5-fluoro-dC-CE phosphoramidites, Ac-dC-CPG ABI and *p-i*Pr-PAC-dG-CPG ABI columns, and all other reagents necessary for automated DNA synthesis were purchased from Glen Research (Sterling, VA). 5′-*O*-(4,4′-dimethoxytrityl)-3′-*O*-(2-cyanoethyl)-*N*,*N*-diisopropyl-phosphoramidite of 6-chloropurine-2′-deoxyriboside was purchased from ChemGenes Corp. (Wilmington, MA). Synthetic DNA oligodeoxynucleotides were synthesized by solid phase synthesis using an ABI 394 DNA synthesizer (Applied Biosystems, CA). *E. coli* mismatch uracil DNA-glycosylase (Mug) and human AP endonuclease were purchased from Trevigen (Gaithersburg, MD). T4 polynucleotide kinase (T4-PNK) was obtained from New England Biolabs (Beverly, MA), while γ-^{32}P ATP was purchased from Perkin-Elmer Life Sciences (Boston, MA). 40% 19:1 acrylamide/bis solution and micro bio-spin 6 columns were purchased from Bio-Rad (Hercules, CA). Illustra NAP-5 desalting columns and Sep-Pak C18 SPE cartridges were obtained from GE Healthcare (Pittsburg, PA) and Waters (Milford, MA), respectively. All other chemicals and solvents were purchased from Sigma-Aldrich (Milwaukee, WI) and used without further purification.

2.2.2 Preparation of Human Fibrosarcoma Nuclear Extracts

Nuclear extracts were prepared as previously described.[361] In brief, human fibrosarcoma cells (HT1080) were grown in Dulbecco's modified Eagle's media supplemented with 9% fetal bovine serum (Life Technologies, Grand Island, NY) in 150 mm tissue culture dishes. Cells were cultured in a humidified atmosphere of 5% carbon dioxide and 95% air at 37 °C. HT1080 cells (30–50 million cells/dish) were collected from 15 confluent 150 mm tissue culture dishes, washed thrice with ice-cold phosphate buffered saline, and resuspended in 2 mL of buffer A (10 mM Tris [pH 7.4] containing 10 mM KCl, 10 mM $MgCl_2$, and 10 mM DTT). Following 15 min incubation on ice, phenylmethylsulfonyl fluoride (PMSF) was added to a final concentration of 1 mM, and the cells were mechanically disrupted by 20 strokes in a Dounce homogenizer (tight pestle). The released nuclei were sedimented and resuspended in 2 mL buffer B (comprised of buffer A supplemented with 350 mM NaCl, 1mM PMSF, 0.5 μg/ml leupeptin, 1.0 μg/ul aprotinin and 0.7 μg/ml pepstatin) and incubated for 60 min on ice. This material was centrifuged at 70,000 rpm in a Beckman TL-100.3 rotor at 4 °C for 30 min. Following the addition of glycerol and β-mercaptoethanol to the final concentrations of 10% and 10mM, respectively, the extract was stored at -80 °C. The total protein concentration was measured using the Bradford assay (2–2.9 mg/mL).

2.2.3 Synthesis of Site-Specifically Modified DNA Substrates

Synthetic oligodeoxynucleotides containing site- and stereospecific (*S*)-N^6-HB-dA, (*R,S*)-1,N^6-HMHP-dA and (*R,R*)-N^6,N^6-DHB-dA lesions at position ***X*** (5'-TCA T***X***G AAT CCT TCC CCC-3' and 5'-CGG AC***X*** AGA AG-3') were synthesized by the post-oligomerization methodology developed by Tretyakova and coworkers.[362] Briefly, (*R,S*)-1,N^6-HMHP-dA adducted DNA strands were prepared by coupling (*R,R*)-*N*-Fmoc-1-amino-2-hydroxy-3,4-epoxybutane with the oligomers containing site-specific 6-chloropurine at position ***X***. The resulting (*R,R*)-N^6-(2-hydroxy-3,4-epoxybutan-1-yl)-adenine containing oligodeoxynucleotides were isolated by HPLC and subjected to cyclization in water to afford the corresponding (*R,S*)-1,N^6-HMHP-dA strands. The corresponding oligomers containing (*S*)-N^6-HB-dA and *(R,R)*-N^6,N^6-DHB-dA adducts were prepared by carrying out nucleophilic aromatic substitution of 6-chloropurine containing DNA (on solid support) with (*S*)-*N*-Fmoc-1-aminobut-3-en-2-ol and (*R,R*)-pyrrolidine-3,4-diol, respectively. The modified oligodeoxynucleotides were cleaved off solid support using 0.1 M NaOH for 3 days at room temperature. DNA 18-mer strands containing 5-flouro-dU, 8-oxo-dG and 1,N^6-etheno-dA modifications at position ***X*** (positive controls for BER experiments) were prepared by solid phase synthesis using commercially available phosphoramidites (Glen Research, Sterling, VA). The modified base was added using offline manual coupling protocol. The 8-oxo-dG-containing oligomer was subsequently oxidized with Na_2IrCl_6 to generate guanidinohydantoin.[363] All DNA strands were purified by reversed phase high

performance liquid chromatography, characterized by liquid chromatography-mass spectrometry, and quantified by UV spectrophotometry as reported previously.[362]

2.2.4 Preparation of Radiolabeled Double-Stranded Oligodeoxy-nucleotides

Single stranded oligodeoxynucleotides containing the modified bases (250 pmol, **Scheme 2-2**C) in water were radiolabeled by incubation with T4 PNK (3 μL) and γ-^{32}P ATP (3 μL) at 37 °C for 60 min in PNK buffer (final volume = 20 μL). The reaction mixture was heated at 65 °C for 10 min to inactivate the enzyme and filtered through Illustra microspin G25 column (GE Healthcare, Pittsburgh, PA) to remove excess γ-^{32}P ATP. The 5'-^{32}P-labeled oligomers were mixed with 1–1.2 molar equivalents of the complementary strands in an annealing buffer (10 mM Tris [pH 7] containing 50 mM NaCl or 20 mM Tris-HCl [pH 7.6], 10 mM EDTA, and 150 mM NaCl), heated at 90 °C for 10 min, and allowed to cool slowly overnight to obtain double stranded DNA.

2.2.5 Base Excision Repair Assays with Human Fibrosarcoma Nuclear Extracts

Radiolabeled DNA duplexes containing site-specific BD-dA adducts (50 nM) were incubated at 37 °C in 10 mM HEPES (pH 7.4) buffer containing 100 mM KCl, 1 mM EDTA, 1 mM EGTA and 0.1 mM DTT. For initial assays examining concentration dependence for repair, 0–12 μg of nuclear protein extract from human fibrosarcoma cells

were added. The repair reactions were conducted in a total volume of 20 μL and stopped after 2 h incubation by heating at 95 °C for 5 min.

To observe time dependent repair, 25 μg of nuclear extract was added to 50 nM ^{32}P-endlabled adducted DNA in a total volume of 50 μL. Aliquots (10 μL) were withdrawn at 0, 15, 30, 75, 120, and 180 min, and quenched by the addition of 10 μL of gel loading buffer (20 mM EDTA in 95% formamide containing 0.05% bromophenol blue and 0.05% xylene cyanol).

To establish a role of base excision repair in adduct removal, DNA duplexes were pre-incubated with a known BER inhibitor, methoxyamine (MX, 3 mM) in the reaction buffer at room temperature for 60 min prior to the addition of the nuclear extract, followed by time dependent repair assay as described above.

2.2.6 Base Excision Repair Assays Using Recombinant Enzymes

Glycosylase assays were performed under single-turnover (STO) conditions by incubating 10–20 nM ^{32}P-endlabeled double-stranded DNA substrates with 10–200-fold excess of the enzyme at 37 °C in a final volume of 20 μL. The assay buffer contained 20 mM Tris-HCl (pH 7.6), 10 mM EDTA, 100 μg/mL BSA, while the experiments using edited or unedited hNEIL1 included 60 mM NaCl. The AAG reaction buffer was 20 mM Tris-HCl buffer (pH 7.8) containing 100 mM KCl, 5 mM β-mercaptoethanol, 2 mM EDTA, 1 mM EGTA, and 50 μg/mL BSA. Manufacturer provided reaction buffer was used for Mug (Trevigen, Gaithersburg, MD). Reactions were allowed to go for 60 min

and quenched by the addition of NaOH to a final concentration of 0.2 M. The solutions were heated at 90 °C for 5 min and then placed on ice until ready to load on a gel.

For selected enzymes (AAG, NEIL1 and Mug), strand breaks were initiated using AP endonuclease. In these cases, repair reactions were conducted for 60 min, followed by heating at 90 °C for 5 min and cooling to room temperature. Human AP endonuclease (0.5 units, Trevigen, Gaithersburg, MD) was added, and the mixtures were incubated at 37 °C in commercial reaction buffer for 60 min, followed by heating at 90 °C for 5 min. Samples were placed on ice until analysis.

2.2.7 Monitoring Base Excision Repair by Gel Electrophoresis

For gel electrophoresis separations, reaction mixtures were mixed with formamide denaturing dye (80% formamide, 0.025% xylene cyanol, 0.025% bromophenol blue in TBE running buffer) or gel loading buffer (20 mM EDTA in 95% formamide containing 0.05% bromophenol blue and 0.05% xylene cyanol) immediately prior to loading on to a denaturing PAGE gel.

Method 1: Samples were loaded onto a 15% denaturing polyacrylamide gel pre-ran in TBE running buffer (89 mM Tris, 89 mM boric acid, and 2 mM EDTA) and run at 1200 V for 2 hours. Radiolabeled DNA fragments were visualized using autoradiography by exposure to a storage phosphor screen overnight. Gels were imaged using storage phosphor autoradiography, and the bands were visualized using GraFit 5.0 (Erithacus Software Ltd., Horley, Surrey, UK).

Method 2: A 20 % denaturing polyacrylamide gel containing 7 M urea was pre-run at a constant power of 15 W for 30 min in TBE running buffer. The samples from repair experiments in gel loading buffer were loaded onto the gel and ran at a constant power of 15 W at ambient temperature. Radiolabeled 18-mer DNA strands and 5′ excision products were detected on a Typhoon FLA 7000 (GE Healthcare, Pittsburgh, PA). The extent of repair was evaluated by volume analysis using ImageQuant TL 8.0 (GE Healthcare, Pittsburgh, PA). The rate constants were calculated by fitting the percent DNA remaining vs. time plots to first order rate equation using Origin 9.1 software (OriginLab Corp., Northampton, MA).

2.2.8 Analysis of Base Excision Repair by Liquid Chromatography-Tandem Mass Spectrometry

For mass spectrometry experiments, repair assays were performed using 10 μM unlabeled adducted DNA duplexes (**Scheme 2-2**D) in 10 mM HEPES (pH 7.4) buffer containing 100 mM KCl, 1 mM EDTA, 1 mM EGTA and 0.1 mM DTT in a total volume of 20 μL. Following the addition of nuclear protein extract (5 μg), the reaction mixtures were incubated at 37 °C for 3 h. Samples were subjected to solid phase extraction using Sep-Pak C18 cartridges (100 mg/1mL) according to the manufacturer suggested protocol (Waters, Milford, MA). SPE fractions containing DNA (70% methanol elutions) were dried *in vacuo* and reconstituted in water (20 μL) prior to HPLC-ESI$^-$-MS/MS analysis.

Capillary HPLC-ESI$^-$-MS/MS analysis was performed on an Agilent 1100 capillary HPLC-ion trap mass spectrometer (Agilent Technologies, Inc., Santa Clara,

CA). Liquid chromatography was performed on a Zorbax SB-C18 column (150 mm × 0.5 mm, 5 μm, Agilent Technologies, Inc., Santa Clara, CA). The column was eluted at a flow rate of 15 μL/min using a gradient of 15 mM ammonium acetate (A) and acetonitrile (B). The column temperature was maintained at 25 °C. The solvent composition was changed linearly from 0 to 20% B over 20 min, then to 80% B over 2 min, kept at 80% B for further 2 min, and decreased to 0% B in 1 min. The mass spectrometer was operated in the ESI^- mode with target ion abundance value set to 50000, and the maximum accumulation time was 300 ms. In a typical experiment, 6–10 average scans were taken in the mass range of m/z 200–1600. Nitrogen was used as a nebulizing gas (15 psi) and a drying gas (5 L/min, 200 °C), while electrospray ionization was achieved at a spray voltage of 3–3.5 kV. The identities of the expected 5' excision products were confirmed by comparing their MS/MS spectra to experimental MS/MS spectra of synthetic standards (5'-TCA T-3' and 5'-CGG AC-3') and theoretical fragmentation patterns from Mongo oligo mass calculator version 2.06 (The RNA Institute, College of Arts and Sciences, State University of New York at Albany).

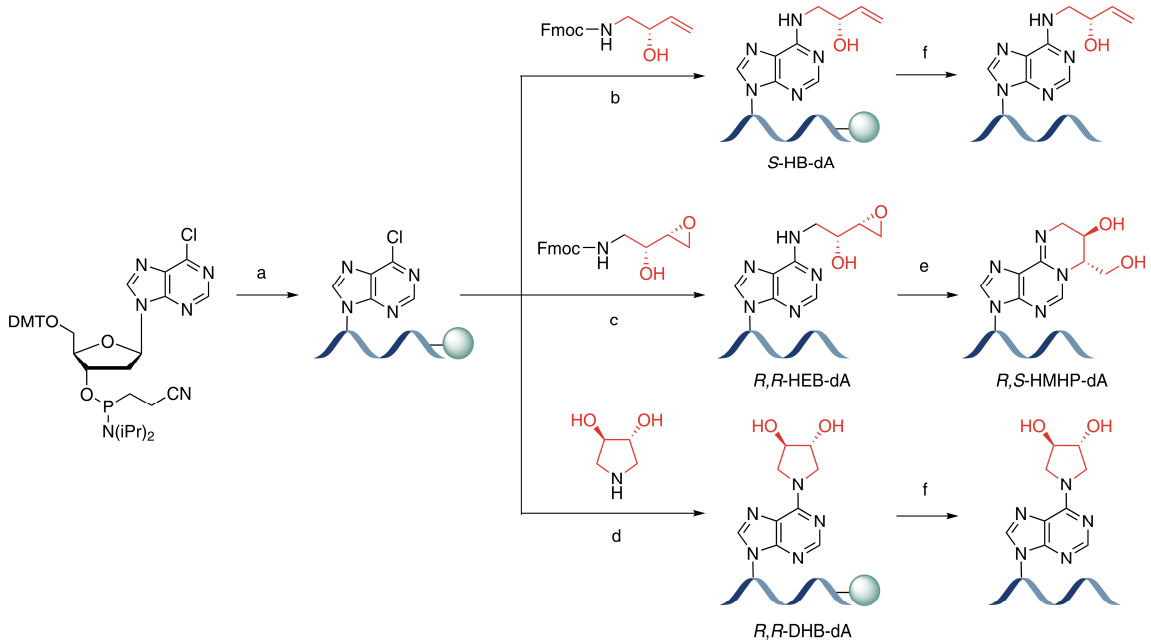
2.3 RESULTS

2.3.1 Synthesis, Purification and Characterization of Adducted DNA Oligodeoxynucleotides

Two DNA sequences were selected for the present study. DNA 11-mer, 5'-C GGA C***X***A GAA G-3', was derived from the human *N-ras* proto-oncogene, with BD-dA adducts (***X***) inserted at the second position of the *N-ras* codon 61.[364] The sequence of the DNA 18-mer, 5'-TCA T***X***G AAT CCT TCC CCC-3' (***X*** = BD-dA adduct), was engineered to be used in future *in vitro* polymerase bypass assays. A post-oligomerization methodology was used to generate synthetic DNA oligonucleotides containing site-specific (*S*)-N^6-HB-dA, (*R*,*R*)-N^6,N^6-DHB-dA, or (*R*,*S*)-1,N^6-HMHP-dA (**Scheme 2-3**).[362] In this approach, which was originally introduced by Harris *et al.*,[365] the inherent electrophilic-nucleophilic functionalities of a carcinogen and a DNA nucleobase are reversed (**Scheme 2-3**). Solid phase synthesis with commercial 6-chloropurine-2'-deoxyribose phosphoramidite (ChemGenes Corp., Wilmington, MA) was used to site-specifically incorporate a 6-chloropurine moiety at position ***X*** (5'-TCA T***X***G AAT CCT TCC CCC-3' and 5'-C GGA C***X***A GAA G-3'). Synthons representing 1,3-butadiene derived side chains conjugated to an amino group were synthesized by previously reported methods [(*S*)-*N*-Fmoc-1-aminobut-3-en-2-ol and (*R,R*)-*N*-Fmoc-1-amino-2-hydroxy-3,4-epoxybutane][78] or hydrogenation of commercially obtained (*3R,4R*)-1-benzyl-3,4-pyrrolidinediol (Sigma-Aldrich, Milwaukee, WI). Nucleophilic aromatic substitution (S_NAr) reactions were carried out at position ***X*** on synthetic DNA to displace

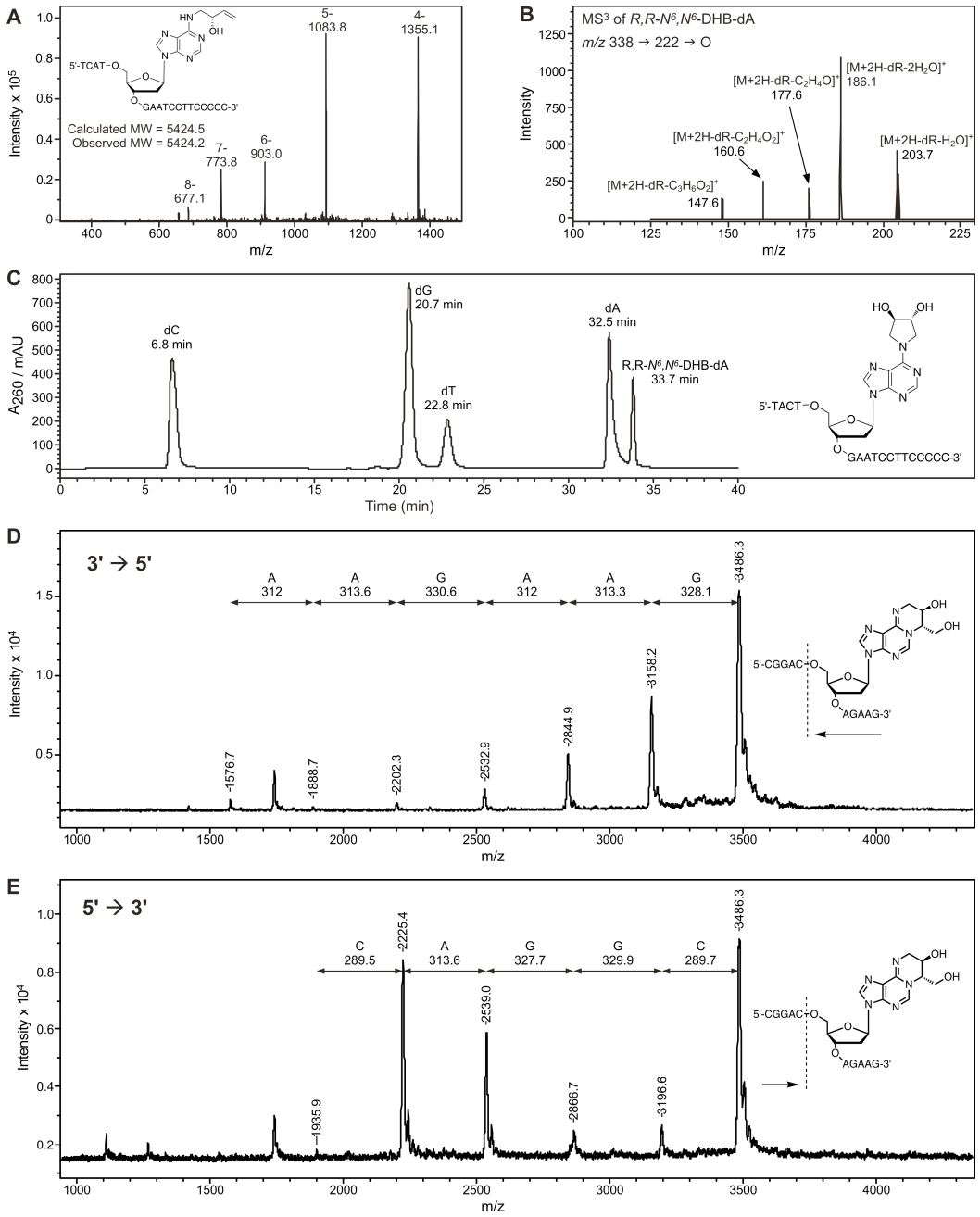
the chloro group of 6-chloropurine with the amine-functionalized synthon to generate the corresponding regio- and stereospecifically modified oligomers (**Scheme 2-3**).[362] In the case of (*R,S*)-1,N^6-HMHP-dA adducts, an additional cyclization step is required following S_NAr coupling between (*R,R*)-*N*-Fmoc-1-amino-2-hydroxy-3,4-epoxybutane and 6-chloropurine containing DNA (**Scheme 2-3**).[362] In order to avoid multiple HPLC purifications and to maximize the yield, (*S*)-*N*-Fmoc-1-amino-3-buten-2-ol and (*R,R*)-pyrrolidine-3,4-diol were coupled to 6-chloropurine containing oligomers on solid support to generate site- and stereospecific (*S*)-N^6-HB-dA and (*R,R*)-N^6,N^6-DHB-dA lesions, respectively (**Scheme 2-3**). All structurally modified DNA oligomers were purified by reversed phase HPLC and characterized by capillary HPLC-ESI^--MS, HPLC-UV and ESI^+-MS^3 analyses of enzymatic digests, and MALDI-TOF-MS of controlled exonuclease digests to ensure their identity and purity (**Figure 2-1**). HPLC pure oligomers were quantified by UV spectrophotometry.

Scheme 2-3 Synthesis of site- and stereospecific oligodeoxynucleotides containing 1,3-butadiene-induced 2'-deoxyadenosine adducts using a post-oligomerization approach.[£]



[£] (a) solid phase synthesis, (b) DIPEA, DMSO, 60 °C, 24 h, (c) i. 0.1 M NaOH, rt, 3 d, ii. DIPEA, DMSO, 37 °C, 24 h, (d) DIPEA, DMSO, 37 °C, 72 h, (e) H_2O, rt, 3h, (f) 0.1 M NaOH, rt, 3d.

Figure 2-1 Mass spectrometry-based characterization of 1,3-butadiene-induced 2'-deoxyadenosine adducts: a representative capillary HPLC-ESI$^-$-MS spectrum of 18-mer (*S*)-N^6-HB-dA (A), ESI$^+$-MS^3 spectrum (B) and HPLC-UV trace (C) of enzymatic digests of 18-mer (*R,R*)-N^6,N^6-DHB-dA, and MALDI-TOF-MS spectra of controlled exonuclease digests of 11-mer (*R,S*)-1,N^6-HMHP-dA (D and E).



Site-specific 18-mer oligodeoxynucleotides containing 5-flouro-dU, guanidino-hydantoin, and 1,N^6-etheno-dA modifications (**Scheme 2-2**B) were employed as positive controls as they are known substrates of mismatch uracil DNA glycosylase (MUG),[366] NEIL1,[367] and alkyl adenine DNA glycosylase (AAG),[126] respectively. Synthetic DNA containing 5-flouro-dU, 8-oxo-dG and 1,N^6-etheno-dA were prepared by solid phase synthesis. 8-oxo-dG was oxidized with Na_2IrCl_6 to generate guanidinohydantoin adducted DNA as reported elsewhere.[363] All DNA strands were purified by reversed phase high performance liquid chromatography, characterized by liquid chromatography-mass spectrometry (**Table 2-1**), and quantified by UV spectrophotometry.

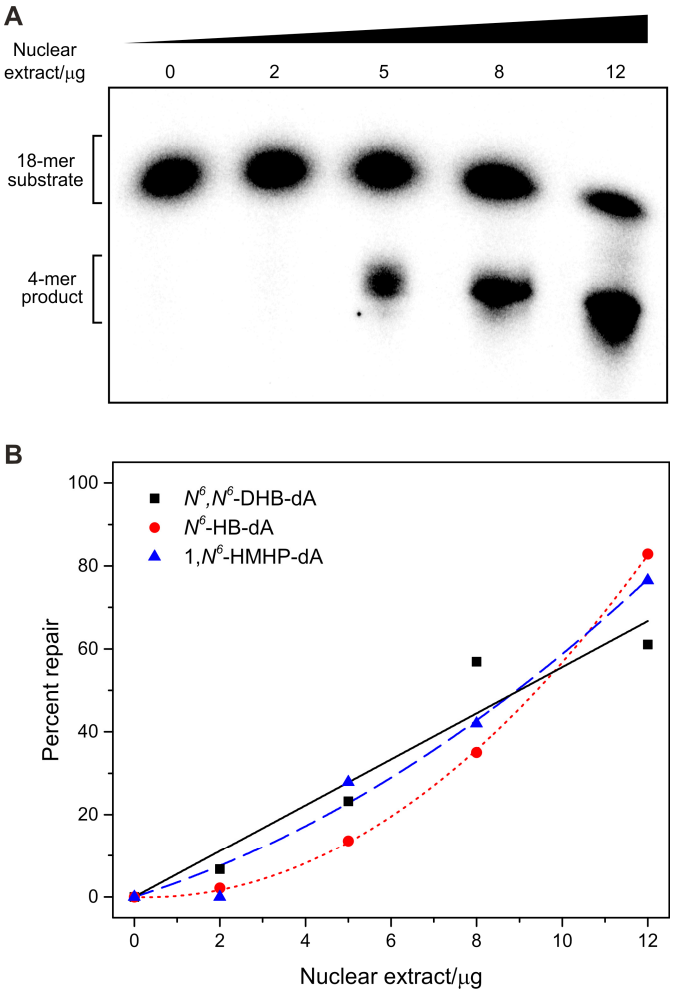
Table 2-1 Capillary HPLC-ESI$^-$-MS characterization of synthetic DNA oligomers.

Oligodeoxynucleotide sequence	*Molecular weight/kDa*	
	Calculated	*Observed*
5'-TCA T***X***G AAT CCT TCC CCC-3'; ***X*** = 6-Cl-Pu	5373.4	5373.5
5'-TCA T***X***G AAT CCT TCC CCC-3'; ***X*** = *S*-N^6-HB-dA	5424.5	5424.2
5'-TCA T***X***G AAT CCT TCC CCC-3'; ***X*** = *R*,*S*-1,N^6-HMHP-dA	5440.5	5440.3
5'-TCA T***X***G AAT CCT TCC CCC-3'; ***X*** = *R*,*R*-1,N^6-DHB-dA	5440.5	5440.1
5'-TCA T***X***G AAT CCT TCC CCC-3'; ***X*** = 1,N^6-εdA	5378.5	5378.9
5'-TCA T***Y***G AAT CCT TCC CCC-3'; ***Y*** = 8-oxo-dG	5386.5	5386.1
5'-TCA T***Y***G AAT CCT TCC CCC-3'; ***Y*** = Gh	5376.5	5376.5
5'-TCA T***Z***G AAT CCT TCC CCC-3'; ***Z*** = 5F-dU	5349.4	5349.0
5'-GGG GGA AGG ATT C***T***A TGA-3'	5643.7	5643.4
5'-GGG GGA AGG ATT C***C***A TGA-3'	5628.7	5628.6
5'-GGG GGA AGG ATT C***A***A TGA-3'	5652.8	5652.2
5'-C GGA C***X***A GAA G-3'; ***X*** = 6-Cl-Pu	3418.2	3418.1
5'-C GGA C***X***A GAA G-3'; ***X*** = *R*,*S*-1,N^6-HMHP-dA	3485.2	3485.0
5'-C TTC TTG TCC G-3'	3274.1	3273.8
5'-TCA T-3'	1148.2	1148.2
5'-CGG AC-3'	1487.3	1487.1

2.3.2 Base Excision Repair of 1,3-Butadiene-Induced 2'-Deoxyadenosine Adducts Using Human Fibrosarcoma Nuclear Extracts

To determine whether the BD-dA adducts can be recognized by base excision repair mechanism, radiolabeled 18-mer duplexes (^{32}P-5'-TCA T***X***G AAT CCT TCC CCC-3') containing site- and stereospecific (*R,S*)-1,N^6-HMHP-dA, (*S*)-N^6-HB-dA and (*R,R*)-N^6,N^6-DHB-dA at position ***X*** were incubated with nuclear protein extracts from human fibrosarcoma (HT1080) cells. Repair reactions were resolved on a 20% denaturing polyacrylamide gel, and the percentage of the excision product (5'-TCA T-3') was calculated by volume analysis. We found that all three BD-dA adducts used in this study were excised in a concentration dependent manner, with up to 90% excision observed in the presence of 12 μg of the extract (**Figure 2-2**).

Figure 2-2 Concentration dependent repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts by human fibrosarcoma nuclear extracts. 50 nM ^{32}P-endlabeled 18-mer DNA (5'-TCA T***X***G AAT CCT TCC CCC-3') duplexes were incubated in 10 mM HEPES (pH 7.4), 100 mM KCl, 1 mM EDTA, 1 mM EGTA and 0.1 mM DTT with increasing amounts of HT1080 nuclear extracts at 37 °C for 2 h. Samples were resolved on a 20% denaturing PAGE gel and visualized by phosphorimaging. (A) A representative PAGE gel for concentration dependent incision of 18-mer containing (*R,S*)-1,N^6-HMHP-dA, and (B) volume analysis showed increasing amounts of incision products for BD-dA adducted dsDNA with increasing amounts of nuclear extract.



To investigate the kinetics of repair of BD-dA adducts by human nuclear proteins, the same radiolabeled duplexes were incubated with HT1080 nuclear extracts, and the reactions were quenched at pre-selected time points. The cleavage products were analyzed by phosphorimaging following separation on a denaturing urea PAGE gel (**Figure 2-3**A). The repair assay was conducted in triplicate, and percent repair was calculated by volume analysis. We found that all three BD-dA adducts were repaired efficiently by human nuclear extracts (**Figure 2-3**B–D, solid lines). First order rate constants (**Table 2-2** and **Figure 2-4**A) revealed that the efficiency of repair decreased in the order of (*S*)-N^6-HB-dA > (*R*,*R*)-N^6,N^6-DHB-dA > (*R*,*S*)-1,N^6-HMHP-dA (0.0208, 0.0152 and 0.0119 min^{-1}, respectively). The extent and the rate of repair were decreased significantly in the presence of a known BER inhibitor, methoxyamine (**Figure 2-3**B–D, dotted lines). First order rate constants were ~3–8-fold smaller in the presence of the inhibitor (**Table 2-2**). Interestingly, the extent of inhibition was greater for (*R,S*)-1,N^6-HMHP-dA (~5-fold) and (*R,R*)-N^6,N^6-DHB-dA (~8-fold) as compared to (*S*)-N^6-HB-dA (~3-fold) (**Table 2-2** and **Figure 2-3**).

Figure 2-3 Time dependent repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts by human fibrosarcoma nuclear extracts: a representative PAGE gel of incision of 18-mer (5'-TCA T***X***G AAT CCT TCC CCC-3') containing (*R,S*)-1,N^6-HMHP-dA (A), time dependent incision of 18-mers containing (*S*)-N^6-HB-dA (B), (*R,S*)-1,N^6-HMHP-dA (C), (*R,R*)-N^6,N^6-DHB-dA (D) in the presence (dotted lines) and absence (solid lines) of BER inhibitor, methoxyamine. 50 nM ^{32}P-endlabeled dsDNA 18-mers were incubated with 0.5 µg/µL nuclear extract in 10 mM HEPES (pH 7.4), 100 mM KCl, 1 mM EDTA, 1 mM EGTA and 0.1 mM DTT at 37 °C. Aliquots of the reaction mixture were quenched at preselected time points, samples were resolved on a 20% denaturing PAGE gel and visualized by phosphorimaging. Volume analysis showed increasing amounts of incision products with increasing incubation time (n = 3).

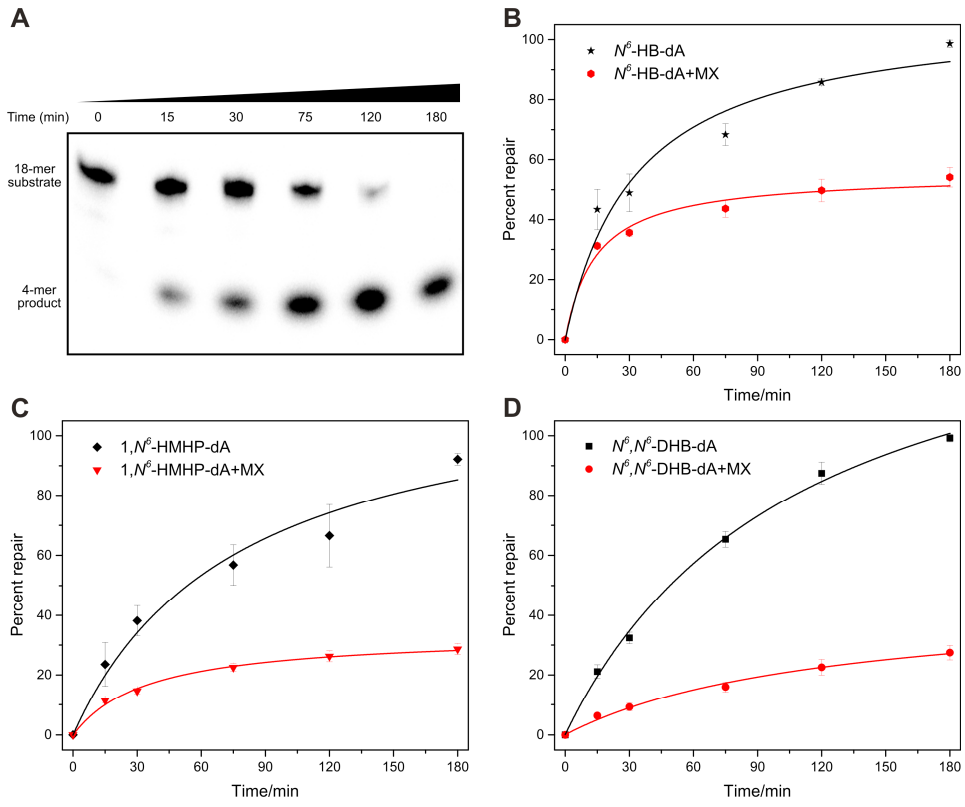
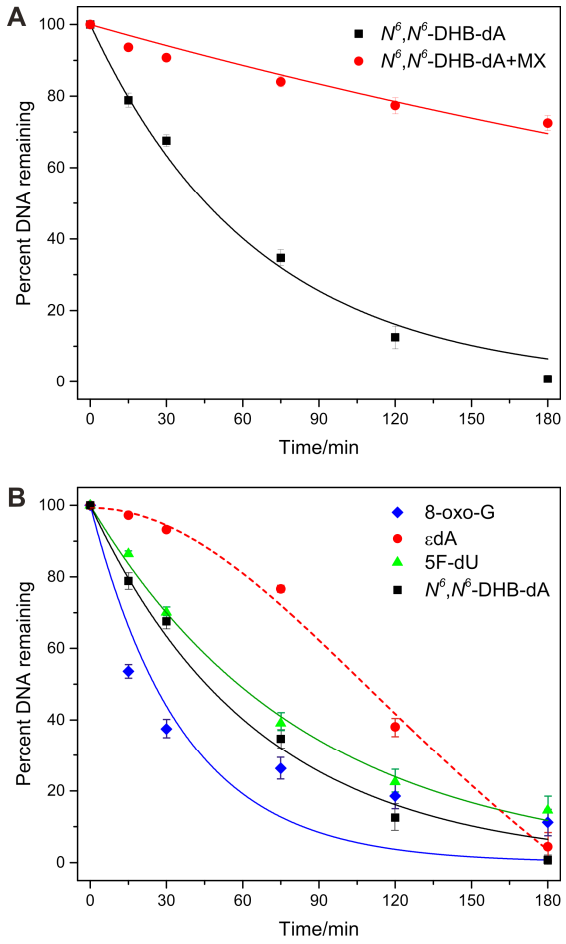


Table 2-2 First order rate constants observed for the repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts and known BER substrates by human fibrosarcoma nuclear extracts.[€]

DNA adduct	*BER inhibitor (MX)*	*Rate constant ± SE/min^{-1}*
(*S*)-N^6-HB-dA	×	0.0208±0.0034
(*S*)-N^6-HB-dA	✓	0.0064±0.0002
(*R*,*S*)-1,N^6-HMHP-dA	×	0.0119±0.0013
(*R*,*S*)-1,N^6-HMHP-dA	✓	0.0024±0.0004
(*R*,*R*)-N^6,N^6-DHB-dA	×	0.0152±0.0009
(*R*,*R*)-N^6,N^6-DHB-dA	✓	0.0020±0.0002
8-oxo-dG	×	0.0275±0.0061
5-F-dU	×	0.0119±0.0004

[€] SE = standard error.

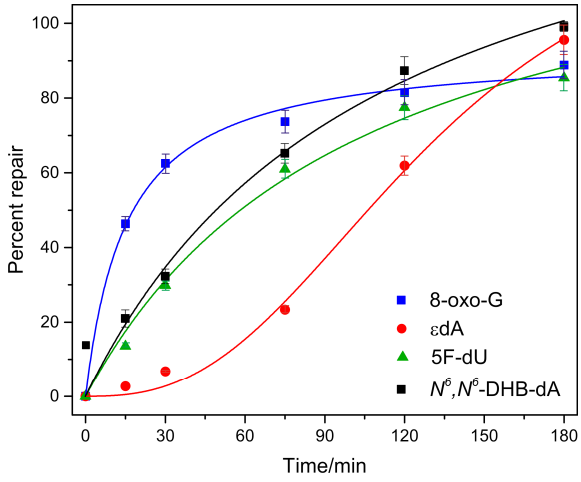
Figure 2-4 Repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts and known BER substrates by human fibrosarcoma nuclear extracts. Percent DNA remaining was plotted against time and data were fitted to the first order rate equation to calculate the rate constants. (A) Repair of (*R,R*)-N^6,N^6-DHB-dA in the presence and absence of BER inhibitor, methoxyamine (MX) and (B) repair of known BER substrates.[¥]



[¥] εdA showed cooperative binding, hence data cannot be fitted to a first order rate equation.

To compare the efficiency of BER of the BD-dA adducts by human fibrosarcoma nuclear extracts to that of known BER substrates, we have prepared DNA duplexes containing 1,N^6-etheno-dA (a known AAG substrate),[126] 5-fluoro-dU (a known MUG substrate),[366] and 8-oxo-dG (a known OGG1 substrate).[368] Repair reactions were conducted the same way as for BD-dA adducts. We found that BD-dA adducts were repaired with similar efficiency as compared to known BER substrates (**Figure 2-5**). According to the first order rate constants (**Table 2-2** and **Figure 2-4**B), all three BD-dA adducts were repaired 1–1.7-fold faster as compared to 5F-dU, while the rates were 1.3–2.3-fold slower with respect to 8-oxo-dG. Since εdA showed cooperative binding (percent DNA remaining vs. time plot yielded a sigmoidal curve),[369] it was not considered for rate comparisons (**Figure 2-4**B).

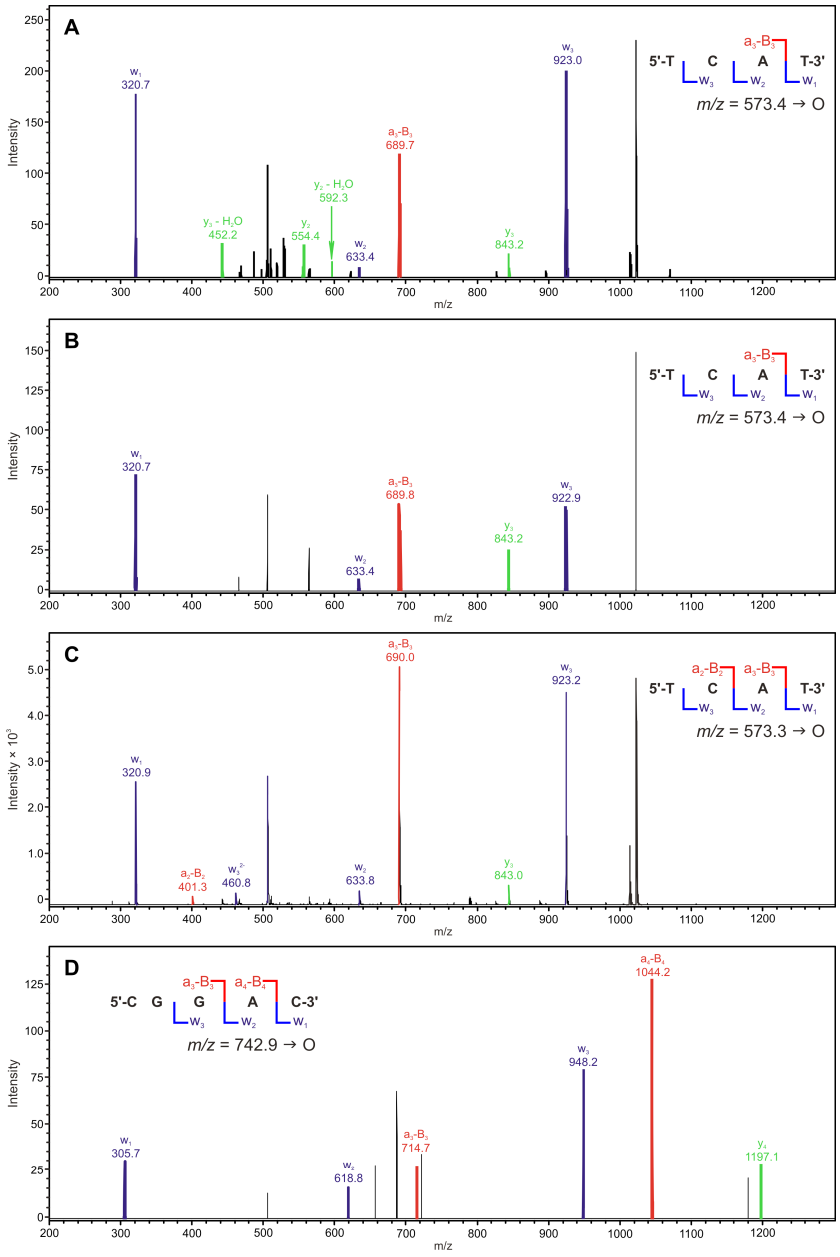
Figure 2-5 Comparison of repair rates of 1,3-butadiene-induced 2'-deoxyadenosine adducts against known BER substrates in human fibrosarcoma nuclear extracts. 50 nM ^{32}P-endlabeled dsDNA 18-mers were incubated with 0.5 μg/μL nuclear extract in 10 mM HEPES (pH 7.4), 100 mM KCl, 1 mM EDTA, 1 mM EGTA and 0.1 mM DTT at 37 °C. Aliquots of the reaction mixture were quenched at preselected time points, the samples were resolved on a 20% denaturing PAGE gel and visualized by phosphorimaging. Volume analysis showed increasing amounts of incision products with increased incubation time (n = 2).



2.3.3 Analysis of Base Excision Repair by Liquid Chromatography-Tandem Mass Spectrometry

A capillary HPLC-ESI$^-$-MS/MS methodology was used to further confirm the mechanism of base excision repair of BD-dA lesions. Following incubation of BD-dA containing 18-mers (5'-TCA T***X***G AAT CCT TCC CCC-3') with nuclear extracts from HT1080 cells, the reaction mixtures were purified by solid phase extraction and subjected to HPLC-ESI$^-$-MS/MS analysis. Tandem mass spectrometry employing collision-induced dissociation was used to sequence the oligonucleotide products. The expected BER excision product, (5'-TCA T-3') was observed for all three BD-dA adducts (**Figure 2-6**A–C). Similar results were obtained for repair reactions of unlabeled 11-mers containing (*R,S*)-1,N^6-HMHP-dA (5'-C GGA C***X***A GAA G-3'). HPLC-MS/MS analysis confirmed the generation of the 5' excision product, 5'-CGG AC-3' under the assay conditions (**Figure 2-6**D). These mass spectrometry results corroborate the gel electrophoresis data in **Figure 2-3**, confirming that BD-dA adducts can be efficiently recognized by the BER mechanism in human cells.

Figure 2-6 Liquid chromatography-tandem mass spectrometry analysis of repair products of 1,3-butadiene-induced 2'-deoxyadenosine adducts by human fibrosarcoma nuclear extracts. Collision induced dissociation (CID) spectra of the 5' excision products of 18-mers containing (*S*)-N^6-HB-dA (A), (*R,S*)-1,N^6-HMHP-dA (B), (*R,R*)-N^6,N^6-DHB-dA (C), and 11-mer containing (*R,S*)-1,N^6-HMHP-dA (D).



2.3.4 Activity of Recombinant BER Glycosylases against BD-dA Adducts

In an attempt to identify specific BER glycosylases involved in repair of BD-dA adducts, the glycosylase activity of 12 human and bacterial recombinant BER enzymes on site-specifically modified substrates was evaluated under single-turnover conditions (STO, enzyme concentration > DNA substrate concentration, **Figures 2-7** and **2-8**). We employed human AAG, Mpg, edited and unedited NEIL1 and OGG1, as well as six *E. coli* enzymes EndoIII, Fpg, Mug, MutY, MutYH and Nei, and *Archaeoglobus fulgidus* glycosylase, UDG. Although these glycosylases were active against their known substrates (results not shown), they showed no activity towards (*R*,*S*)-1,N^6-HMHP-dA, (*R*,*R*)-N^6,N^6-DHB-dA, and (*S*)-N^6-HB-dA (**Figure 2-7**). No cleavage was detected even when using 100–200-fold excess of AAG, NEIL1, Fpg and Mug or when incubating the reaction mixtures with AP endonuclease to maximize the formation of cleaved products (**Figure 2-8**). These results confirm that the BER enzymes tested here are not responsible for repairing BD-dA lesions.

Figure 2-7 Base excision repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts by human and bacterial BER enzymes. Repair of 18-mers containing (*R,R*)-N^6,N^6-DHB-dA (A) and (*S*)-N^6-HB-dA (B), and 11-mer containing (*R,S*)-1,N^6-HMHP-dA (C). 20 nM ^{32}P-endlabeled dsDNA substrates were incubated at 37 °C with 200 nM enzymes in 20 mM Tris-HCl (pH 7.6), 10 mM EDTA, 100 μg/mL BSA (buffer solution for hNEIL1 included 60 mM NaCl). After 60 min, reactions were quenched by the addition of NaOH, samples were resolved on a 20% denaturing PAGE gel and visualized by phosphorimaging.

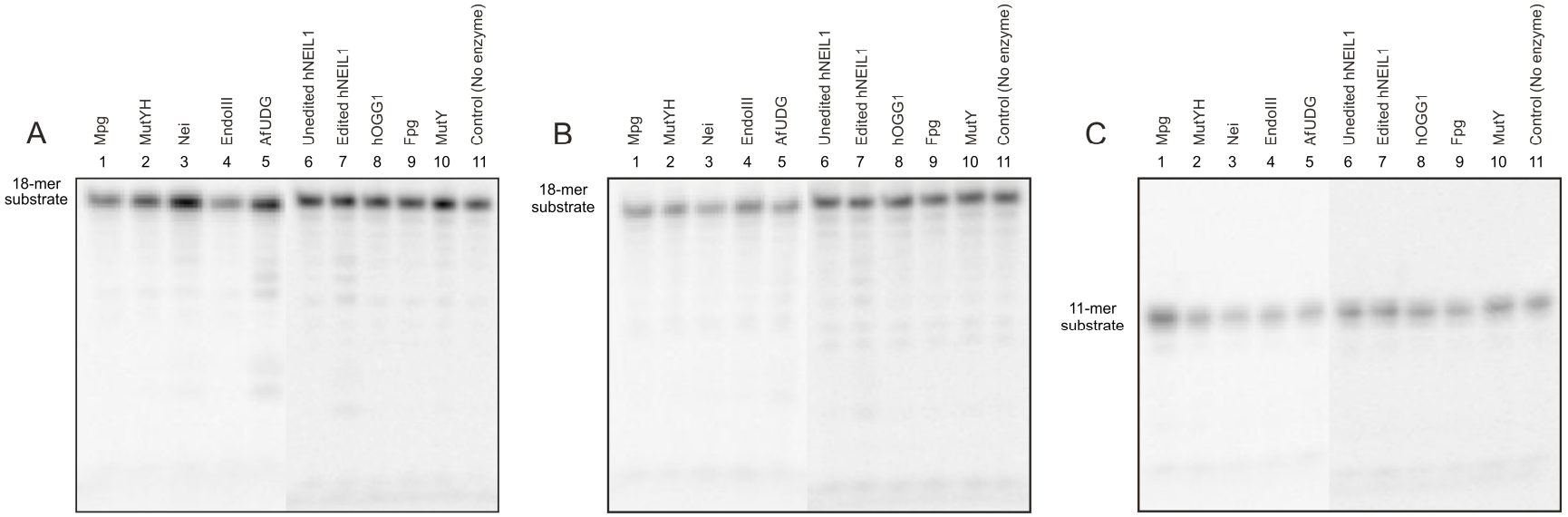
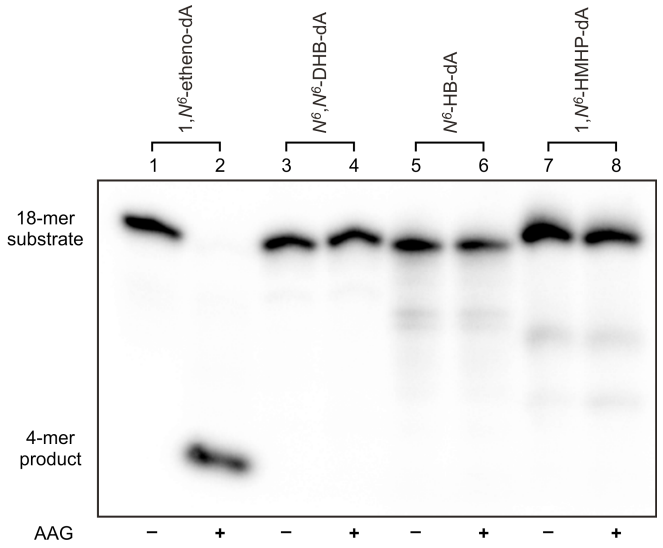


Figure 2-8 Base excision repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts by recombinant glycosylases: a representative PAGE gel of the repair assay in the presence of human alkyladenine DNA glycosylase (AAG). 10 nM ^{32}P-endlabeled synthetic DNA duplexes containing BD-dA adducts were incubated with 2 μM AAG in a buffered solution containing 20 mM Tris-HCl buffer (pH 7.8), 100 mM KCl, 5 mM β-mercapto-ethanol, 2 mM EDTA, 1 mM EGTA and 50 μg/mL BSA at 37 °C for 60 min. The reaction mixtures were incubated with human AP endonuclease (0.5 units) for another 60 min. Samples were resolved on a 20% denaturing PAGE gel and visualized by phosphor autoradiography. Although 1,N^6-etheno-dA, a known substrate of AAG was repaired (lane 1–2), none of the BD-dA adducts were recognized by the enzyme (lane 3–8).



2.4 DISCUSSION

The integrity and stability of genomic DNA is essential for life since DNA is the repository of hereditary information in cells.[1-3] DNA is not chemically inert, but possesses multiple reactive sites that can be modified by various chemical and physical agents resulting in damaged DNA.[1,3,6] Multiple cellular DNA damage response mechanisms have evolved to rescue the living cells from DNA modifications.[1,135] A number of DNA repair pathways operate in eukaryotic systems including direct repair,[1] mismatch repair (MMR),[1,7,121] base excision repair (BER),[1,6,121] nucleotide excision repair (NER),[1,116,118,121] homologous recombination (HR)[1,119] and non-homologous end joining (NHEJ) (see section 1.3.1).[1,120] Adduct size and shape, the extent of helix distortion and thermodynamic destabilization can influence the mechanisms of repair.[1,121]

In the present study, we have focused on the ability of human cells to recognize and repair three deoxyadenosine adducts induced by 1,3-butadiene: the monoadduct, N^6-(2-hydroxy-3-buten-1-yl)-2'-deoxyadenosine (N^6-HB-dA)[370] and two exocyclic adducts, 1,N^6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2'-deoxyadenosine (1,N^6-HMHP-dA) and N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine (N^6,N^6-DHB-dA).[362] Our previous circular dichroism (CD) studies suggested that DNA duplexes containing the three adducts maintain B-type DNA conformation.[362] NMR studies of (*R*,*R*)- and (*S*,*S*)-N^6,N^6-DHB-dA-containing DNA duplexes have revealed that the adduct is accommodated in the major groove of DNA. The modified nucleotide exists in the anti-conformation around the glycosidic bond, and with an out-of-plane rotation around the C6-N^6 bond to accommodate the bulky 3,4-dihydroxypyrrolidine moiety.[371] This

orientation of the adducted adenine disrupts the stacking interactions of both the adduct and its base pair with the neighboring 5' and 3' nucleotides.[371] Further, *bis*-alkylation of N^6 of adenine prevents hydrogen bonding with the thymidine in the opposite strand.[371] The (*S*)-N^6-HB-dA adduct was also positioned in the major groove such that the 2-hydroxy-3-butenyl moiety is oriented in the 3' direction with the anti-conformation around the glycosidic bond.[372] Only minor perturbations in stacking or hydrogen bonding interactions were observed.[372] These structural data suggests that NER mechanism, which recognizes bulky and helix distorting lesions may not be the major repair pathway responsible for recognizing these BD-dA adducts. However, the significant lowering of the melting temperatures of the BD-dA adducts placed opposite the correct base, thymidine suggests that the adducted DNA duplexes are thermodynamically destabilized. In addition, when these adducts were mispaired with adenine, the melting studies showed considerable stability in comparison to A:d mismatch.[362,370] This can possibly mean that these adducts can be repaired or replicated in an error-prone manner leading to mutations. Moreover, relatively mild effects of these adducts on DNA structure as revealed by NMR studies and their ability to destabilize DNA duplex observed in thermal melting studies suggest that BER can be responsible for the removal of these lesions.

Herein, we investigated the ability of BER to recognize and excise three BD-dA adducts: (*S*)-N^6-HB-dA, (*R*,*S*)-1,N^6-HMHP-dA and (*R*,*R*)-N^6,N^6-DHB-dA. We incorporated the 2'-deoxyadenosine adducts of interest site- and stereospecifically into synthetic DNA oligomers using solid phase DNA synthesis and post-oligomerization approach. Gel electrophoresis-based repair assays conducted using human fibrosarcoma

(HT1080) nuclear extracts have revealed that all three adducts were efficiently recognized and processed by BER pathway (**Figure 2-3**). The assay was designed such that the initial cleavage of the modified nucleobase by a BER glycosylase, and subsequent excision of the apurinic/apyrimidinic (AP) site by AP endonuclease (APE) can be detected using 5'-radiolabeled DNA. Inhibitor assays performed using a known BER inhibitor, methoxyamine (MX) have shown 50–75% reduction in repair, when compared to the assays conducted in the absence of the inhibitor. MX is a small molecule that binds to the AP site generated, when the adducted nucleobase is excised by a BER glycosylase, to form a Schiff base.[373,374] We also employed a liquid chromatography-tandem mass spectrometry assay to sequence the 5' excision products (**Figure 2-6**). We were able to detect the excision products under two different DNA sequence contexts, confirming that BER is responsible for the repair of these BD-dA adducts. However, our efforts to identify the glycosylases involved in removal of these lesions were not fruitful. We tested 12 human and bacterial BER glycosylases under single-turnover (STO) conditions. Yet, none of them showed any activity towards the adducts of interest (**Figures 2-7** and **2-8**). Currently, a mass spectrometry-based proteomics methodology is being developed in our laboratory to identify the individual glycosylases responsible for the repair of these BD-dA adducts.

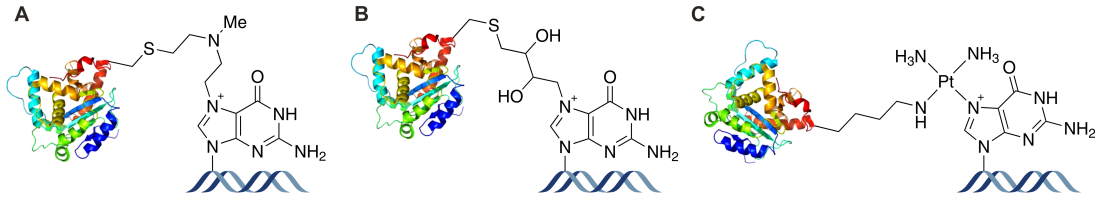
3 SYNTHESIS OF SEQUENCE-SPECIFIC DNA-PROTEIN CONJUGATES CROSS-LINKED TO 7-DEAZAGUANINE *VIA* A REDUCTIVE AMINATION STRATEGY

3.1 INTRODUCTION

Exposure to common antitumor drugs,[158,159,294] environmental toxins,[85,86,244] transition metals, [211,213,214] UV light,[259] ionizing radiation,[160] and free radical-generating systems[164,182,184] can result in cellular proteins becoming covalently trapped on DNA.[4] The resulting DNA-protein cross-links (DPCs) are unusually bulky, structurally diverse, and highly heterogeneous DNA lesions involving proteins of varying size, hydrophobicity, and cellular functions.[4,5,85,158] Mass spectrometry-based studies performed by Tretyakova and coworkers have revealed that a wide range of proteins can become covalently bound to genomic DNA upon treatment of human cells with clinically relevant concentrations of chemotherapeutic drugs (cisplatin and mechlorethamine) and metabolically activated carcinogens such as 1,2,3,4-diepoxybutane.[85,86,158] Some examples of the participating proteins include HSP 90, tubulins, DNA helicases, PCNA, Fen-1, Ku70, Ku86, Ref-1, PARP, and DNA polymerase δ.[85,158] MS/MS sequencing has shown that DNA-protein cross-linking is non-random, with specific amino acid side chains (cysteine, lysine, histidine, or arginine) participating in covalent conjugate

formation to the *N7* position of guanine in DNA (**Scheme 3-1**).[85,158,161] DPCs have been shown to accumulate in the brain and heart tissues with age.[36,37,167,181,198,234,250,251,308] Recent studies with laboratory mice deficient in the Fanconi Anemia DNA repair pathway have implicated DPC formed by formaldehyde in the observed cellular toxicity.[306,307]

Scheme 3-1 Structures of DNA-protein cross-links induced by *bis*-electrophiles. *N*-[2-(guan-7-yl)ethyl]alkylamine adducts formed by mechlorethamine (A); 4-(guan-7-yl)-2,3-butanediol adducts formed by 1,2,3,4-diepoxybutane (B), and 1,1-*cis*-diammine-2-alkylamino-2-(2'-deoxyguanosine-7-yl)-platinum(II) adducts induced by cisplatin (C).



Cys-N7G-EMA Cys-N7G-BD Lys-N7G-Pt

Due to their enormous size as compared to other DNA lesions, DPCs are believed to compromise genetic stability and cellular viability by interfering with normal DNA-protein interactions required for DNA replication, transcription, and repair.[4] We have recently engineered protein monoepoxide agents that specifically induce chromosomal DPCs, when electroporated in to cells.[304] These DNA-reactive proteins induced significant levels of mutations and toxicity when introduced into human cells,[304] probably because of the ability of the resulting DPCs to block DNA replication, transcription, and repair. However, relatively little is known about the influence of DPC adducts on DNA and RNA polymerases or their repair mechanisms in mammalian cells. Consequently, there is a pressing need to examine DNA replication and transcription using site-specific DPC lesions and to identify DNA repair mechanisms responsible for their removal in mammalian cells.

Any mechanistic investigations of the biological effects of DPC lesions in human cells require the availability of structurally defined DNA substrates containing DPC lesions at a specified site of DNA. However, the access to such DPC substrates has been limited due to the synthetic challenge of covalently linking two complex biomolecules (DNA and proteins) in a site-specific manner. Previously, model DPCs have been generated by covalently trapping various enzymes on their DNA substrates. For example, the Schiff base intermediate produced between T4 pyrimidine dimer glycosylase/AP lyase (T4-pdg) and apurinic/apyrimidinic site of DNA can be reduced to form a stable T4-pdg-DNA conjugate.[324] A disulfide trapping strategy was used to attach N149C mutant of human 8-oxoguanine DNA glycosylase I (hOGG1) protein to a DNA duplex

containing alkanethiol tether at the *N4* position of cytosine.[347] DNA methyltransferase has been cross-linked to *C6* position of 5-azacytosine.[331] More recently, a reductive amination strategy was used to generate DNA-protein/peptide conjugates by the reaction of the N^2-guanine aldehyde functionality derived from acrolein-induced 3-(2'-deoxyrobo-1'-syl)-5,6,7,8-tetrahydro-8-hydroxypyrimido[1,2*a*]purin-10(3*H*)-one (γ-HOPdG) with proteins and peptides.[246,336]

The most common site of DNA involved in DPC formation following treatment with *bis*-electrophiles is the *N7* of guanine (**Scheme 3-1**).[85,158,159,294] However, to our knowledge, no methods exist in the literature to generate *N7* guanine conjugated DPCs. One formidable obstacle in accomplishing this goal is that *N7* guanine alkylation destabilizes the *β*-glycosidic bond of the modified nucleoside, leading to spontaneous depurination.[375] In the present study, we have developed a new methodology to create hydrolytically stable structural mimics of *N7* guanine conjugated DPCs by reductive amination reactions between the Lys and Arg side chains of proteins and acetaldehyde functionalities of the modified 7-deazaguanine residues within DNA. The resulting model DPCs are structurally analogous to *N7* guanine adducts generated by antitumor nitrogen mustards, 1,2,3,4-diepoxybutane (**Scheme 3-1**) and chlorooxirane.[85,158,161,376]

3.2 MATERIALS AND METHODS

3.2.1 Materials

Synthetic oligodeoxynucleotides containing 7-deaza-7-(2,3-dihydroxypropan-1-yl)-2'-deoxyguanosine (deaza-DHP-dG) and deaza-DHP-dG nucleoside were prepared as previously described.[377] Fluorescein-dT phosphoramidite, protected 2'-deoxyribo-nucleoside-3'-phosphoramidites (dA-CE, Ac-dC-CE, dmf-dG-CE, dT-CE), Ac-dC-CPG ABI, dmf-dG-CPG ABI columns, and all other reagents required for automated DNA synthesis were purchased from Glen Research (Sterling, VA). Synthetic DNA oligonucleotides were synthesized by solid phase synthesis using an ABI 394 DNA synthesizer (Applied Biosystems, CA). All solvents and chemical reagents were obtained from commercial sources and used without further purification.

3.2.2 Preparation of Radiolabeled DNA Duplexes

Single stranded oligodeoxynucleotides 5'-G TCA CTG GTA ***deaza-DHP-dG***CA AGC ATT G-3' and 5'-C AGT GAC CAT C***deaza-DHP-dG***T TCG TAA C-3' (2 nmol in 12 μL of water) were radiolabeled with γ-^{32}P ATP using standard methods. Following heating at 65 °C for 10 min to inactivate the enzyme, excess γ-^{32}P ATP was removed using Illustra microspin G25 columns (GE Healthcare, Pittsburgh, PA). To obtain double stranded DNA, 5'-^{32}P-endlabeled oligomers were mixed with equimolar amounts of the complementary strands in 10 mM Tris buffer (pH 7) containing 50 mM NaCl and heated at 90 °C for 10 min, followed by gradual cooling overnight.

3.2.3 Reductive Amination to Generate DNA-Protein Cross-Links

^{32}P-endlabeled DNA strands (50 pmol in 8 μL water) were oxidized in the presence of 50 mM $NaIO_4$ (5 μL) in 15 mM sodium phosphate buffer (pH 5.4, 5 μL) for 6 h at 4 °C in the dark to unmask the aldehyde moiety on deaza-DHP-dG. Excess $NaIO_4$ was quenched with 55 mM Na_2SO_3 (5 μL). Proteins and peptides of interest (0.5–2.5 nmol) were incubated with the aldehyde-containing DNA (50 pmol) in the presence of 25 mM $NaCNBH_3$ at 37 °C overnight to generate stable DNA-protein cross-links. Aliquots of the reaction mixtures were withdrawn and resolved by 12% SDS-PAGE with or without proteinase K digestion (6 units, 48 h at 37 °C).

3.2.4 Gel Electrophoretic Analysis of DNA-Protein Cross-Links Generated by Reductive Amination

12% SDS-PAGE gel plates were pre-run at a constant voltage of 150 V for 30 min in 1×SDS running buffer. DPC reaction mixtures were dissolved in 0.1% TFA or 10% SDS (2 μL). The samples were reconstituted in SDS loading buffer and heated at 90 °C for 5 min prior to loading. The gels were run at a constant voltage of 150 V at ambient temperature. Radiolabeled DNA strands and DPCs were detected with a Storm 840 phosphorimager (Amersham Biosciences Corp., Piscataway, NJ) or a Typhoon FLA 7000 instrument (GE Healthcare, Pittsburgh, PA). Covalent DPCs were observed as slowly moving bands on the gel, and the reaction yields were calculated by volume analysis using Image Quant TL 8.0 (GE Healthcare, Pittsburgh, PA).

To visualize the proteins participating in DPC formation to aldehyde-containing DNA, NuPAGE Novex 12% Bis-Tris gels (Life Technologies, Grand Island, NY) were pre-run at a constant voltage of 100 V for 30 min in 1×NuPAGE MOPS SDS running buffer (Life Technologies, Grand Island, NY). The reaction mixtures obtained from DNA-protein cross-linking were dissolved in 10% SDS (2 μL) and reconstituted in NuPAGE LDS sample buffer (Life Technologies, Grand Island, NY). The samples were heated at 70 °C for 10 min prior to loading onto a gel. The gels were run at a constant voltage of 100 V at ambient temperature. The unreacted protein and DPC bands were visualized by staining with SimplyBlue SafeStain (Life Technologies, Grand Island, NY).

3.2.5 Sample Processing for Mass Spectrometry Analysis

DPCs were generated by reductive amination as described above using unlabeled DNA and proteins (15–20-fold excess) and either directly processed for mass spectrometric analysis or purified by PAGE as described below.

In direct processing experiments, the DNA component of DPCs was digested with PDE I (120 mU), PDE II (105 mU), DNase (35 U) and alkaline phosphatase (22 U) in 10 mM Tris-HCl/15 mM $MgCl_2$ (pH 7) buffer at 37 °C overnight. The resulting protein-nucleoside conjugates were dried *in vacuo*, reconstituted in 100 mM NH_4HCO_3 (pH 7.9, 90 μL), and subjected to trypsin digestion using MS grade Trypsin Gold (2.5 μg, Promega, Madison, WI) at 37 °C for 20 h. The digests were dried *in vacuo*, desalted using ZipTip with 0.6 μL C18 resin (Millipore, Billerica, MA), and the resulting peptides were reconstituted in 0.1% formic acid (10 μL) prior to MS analysis.

Alternatively, DNA-protein conjugates were first purified by 12% SDS-PAGE and stained with SimplyBlue SafeStain (Life Technologies, Grand Island, NY). DPC-containing gel bands were cut into slices and subjected to reduction using 300 mM DTT (10 μL) followed by alkylation with iodoacetamide (10 μL in 100 μL of 25 mM NH_4HCO_3, pH 7.9). Gel pieces were dehydrated with acetonitrile, dried under vacuum, reconstituted in 25 mM NH_4HCO_3 (pH 7.9, 75 μL) and incubated with PDE I (120 mU) at 37 °C overnight. Next, the samples were subjected to tryptic digestion and ZipTip desalting as described above, followed by MS analysis.

3.2.6 Characterization of DNA-Protein Cross-Links by Mass Spectrometry

All HPLC-ESI^+-MS/MS analyses were conducted with a Thermo Scientific LTQ Orbitrap Velos mass spectrometer interfaced with an Eksigent NanoLC-Ultra 2D HPLC system. Peptide mixtures (5 μL) were loaded onto a nano HPLC column (75 μm ID, 10 cm packed bed, 15 μm orifice) created by hand packing commercially purchased fused-silica emitters (New Objective, Woburn, MA) with Luna C18, 5 μm separation media (Phenomenex, Torrance, CA). Liquid chromatography was carried out at an ambient temperature at a flow rate 0.3 μL/min using 0.1% formic acid (A) and acetonitrile (B). The solvent composition was changed linearly from 2% to 70% B over 60 min, then to 95% B over 1 min, kept at 95% B for further 5 min, and decreased to 2% B in 1 min. Finally, the flow rate was increased to 1 μL/min and kept at 2% B for an additional 7 min. Mass spectrometry was performed using the FTMS mass analyzer with

a resolution of 60,000 ppm and with a scan range of *m/z* 300–2000. Peptide MS/MS spectra were collected using data-dependent scanning in which one full scan mass spectrum was followed by eight MS/MS spectra using an isolation width of 2.5 *m/z,* normalized CID collision energy of 35%, 1 repeat count, 20 s exclusion duration, with an exclusion mass width of ±5 ppm.

Spectral data were analyzed using Thermo Proteome Discoverer 1.3 (Thermo Scientific, 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 the protein FASTA database. Search parameters included trypsin specificity and up to 2 missed cleavage sites. Protein *N*-terminus, lysine, or arginine residues were specified as possible modification sites by specifying the following dynamic modifications: (A) ethan-1,2-diyl cross-link to 7-deaza-7-ethan-1-yl-2'-deoxy-guanosine, +292.1172 Da ($C_{13}H_{16}N_4O_4$); (B) cross-link to 7-deaza-7-ethan-1-yl-guanine, +176.0698 Da ($C_8H_8N_4O$); or (C) cross-link to 7-deaza-7-ethan-1-yl-2'-deoxyguanosine 5'-monophosphate, +372.0835 Da ($C_{13}H_{17}N_4O_7P$).

3.2.7 Synthesis and Characterization of 7-Deaza-7-(2-(N-acetyllysine) ethan-1-yl)-2'-deoxyguanosine and 7-Deaza-7-(2-(N-acetylarginine) ethan-1-yl)-2'-deoxyguanosine Conjugates

Synthetic 7-deaza-7-(2,3-dihydroxypropan-1-yl)-2'-deoxyguanosine[377] (10 nmol) was oxidized in the presence of 50 mM $NaIO_4$ (4 μL) in 1 M sodium phosphate buffer

(pH 5.4, 6 μL) for 6 h at 4 °C in the dark to generate 7-deaza-7-(formylmethan-1-yl)-2'-deoxyguanosine. Excess $NaIO_4$ was quenched with 55 mM Na_2SO_3 (4 μL). *N*-acetyl protected Lys or Arg (4 μL of 5 mM solution) was added to the reaction mixture followed by 0.5 M $NaCNBH_3$ (4 μL), and incubated at 37 °C overnight to generate amino acid-nucleoside conjugates. The amino acid-nucleoside conjugates were isolated by HPLC on a Supelcosil-LC-18-DB (4.6 × 250 mm, 5 μm) column (Sigma Aldrich, Milwaukee, WI) using a gradient of 0.1% formic acid (A) and acetonitrile (B). The solvent composition was changed from 0 to 24% B over 24 min, then to 75% B over 6 min, and finally to 0% B over 2 min. HPLC fractions containing major components were collected, concentrated *in vacuo* and analyzed by capillary HPLC-ESI^+-MS^n on an Agilent 1100 capillary HPLC-ion trap mass spectrometer (Agilent Technologies, Inc., Wilmington, DE).

3.2.8 Synthesis and Characterization of Nucleoside-Peptide Conjugates

Synthetic nucleoside-peptide conjugates were prepared using 7-deaza-7-(2,3-dihydroxypropan-1-yl)-2'-deoxyguanosine and angiotensin I (DRVYIHPFHL) or substance P ($RPKPQQFFGLMNH_2$) peptides by reductive amination procedure as described above. The reaction mixtures were dried *in vacuo* and desalted using ZipTip with 0.6 μL C18 resin (Millipore, Billerica, MA). The resulting mixtures were reconstituted in 0.1% formic acid (10 μL) and analyzed on a Thermo Scientific LTQ Orbitrap Velos mass spectrometer interfaced with an Eksigent NanoLC-Ultra 2D HPLC system as described below.

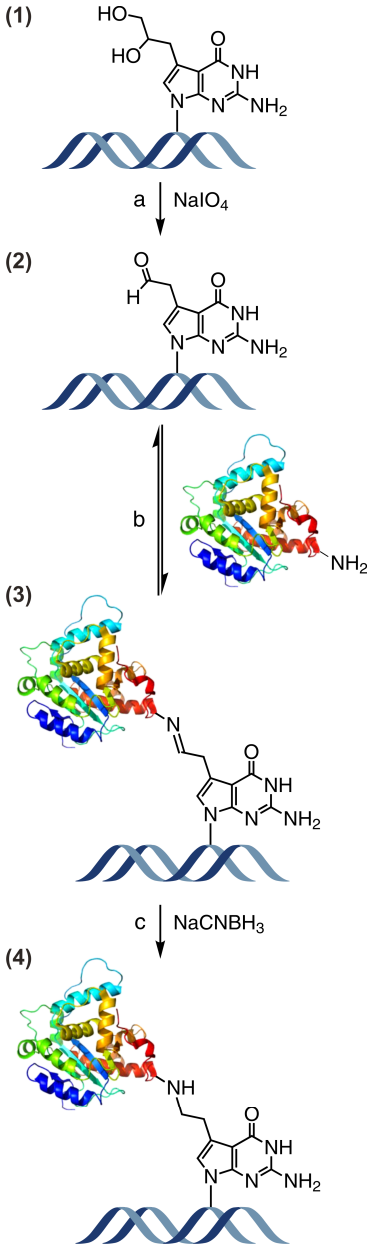
Desalted reaction mixtures (2 μL) were loaded onto a nano HPLC column (75 μm ID, 10 cm packed bed, 15 μm orifice) created by hand packing commercially purchased fused-silica emitters (New Objective, Woburn, MA) with Zorbax SB-C18, 5 μm separation media (Phenomenex, Torrance, CA). Liquid chromatography was carried out using 0.1% formic acid (A) and acetonitrile (B) as solvents at an ambient temperature at an initial flow rate of 1 μL/min at 2% B for 5.5 min. The flow rate was reduced to 0.3 μL/min in 30 s and the solvent composition was changed linearly from 2% to 50% B over 20 min, then to 95% B over 1 min, kept at 95% B for further 5 min, and decreased to 2% B in 1 min. Finally, the HPLC flow rate was increased to 1 μL/min and kept at 2% B for an additional 6 min. Mass spectrometry was performed using the FTMS mass analyzer with a resolution of 60,000 ppm and with a scan range of *m/z* 300–2000 in the full scan mode. MS^2 spectra were collected using the iontrap with an isolation width of 2.5 *m/z,* normalized CID collision energy of 35%, while MS^3 spectra were collected using the orbitrap with an isolation width of 2.5 *m/z,* and a normalized HCD collision energy of 35%.

3.3 RESULTS

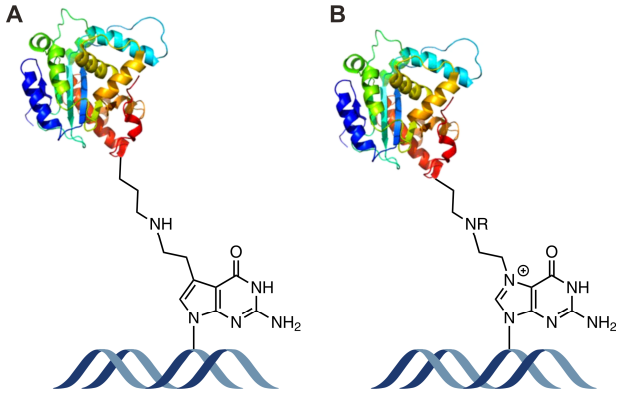
3.3.1 Experimental Strategy for the Generation of Hydrolytically Stable Model DNA-Protein Cross-Links

Since *N7*-guanine alkylation introduces a positive charge on the alkylated base, it destabilizes the *N*-glycosidic bond, leading to spontaneous depurination.[378] Therefore, it is not practical to employ *N7*-guanine adducts in DNA replication and repair experiments. To avoid spontaneous degradation of our model DPC substrates, we have replaced the *N7* nitrogen of guanine with a carbon atom (7-deaza-G).[377] To create a protein reactive group, the 2,3-dihydroxypropyl group was introduced at the same position (deaza-DHP-dG, 1 in **Scheme 3-2**). Treatment with periodate converts the diol group to the corresponding aldehyde (2 in **Scheme 3-2**), which then reacts with free amino groups of proteins (e.g., Lys or Arg side chains) to form a Schiff base (3 in **Scheme 3-2**). The latter can be quantitatively reduced with $NaCNBH_3$ to produce a stable amine linkage (4 in **Scheme 3-2**). The aldehyde substrate (2 in **Scheme 3-2**) is a direct model for *N7*-(2-oxoethyl)-G, which is the major DNA adduct from exposure to chlorooxirane,[376] and the resulting model cross-links are structurally analogous to DPCs formed by chlorooxirane and antitumor nitrogen mustards in cells (**Scheme 3-3**).

Scheme 3-2 Synthesis of DNA-protein cross-links by a post-synthetic reductive amination strategy. (a) Oxidative cleavage to unmask the reactive aldehyde moiety on the 7-deaza-7-(2,3-dihydroxypropan-1-yl)-guanine of the synthetic oligomer, (b) reaction of an amino group on the protein with the aldehyde on the DNA to form a Schiff base, and (c) reduction of the imine to form a stable amine linkage.



Scheme 3-3 Structural similarity between the model DNA-protein cross-links prepared by reductive amination and DNA-protein cross-links formed *in vivo*. DPCs obtained by reductive amination of aldehyde containing DNA and proteins (A) are structurally analogous to cellular DPCs formed upon reaction of antitumor nitrogen mustards with *N7* guanine of DNA and proteins (B).[158] Synthetic DPC substrates (A) are also a direct model for cross-linking of proteins to *N7*-(2-oxoethyl)-G, which is the major DNA adduct formed upon exposure to vinyl chloride.[376]



3.3.2 Characterization of DNA-Protein Cross-Link Formation by Denaturing Gel Electrophoresis

Our initial experiments were conducted with recombinant *E. coli* AlkB protein, a DNA repair protein that contains multiple nucleophilic lysine and arginine residues and is known to bind to DNA.[379,380] The formation of covalent AlkB-DNA conjugates was monitored by two independent methods. In the first approach, oligodeoxynucleotides containing the convertible nucleoside (deaza-DHP-dG) were radiolabeled with ^{32}P-ATP. Following the cross-linking reaction, DPC formation was detected as the appearance of a new, low mobility band on denaturing PAGE (**Figure 3-1**). Alternatively, free proteins and DNA-protein conjugates were visualized by protein staining, and the presence of a cross-link was detected as a new protein band with reduced mobility (**Figure 3-2**).

As shown in **Figure 3-1**, the cross-linking reaction between deaza-DHP-dG containing DNA 20-mer and recombinant AlkB protein led to the formation of covalent DPC conjugates as revealed by the appearance of a low mobility band on a denaturing polyacrylamide gel (Lane 3). This band was not observed in control experiments conducted in the absence of protein (Lane 1), and only trace amounts of conjugation were observed in the absence of the reducing agent (Lane 2). The DPC band disappeared when the reaction mixture was subjected to proteinase K digestion, (Lane 4), confirming that it corresponds to covalent DNA-protein conjugates.

Figure 3-1 Denaturing PAGE analysis of DNA-protein cross-links generated by post-synthetic reductive amination. DNA and DNA-protein conjugation products of the reaction between *E. coli* AlkB protein and aldehyde-containing DNA 20-mer, 5'-G TCA CTG GTA ***deaza-DHP-dG***CA AGC ATT G-3', were visualized by ^{32}P-end labeling. The formation of covalent DPC was revealed as a low mobility band on the gel. Lane 1: aldehyde containing oligonucleotide in the presence of a reducing agent (negative control); Lane 2: aldehyde containing oligonucleotide and AlkB protein in the absence of a reducing agent; Lane 3: reaction mixture of aldehyde containing oligonucleotide and the protein in the presence of a reducing agent; Lane 4: reaction mixture from lane 3 subjected to proteinase K digestion.

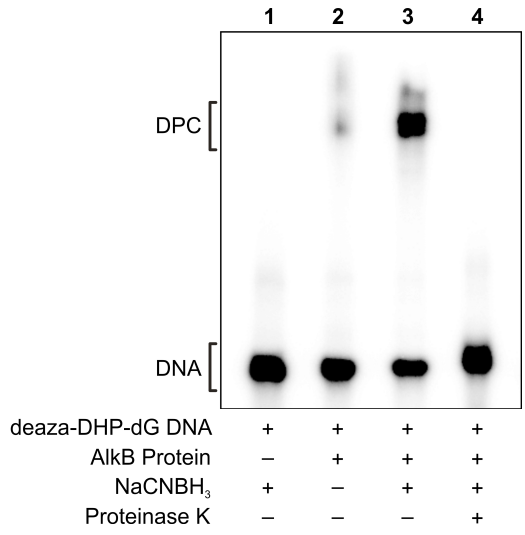
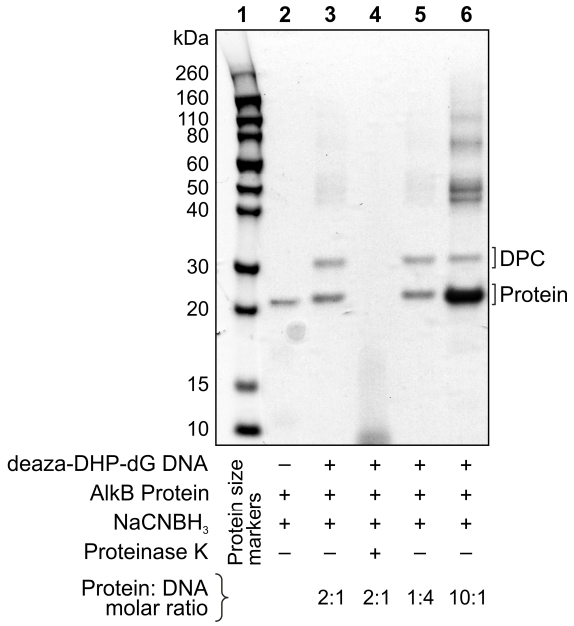


Figure 3-2 SDS-PAGE analysis of reductive amination-mediated DNA-protein cross-linking. DPC formation between aldehyde-containing DNA 20-mer (5'-G TCA CTG GTA ***deaza-DHP-dG***CA AGC ATT G-3') and *E. coli* AlkB protein was visualized *via* SimplyBlue protein staining. Lane 1: protein size markers; Lane 2: AlkB protein (negative control); Lane 3: DPCs generated using 1:2 ratio of DNA to protein; Lane 4: proteinase K digested reaction mixture from lane 3; Lane 5: DPCs generated using 4:1 molar ratio of DNA to protein; Lane 6: DPCs generated using 1:10 ratio of DNA to protein. Multiple bands were observed in the region >40 kDa due to protein aggregation.



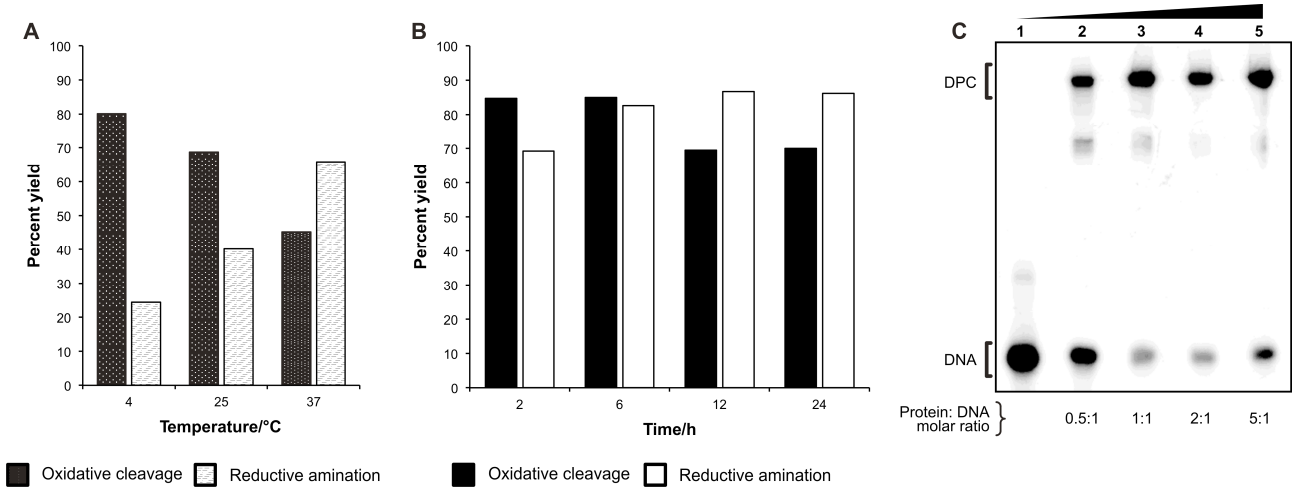
These results were further confirmed using protein staining to visualize the protein and the DPCs (**Figure 3-2**). Following reductive amination reaction between AlkB protein and deaza-DHP-dG containing 20-mer, a new band was observed with an increased molecular weight as compared to unreacted AlkB protein (22.9 kDa) (Lane 3). The size of the newly formed conjugate (~29 kDa) was consistent with the addition of 18-mer oligodeoxynucleotide (6.2 kDa) to the protein, and the cross-linking yield was dependent on DNA:protein ratio (Lanes 3, 5 and 6). In addition, several higher molecular weight bands (>40 kDa) were observed due to the propensity of the AlkB protein to oligomerize when present at a high concentration.

3.3.3 Effects of Reaction Conditions on DNA-Protein Cross-Link Yields

The experimental conditions for each of the reaction steps (**Scheme 3-2**) were optimized by varying the reaction temperature, reaction time, and molar ratios. We found that the highest DPC yields were observed when the $NaIO_4$-mediated oxidative cleavage (step **a** in **Scheme 3-2**) was conducted at 4 °C, and the optimal temperature for reductive amination reaction (steps **b** and **c** in **Scheme 3-2**) was 37 °C (**Figure 3-3**A).

When the effect of reaction times on DPC yields was assessed, the best results were achieved when the time of periodate-mediated oxidative cleavage (step **a** in **Scheme 3-2**) was limited to 2–6 h (**Figure 3-3**B). This can be explained by a limited stability of aldehydes under oxidizing conditions. In contrast, the best yields of reductive amination reaction (steps **b** and **c** in **Scheme 3-2**) were achieved at extended reaction times (12–24 h, **Figure 3-3**B).

Figure 3-3 Effects of reaction conditions on the yields of DNA-protein cross-links. Reductive amination-mediated cross-linking between deaza-DHP-dG containing DNA and the AlkB protein was performed, DPCs were visualized by phosphorimaging and quantified by volume analysis. (A) Influence of reaction temperature on DPC yields. Oxidative cleavage was most efficient at 4 °C and cross-linking reaction gave the highest yields at 37 °C. (B) Influence of reaction time on DPC yields. Oxidative cleavage was high yielding at shorter reaction times (2–6 h), while cross-linking reaction gave the highest yields of DPCs when incubated for longer reaction times (12–24 h). (C) Variation of DPC yields with increased protein:DNA molar ratio. Cross-linking reaction was carried out by varying the molar equivalents of AlkB protein to DNA between 0.5:1 to 5:1. The production of covalent conjugates was increased in the presence of excess protein.



The influence of DNA:protein molar ratios on reaction yields was examined by keeping the concentration of one of the reagents (protein or DNA) constant while varying the molar equivalents of the other. When protein amounts were varied with respect to constant amounts of radiolabeled DNA, the highest DPC yields (~85%) were achieved when using an excess of the protein (**Figure 3-3**C). When protein amounts were held constant and protein staining was employed to follow the reaction, increasing DNA concentrations similarly has led to increased DPC yields (**Figure 3-2**). These results suggest that the reversible Schiff base formation between the aldehyde functionality within DNA and the basic amino acid side chains of the protein can be driven forward towards product formation by employing a large molar excess of the other reagent. In a practical sense, generation of DNA substrates for replication and repair studies requires an excess of the protein in order to maximize the yields of DPC-containing DNA.

3.3.4 Influence of Protein Identity on DNA-Protein Cross-Link Formation

In theory, any protein containing lysine or arginine side chains can be cross-linked to deaza-DHP-dG containing DNA using the reductive amination strategy (**Scheme 3-2**). However, the reactivity and the accessibility of basic residues may vary depending on the protein identity. Therefore, the general applicability of our approach was examined using a range of proteins of different sizes and structures (**Table 3-1** and **Figure 3-4**A). We found that while some proteins (AlkB, NEIL1, Histone H4, GAPDH) formed DPCs in a high yield (75–95%), significantly lower DPC yields (<20%) were

observed for others (trypsin, carboxypeptidase, myoglobin, **Table 3-1**). In general, DNA binding proteins such as those involved in chromatin condensation and DNA repair produced DPCs in a higher yield than proteins that do not have an affinity for DNA. These results suggest that the formation of DPCs by reductive amination is facilitated by reversible DNA-protein interactions, which brings the two biomolecules into a close proximity to each other. Additional low mobility bands were observed for some proteins due to their dimerization (e.g. histone H4 and myoglobin, **Figure 3-4**A).

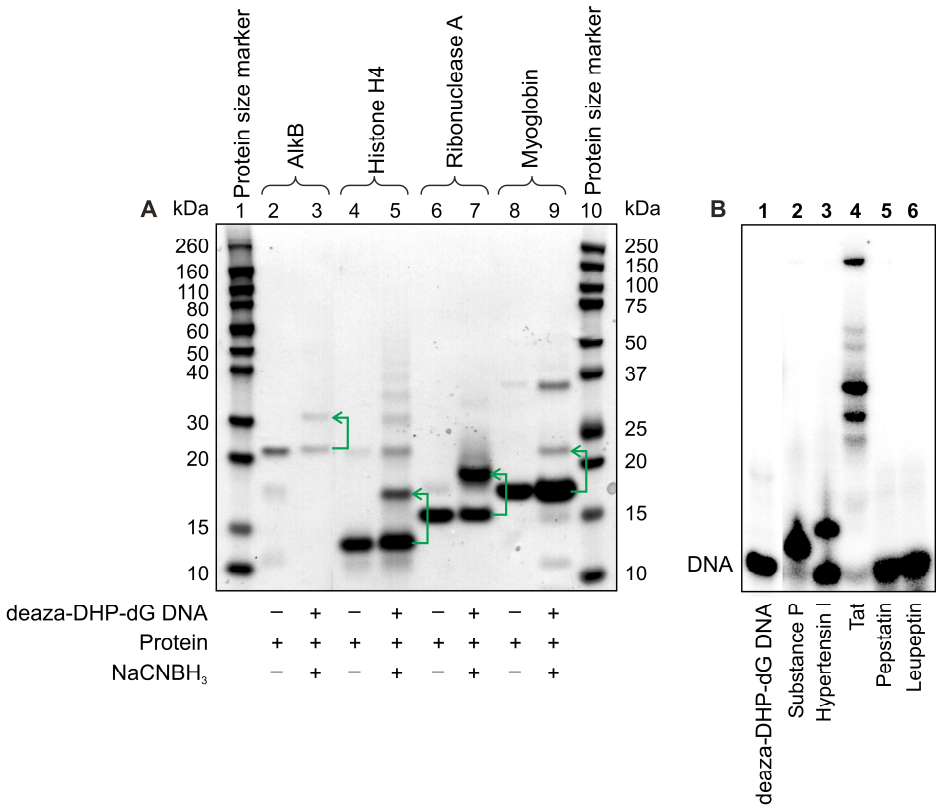
To gain insight into the identities of the amino acid residues participating in DPC formation, the cross-linking reactions were carried out with peptides of differing amino acid composition. High abundance DPC bands were observed for Tat and Substance P, which are rich in lysine and arginine residues (**Figure 3-4**B, lanes 2 and 4) suggesting that amino groups of Lys and Arg may be involved in crosslinking to DNA. In contrast, no covalent conjugates were detected for pepstatin, which has no Lys or Arg residues (**Figure 3-4**B, lane 5). A single DPC band was observed for hypertensin I, which has only one Arg residue and no Lys (**Figure 3-4**B, lane 3). Taken together, these results suggest that amino side chains of Lys and Arg within proteins can participate in reductive amination reactions with deaza-DHP-dG containing DNA.

Table 3-1 Proteins and peptides used to generate DNA-protein cross-links using the reductive amination reactions with deaza-DHP-dG containing DNA.[ǂ]

Protein/peptide	*Molecular weight/kDa*	*No. of Lys per monomer*	*No. of Arg per monomer*	*% Yield of DPC*
Histones	11-15	14	9	99
Substance P	1.3	1	1	94
Tat	1.8	2	6	92
Ribonuclease A	14.7	8	4	89
Histone H4	11.5	11	14	89
GAPDH	38	26	11	74
Aprotinin	6.5	4	7	73
NEIL1	44.5	23	42	64
Apomyoglobin	17.3	19	2	63
Insulin	5.8	2	5	63
AlkB	22.9	8	13	59
RNase	14.7	8	4	56
Hypertensin I	0.9	0	1	43
Pepsin	34.6	11	4	25
AGT	21.9	12	6	20
Myoglobin	17.6	19	2	19
T4 PNK	132	29	19	18
Carboxypeptidase A	34	18	17	17
Proteinase K	28.9	14	17	9
Trypsin	23.3	8	3	4
Pepstatin	0.7	0	0	0
Leupeptin	0.5	0	1	<1

[ǂ] The reactions were carried out under optimized conditions using a 20-fold excess of protein/peptide with respect to the starting oligodeoxynucleotide.

Figure 3-4 Influence of protein identity on DPC yield for reductive amination. (A) SDS-PAGE analysis of DPCs visualized *via* protein staining. Lane 1 and 10: protein size markers; Lanes 2–9: DPC formation between aldehyde-containing DNA and AlkB, histone H4, ribonuclease A, and myoglobin. Lanes 2, 4, 6, and 8 correspond to reactions conducted in the absence of DNA and $NaCNBH_3$ (negative controls). The cross-linking reactions were conducted using the 2:1 protein:DNA molar ratio. (B) Denaturing PAGE analysis of DPCs prepared using a set of different peptides that were visualized *via* ^{32}P-endlabeling of DNA. Lane 1: Aldehyde containing oligonucleotide in the presence of the reducing agent (negative control), Lanes 2–6: DPC formation between aldehyde-containing DNA and Substance P, Hypertensin I, Tat, pepstatin and leupeptin.



3.3.5 Mass Spectrometry Characterization of DNA-Protein Cross-Links

A mass spectrometry-based approach was employed to further characterize the structures of DNA-protein conjugates created by reductive amination and to identify the amino acids participating in reactions. The DNA components of DPCs to selected proteins (AlkB, RNase A, histone H4 and myoglobin) were digested with PDE I, PDE II, DNAse, and alkaline phosphatase. The resulting protein-nucleoside conjugates were cleaved with trypsin, and the peptides were analyzed by nano HPLC-ESI$^+$-MS/MS on an Orbitrap Velos mass spectrometer. Tryptic peptides containing cross-links to deaza-dG were identified, and the cross-linking sites were determined by MS/MS sequencing (see an example in **Figure 3-5**). The mass spectral data were processed using Thermo Proteome Discoverer 1.3 (ThermoScientific, San Jose, CA) to identify the cross-linking sites. We found that for all four proteins examined, multiple lysine and arginine residues were engaged in the cross-linking reaction to aldehyde-containing DNA under reducing conditions (**Tables 3-2** and **3-3**).

Figure 3-5 Mass spectrometry chracaterization of AlkB-DNA cross-links. MS/MS spectra of AlkB tryptic peptides NDPLK* (*m/z* 439.7, doubly charged) and R*NDPLK (*m/z* 557.8, doubly charged), where K* and R* contain an ethyl cross-link to 7-deaza-2'-deoxyguanosine and 7-deaza-2'-deoxyguanosine 5'-monophosphate, respectively.

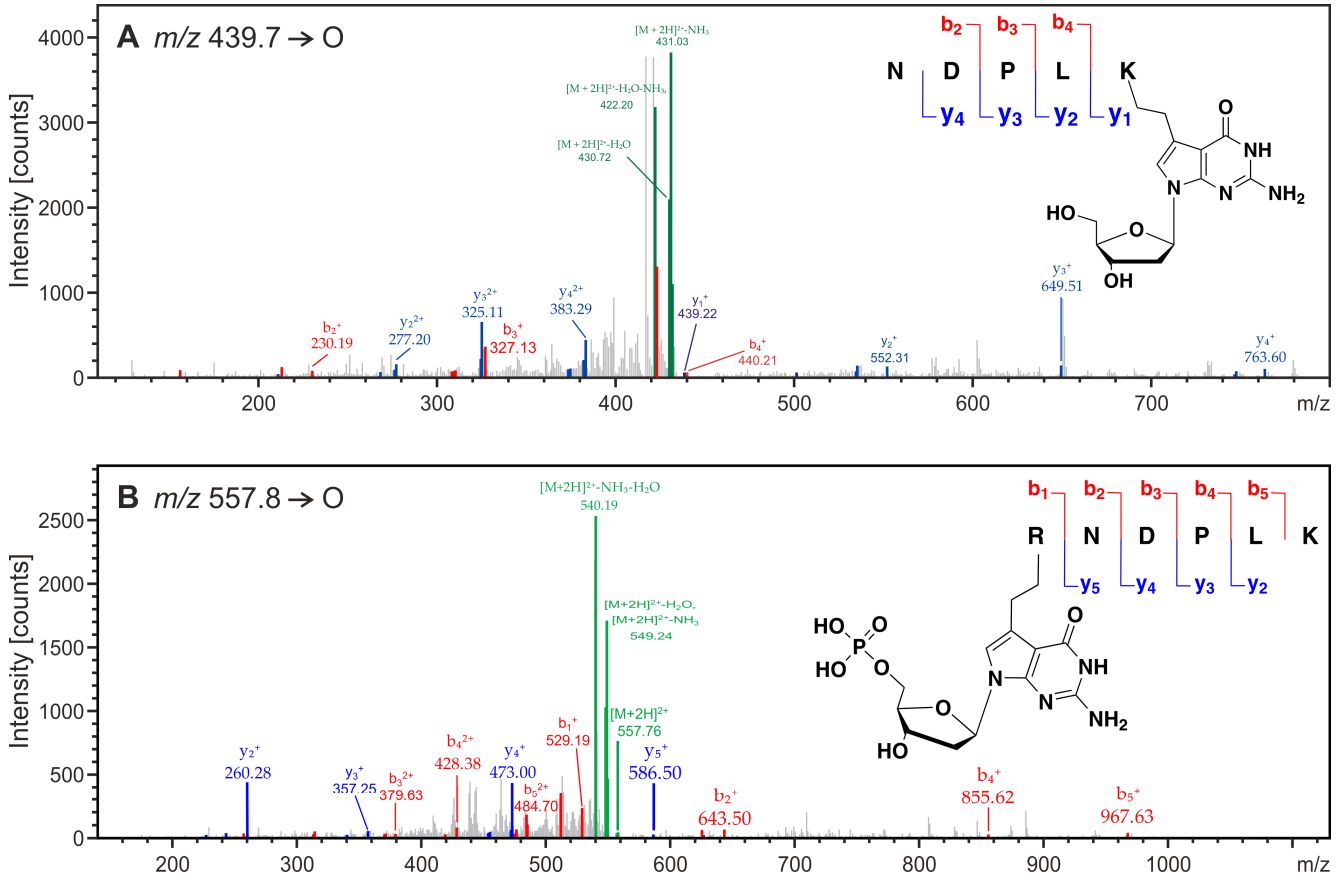


Table 3-2 Sites of reductive amination-mediated cross-linking between deaza-DHP-dG containing DNA and recombinant AlkB protein as identified by nano HPLC-ESI$^+$-MS/MS of tryptic digests.

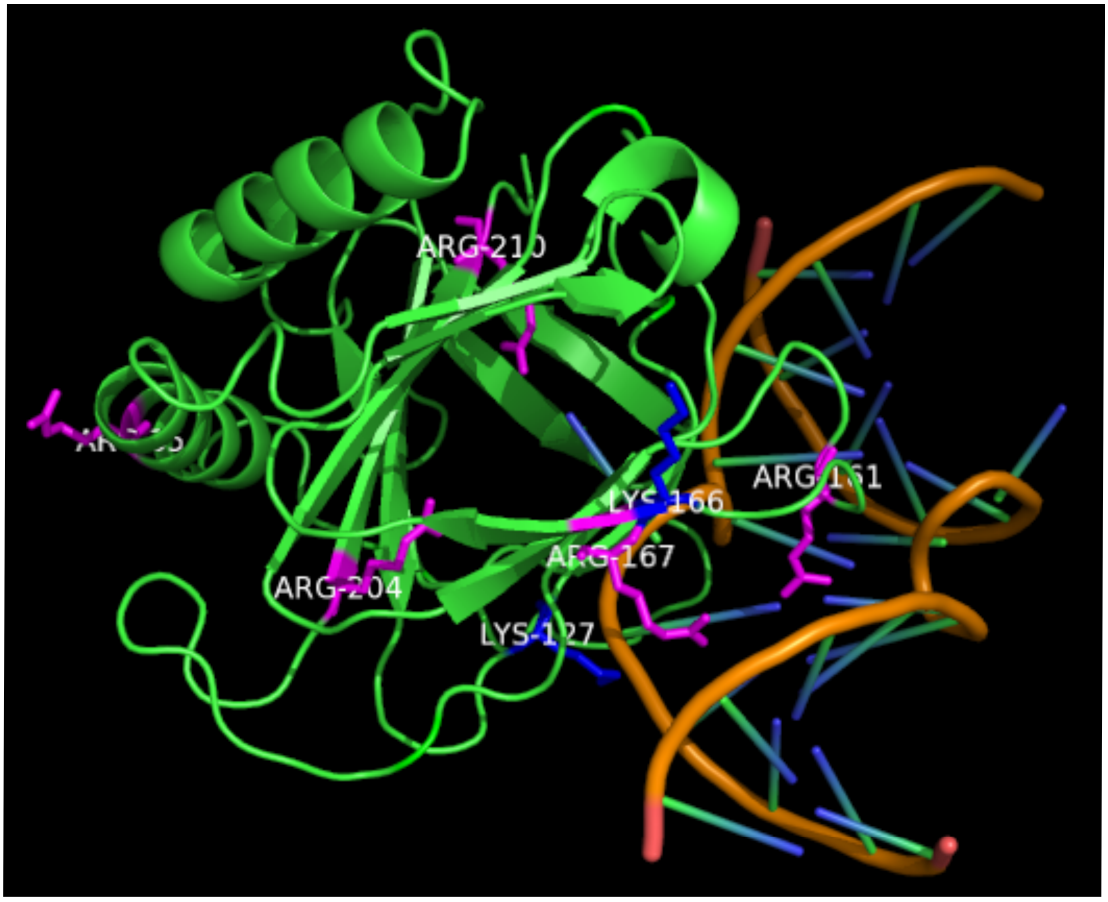
Amino acid positions	*Amino acid sequence*	*Site of modification*	*Location within the protein*
25–35	FAFNAAEQLI***R***	R35	-
122–127	CVPGA***K***	K127	DNA binding groove
122–134	CVPGA***K***LSLHQDK	K127	DNA binding groove
161–166	***R***NDPLK	R161	DNA binding groove
162–166	NDPL***K***	K166	DNA binding groove
162–167	NDPL***K***R	K166	DNA binding groove
162–183	***R***LLLEHGDVVVWGGESR	R167	DNA binding groove
194–204	VGVHPLTTDC***R***	R204	active site
194–210	VGVHPLTTDC***R***YNLTFR	R204	active site
194–210	VGVHPLTTDCRYNLTF***R***	R210	active site
205–215	YNLTF***R***QAGKK	R210	active site

Table 3-3 Sites of reductive amination-mediated cross-linking between deaza-DHP-dG containing DNA and proteins as identified by nano HPLC-ESI$^+$-MS/MS of tryptic digests.

Amino acid positions	*Amino acid sequence*	*Site of modification*
A. Histone H4		
1–6	MSG***R***GK	R4
1–9	MSG***R***GKGGK	R4
10–17	GLG***K***GGAK	K13
10–18	GLG***K***GGAKR	K13
21–24	KVL***R***	R24
21–24	***K***VLR	K21
46–56	***R***ISGLIYEETR	R46
B. Myoglobin		
46–56	FKHL***K***TEAEMK	K50
134–145	ALELF***R***NDIAAK	R139
134–145	ALELFRNDIAA***K***	K145
140–145	NDIAA***K***	K145
140–147	NDIAA***K***YK	K145
C. RNase I		
1–7	***K***ETAAAK	K1
40–61	C***K***PVNTFVHESLADVQAVCSQK	K41
40–61	CKPVNTFVHESLADVQAVCSQ***K***	K61

Nano HPLC-ESI$^+$-MS/MS analysis of AlkB-DNA conjugates was repeated several times, yielding reproducible results. Peptide sequencing by HPLC-ESI$^+$-MS/MS has revealed that two lysine residues (K127, K166) and five arginine residues of AlkB (R35, R161, R167, R204, R210) can participate in the AlkB-DNA cross-linking *via* reductive amination (**Figure 3-6** and **Table 3-2**). Examination of published crystal structures suggests that K134, R204, and R210 are located in the active site of AlkB, while K127, K166, and R167 reside in the DNA binding groove of the protein.[379,380] These results suggest that specific AlkB-DNA binding facilitates covalent DPC formation. However, some of the residues participating in cross-linking (e.g. R35) are located outside of the DNA-binding domain. This can be explained by partial denaturation of the protein under strongly reducing conditions used in our experiments.

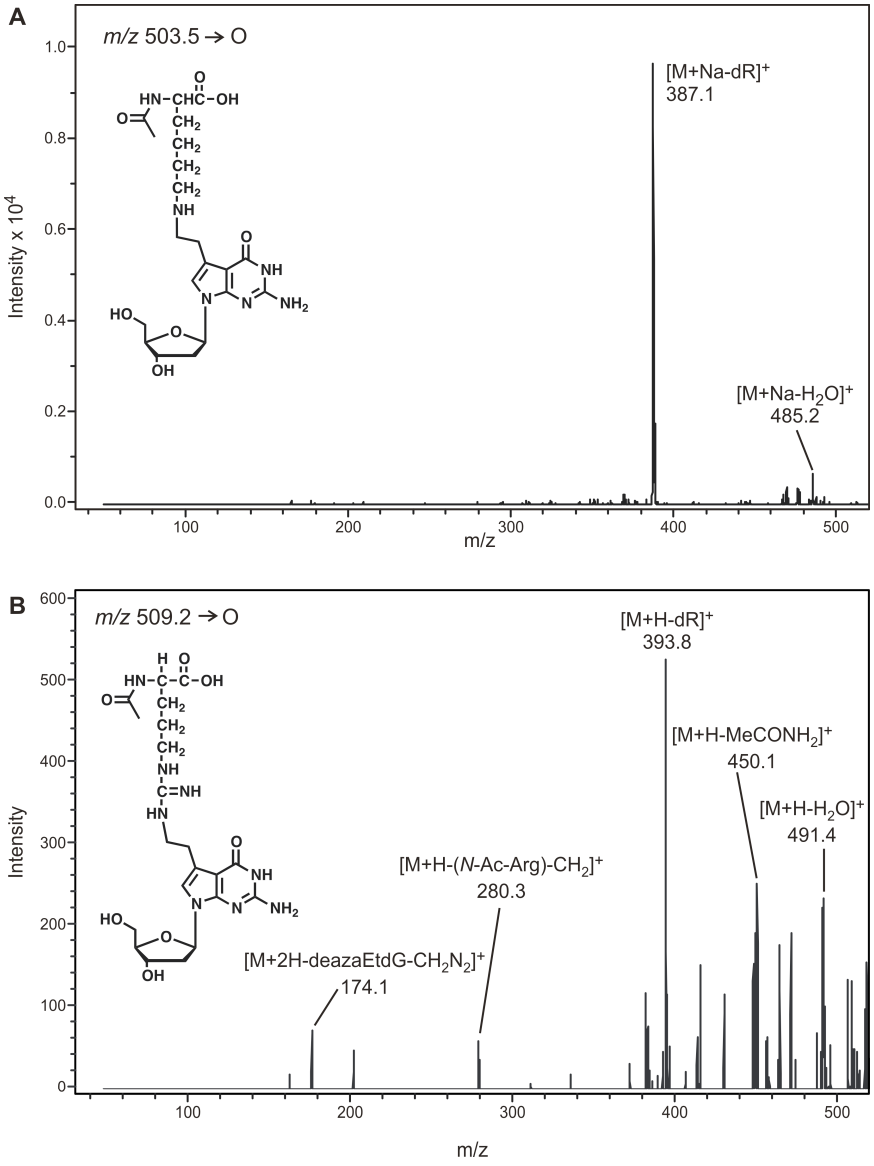
Figure 3-6 Crystal structure of AlkB protein bound to double stranded DNA (PDB ID: 3BI3) showing the amino acids participating in DNA-protein cross-linking.



3.3.6 Synthesis and Structural Characterization of 7-Deaza-7-(2-(N-acetyl lysine)ethan-1-yl)-2'-deoxyguanosine and 7-Deaza-7-(2-(N-acetyl arginine)ethan-1-yl)-2'-deoxyguanosine Conjugates

To confirm the exact chemical structure of the DNA-protein cross-links generated *via* reductive amination (**Scheme 3-2**), synthetic 7-deaza-7-(2,3-dihydroxypropan-1-yl)-2'-deoxyguanosine (deaza-DHP-dG) was allowed to react with *N*-acetyl protected Lys and Arg, and the resulting nucleoside-amino acid conjugates were isolated by HPLC and characterized by mass spectrometry. In the case of *N*-acetyllysine, the major conjugation product was observed at *m/z* 503.2, corresponding to the $[M+Na]^+$ ions of 7-deaza-7-(2-(*N*-acetyllysine)ethan-1-yl)-2'-deoxyguanosine. MS/MS fragmentation pathway of this conjugate (**Figure 3-7**A) was dominated by the product ions at *m/z* 387.1 and *m/z* 485.2, which correspond to the loss of deoxyribose, and a water molecule, respectively (**Figure 3-7**A). Similar results were observed for *N*-acetylarginine. MS^2 fragmentation of protonated 7-deaza-7-(2-(*N*-acetylarginine)ethan-1-yl)-2'-deoxy-guanosine, $[M+H]^+$, has revealed the loss of water ($[M+H-H_2O]^+$), acetylamine ($[M+H-MeCONH_2]^+$) and deoxyribose ($[M+H-dR]^+$, **Figure 3-7**B). Further, fragmentation at the guanidinium moiety of Arg (*m/z* 174.1) and ethylene linker of deaza-dG (*m/z* 280.3) were also observed (**Figure 3-7**B). Taken together, these results are consistent with the cross-linking mechanism shown in **Scheme 3-2**, e.g. Schiff base formation between the lysine amino side chain and the aldehyde functionality within oxidized deaza-DHP-dG, followed by imine reduction to generate a stable amino linkage.

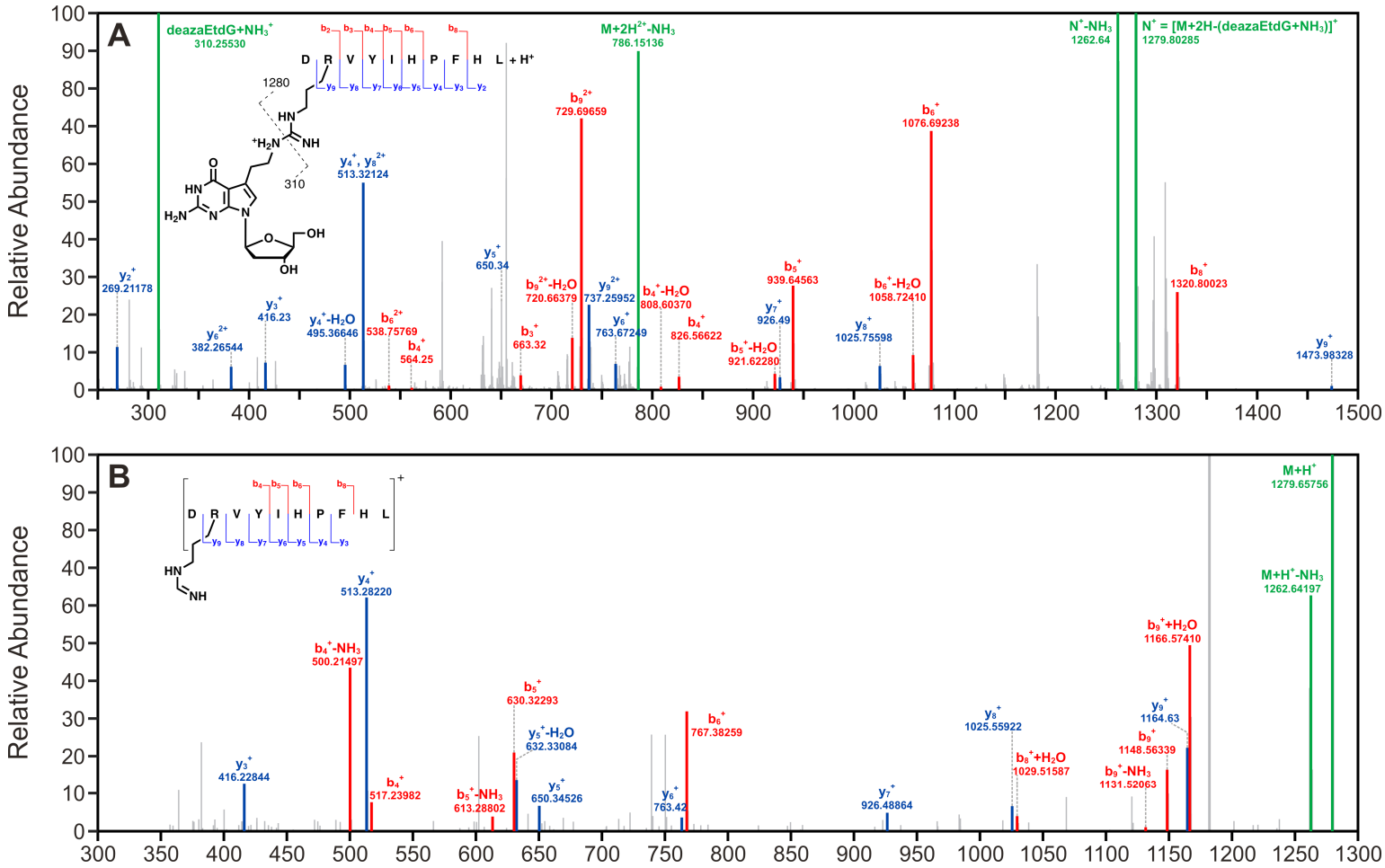
Figure 3-7 HPLC-ESI$^+$-MS/MS characterization of 7-deaza-7-(2-(*N*-acetyllysine)ethan-1-yl)-2'-deoxyguanosine (A) and 7-deaza-7-(2-(*N*-acetylarginine)ethan-1-yl)-2'-deoxyguanosine (B).



3.3.7 Synthesis and Characterization of Nucleoside-Peptide Conjugates

To further confirm the chemical structure and site of cross-linking of DPCs formed by reductive amination strategy, nucleoside-peptide conjugates of 7-deaza-7-(2,3-dihydroxypropan-1-yl)-2'-deoxyguanosine (deaza-DHP-dG) to angiotensin I and substance P were characterized by mass spectrometry. MS^2 spectrum of angiotensin I conjugate has revealed that the site of cross-linking is side chain amino group of arginine (**Figure 3-8**A). The MS^2 spectrum has further revealed that 7-deaza-7-(1-aminoethan-2-yl)-2'-deoxyguanosine is lost upon fragmentation of the doubly charged parent ion, $[M+2H]^+$ = 794.9 *m/z*. The resulting singly charged daughter ion, $[M+H]^+$ = 1279.8 *m/z* was further fragmented to observe additional b- and y-ions of the modified peptide that confirmed the site of cross-linking (**Figure 3-8**B).

Figure 3-8 NanoLC-nanospary-MS^2 (A) and MS^3 (B) spectra of the peptide, Angiotensin I (DRVYIHPFHL) cross-linked to 7-deaza-7-(ethan-1-yl)-2'-deoxyguanosine.



3.4 DISCUSSION

Dynamic DNA-protein interactions are crucial for many cellular functions including chromatin packaging, cell division, DNA replication, gene expression, DNA damage response, and DNA repair.[4,5,381] Proteins reversibly interact with DNA by a combination of electrostatic forces, hydrogen bonding, and stacking interactions, and their ability to dissociate from DNA is critical for their cellular functions. However, exposure to common antitumor drugs, environmental toxins, transition metals, UV light, ionizing radiation, and free radical-generating systems can result in proteins becoming covalently trapped on DNA.[4,5] This generates super bulky, highly heterogeneous DNA-protein cross-links (DPCs) that can block DNA and RNA polymerases, causing toxicity and/or mutations in affected cells.[36,38,85,158,192,308,315,316]

Our previous mass spectrometry-based proteomics studies have revealed that covalent DNA-protein cross-links (DPCs) involving the *N7* position of guanine are readily formed in human cells treated with clinically relevant concentrations of chemotherapeutic drugs (e.g., platinum compounds and nitrogen mustards)[158,294] and metabolically activated carcinogens (e.g., 1,2,3,4-diepoxybutane).[85] Additionally, covalent DPCs have been shown to accumulate in an age-dependent fashion in the brain and heart tissues, probably a result of exposure to endogenous reactive oxygen species, lipid peroxidation products, and transition metals.[36] If not repaired, DPCs may contribute to the development of cancer, cardiovascular disease, and age-related neurodegeneration.[34,36-38,164,382,383]

Conflicting data exist in the literature regarding the mechanisms of cellular repair of DPC lesions. Reardon *et al.* examined the ability of reconstituted bacterial and mammalian excision nuclease systems to recognize the ring-open T4 pyrimidine DNA glycosylase-DNA cross-links.[169] While the excision of DNA-protein conjugates was not detected, DPCs to short polypeptides were recognized and cleaved by mammalian protein extracts, leading to the hypothesis that DPCs are proteolytically degraded prior to their repair *via* the NER pathway.[169] Similar conclusions were drawn by the Lloyd group when using a bacterial *UvrABC* system[324] and by Baker *et al.* who examined model DPCs containing bacterial DNA methyltransferase-DPC (37 kDa) attached to the *C6* position of cytosine.[331] Quievryn *et al.* observed reduced rates of repair of formaldehyde-induced DPCs in the presence of a protease inhibitor.[297] In contrast, Nakano *et al.* reported that cytosolic ATP-dependent proteases are not involved in DPC removal.[326,327] These authors proposed that homologous recombination repair is responsible for removing the majority of DPCs generated *via* oxanine, while only DPCs involving small proteins (<12 kDa in bacteria and <8 – 10 kDa in mammalian cells) are repaired by NER.[326,327]

Structural factors such as protein size, identity and lesion structures (e.g. major or minor groove of DNA) are also likely to affect the DPC lesion's ability to be bypassed by DNA and RNA polymerases.[4,260] For example, *E. coli* polymerase I and HIV-1 reverse transcriptase were completely blocked by *cis*-[diamminedichloroplatinum(II)] (cisplatin) cross-linked to histone H1.[219] The A family human polymerase ν was blocked by DNA-peptide cross-links located in the minor groove *via* N^2-dG.[336] In contrast, chemically and structurally similar lesions located in the major groove of DNA *via* N^6-dA were

efficiently and accurately bypassed by both human polymerase ν and *E. coli* polymerase I.[336,337]

It is likely that the discrepancies between the mechanisms of DPC repair and bypass reported by different groups reflect structural differences between the model DPCs examined. Indeed, these previous studies have employed DNA-conjugates of diverse structure and size, including those where the protein was directly attached to ring open abasic sites in DNA.[5,325,327] This underlines the need to re-examine the replication and repair of DPC-containing DNA using substrates resembling the lesions formed in cells. The most common site of DNA involved in DPC formation is the *N7* of guanine.[5,158,161] However, to our knowledge, no methods have been previously reported in the literature to generate *N7* guanine conjugated DPCs.

As mentioned before, currently available synthetic strategies to generate site-specific DPCs are limited to several main strategies. Lloyd *et al.*[324] and Sancar *et al.*[330] employed a semi-enzymatic approach to trap T4 pyrimidine dimer glycosylase/AP lyase (T4-pdg) on abasic sites of DNA in the presence of sodium borohydride. A similar methodology has been used to attach oxoguanine glycosylase (Ogg) to DNA strands containing 8-oxo-dG. DNA methyltransferase (Dnmt) has been trapped on DNA containing 5-fluorocytosine.[331] Another approach involves the use of oxanine (Oxa) in DNA that spontaneously reacts with amino groups of proteins to give a pyrimidine ring-open structure,[164] and this strategy, which is specific to nitric oxide-induced Oxa lesion, requires a large excess of the protein (425 to 3000-fold), as well as long incubation times (up to 48 h).[164,326] Finally, Schiff base formation between acrolein-induced γ-HOPdG

adducts and lysine residues of proteins and peptides can be stabilized in the presence of $NaCNBH_3$.[198,246,336] Either the N^2 guanine or N^6 adenine aldehyde functionality derived from acrolein-induced 3-(2'-deoxyribos-1'-yl)-5,6,7,8-tetrahydro-8-hydroxypyrimido [1,2*a*]purin-10(3*H*)-one (γ-HOPdG) can be reacted with proteins and peptides to produce a Schiff base, which was subsequently reduced to the corresponding amine with sodium cyanoborohydride.[246,336] To our knowledge, no synthetic methodologies are available to generate *N7*-guanine DPCs such as those formed *in vivo* upon exposure to environmental carcinogens and antitumor agents.[85,158]

In the present study, a post-synthetic reductive amination strategy was employed to create hydrolytically stable structural mimics of *N7* guanine conjugated DPCs by reductive amination reactions between Lys and Arg side chains of proteins and acetaldehyde functionalities of modified 7-deazaguanine residues of DNA (**Scheme 3-2**). The main advantage of this approach is that it generates sequence specific DPC lesions structurally analogous to the lesions formed *in vivo*. Reductive amination methodology is highly versatile, as it can be used to generate DPCs to most proteins and peptides containing Lys and/or Arg residues (**Table 3-1**). The resulting structurally defined and hydrolytically stable DPCs can be used to study the biological fate of DPCs *in vitro* to better understand the effects of these lesions in cells. Furthermore, experimental methods are being developed in our laboratory to incorporate these substrates into plasmid DNA and study their repair in cells to identify the mechanisms responsible for the removal of these lesions *in vivo*.

Model DPC lesions generated in this work resemble the adducts induced by antitumor nitrogen mustards (**Scheme 3-3**).[158,161] We have previously identified 39 proteins that form covalent DPCs in human fibrosarcoma (HT1080) cells treated with mechlorethamine.[159] However, it is not known to what extent DPC formation contributes to toxicity of nitrogen mustards in cancer cells. The availability of hydrolytically stable model DPC substrates will, for the first time, enable structural and biological evaluation of these super-bulky lesions. Based on our recent studies with DNA-reactive protein reagents that specifically induce DPCs in cells,[304] we hypothesize that spontaneous and xenobiotic-induced DPCs, if not repaired, will compromise the efficiency and the accuracy of DNA replication and are responsible for a major portion of the toxicity and mutagenicity induced by *bis*-alkylating agents, UV light, reactive oxygen species, and γ-radiation.

The model DPC substrates created by reductive amination (**Scheme 3-2**) are site-specific with respect to DNA, but may involve multiple possible cross-linking sites within the protein (**Figure 3-6**, **Tables 3-2** and **3-3**). This limited specificity with respect to the protein side chains should not affect the ability to employ these model conjugates in biological studies since the "real" DPC lesions formed in cells are also heterogeneous in nature. However, it may not be practical to use this approach to generate DNA-protein conjugates for structural studies by NMR or X-ray crystallography, especially in the case of proteins that contain multiple basic residues available for reaction with DNA. Other types of conjugations that employ bioorthogonal reactive groups in each biomolecule (protein and DNA) may be more appropriate for this purpose. To this end, we have

employed a copper-catalyzed [3+2] Huisgen cycloaddition (click reaction) between azide-functionalized proteins[352] and alkyne-containing DNA to generate site specific DPC conjugates (Chapter 5).

4 REPLICATION BYPASS OF MODEL DNA-PROTEIN CROSS-LINKS CONJUGATED TO THE 7-DEAZAGUANINE OF DNA

4.1 INTRODUCTION

DNA-protein cross-linking in drug-treated cells appears to be as prominent as the formation of "traditional" DNA adducts. Isotope dilution HPLC-ESI-MS/MS studies of DNA hydrolysates from human fibrosarcoma (HT1080) cells treated with 0–100 μM mechlorethamine for 3 h have detected 1–5 DPC lesions per 10^6 nucleotides, depending on drug concentration.[159] Our more recent studies have revealed significant numbers of DPC lesions in peripheral blood lymphocyte DNA from patients undergoing treatment with cyclophosphamide, representing a possible novel mechanism of target and off-target toxicity of this useful drug (Teshome Gherezghiher and Natalia Tretyakova, unpublished data).

To our knowledge, no reports available in literature on replication bypass of *N7*-guanine adducts, probably because of their transient nature. Recently, we developed a post-synthetic reductive amination strategy to generate hydrolytically stable, site-specific DPCs that are structurally analogous to *N7*-G adducts formed *in vivo* (Chapter 3).[354] In the present work, primer extension experiments with human lesion bypass DNA polymerases were conducted to examine replication bypass of model DPCs of varying size. *In vitro* DNA replication assays have revealed that large protein (AlkB, histone H4

and 6×His-eGFP) and polypeptide (23-mer peptide) cross-links completely block human translesion synthesis polymerases η and κ, while smaller 10-mer peptide cross-links were bypassed. Steady-state kinetics experiments provided evidence for the error-free bypass of the 10-mer peptide cross-link by hPol κ. hPol η incorporated primarily the correct base, dC opposite the adduct, while the incorporation of dT was much less efficient. Our results suggest the requirement of proteolytic degradation of proteins cross-linked to DNA prior to replication and the ability of human TLS polymerases to catalyze accurate replication in the presence of proteolytically cleaved DPC lesions.

4.2 MATERIALS AND METHODS

4.2.1 Materials

Fluorescein-dT phosphoramidite, protected 2'-deoxyribonucleoside-3'-phosphoramidites (dA-CE, Ac-dC-CE, dmf-dG-CE, dT-CE), Ac-dC-CPG ABI, dmf-dG-CPG ABI columns, and all other reagents required for automated DNA synthesis were purchased from Glen Research (Sterling, VA). Recombinant human polymerases hPol η, κ and ι were expressed and purified as described elsewhere.[384-386] Recombinant AlkB protein was a generous donation by Dr. Chuan He (Department of Chemistry, University of Chicago). Expression of 6×His-eGFP protein[387] and synthesis of 23-mer and 10-mer peptides[388] were carried out according to previously reported protocols. T4 polynucleotide kinase (T4-PNK) and recombinant histone H4 were obtained from New England Biolabs (Beverly, MA), while γ-^{32}P ATP was purchased from Perkin-Elmer Life Sciences (Boston, MA). 40% 19:1 Acrylamide/bis solution and micro bio-spin 6 columns were purchased from Bio-Rad (Hercules, CA). The unlabeled dNTPs were obtained from Omega Bio-Tek (Norcross, GA). Size exclusion columns (NAP-5 and Microspin G25) and Sep-Pak C18 SPE cartridges were purchased from GE Healthcare (Pittsburg, PA) and Waters (Milford, MA), respectively. All other chemicals and solvents were purchased from Sigma-Aldrich (Milwaukee, WI) and were of the highest grade available.

4.2.2 Synthesis and Characterization of Oligodeoxynucleotides

Synthetic DNA strands containing site-specific 7-deaza-7-(2,3-dihydroxypropan-1-yl)-2'-deoxyguanosine (5'-G TCA CTG GTA ***deaza-DHP-dG***CA AGC ATT G-3' and 5'-GAA AGA AG***deaza-DHP-dG*** ACA GAA GAG GGT ACC ATC ATA GAG TCA GTG-3') were prepared as previously described.[377] Primers internally labeled with 5-[*N*-((fluoresceinyl)-aminohexyl)-3-acrylimido]-2'-deoxyuridine (5'-CAA ***FAMdT***GC TTG-3' and 5'-CTA TGA ***FAMdT***GG TAC C-3') and unmodified DNA strands (5'-G TCA CTG GTA GCA AGC ATT G-3', 5'-GAA AGA AGG ACA GAA GAG GGT ACC ATC ATA GAG TCA GTG-3', and 5'-CAA TGC TTG-3') were synthesized by standard solid phase synthesis using an ABI 394 DNA synthesizer (Applied Biosystems, CA). All oligodeoxynucleotides were purified by semi-preparative HPLC, desalted by Illustra NAP-5 columns, characterized by HPLC-ESI$^-$-MS, and quantified by UV spectrophotometry.

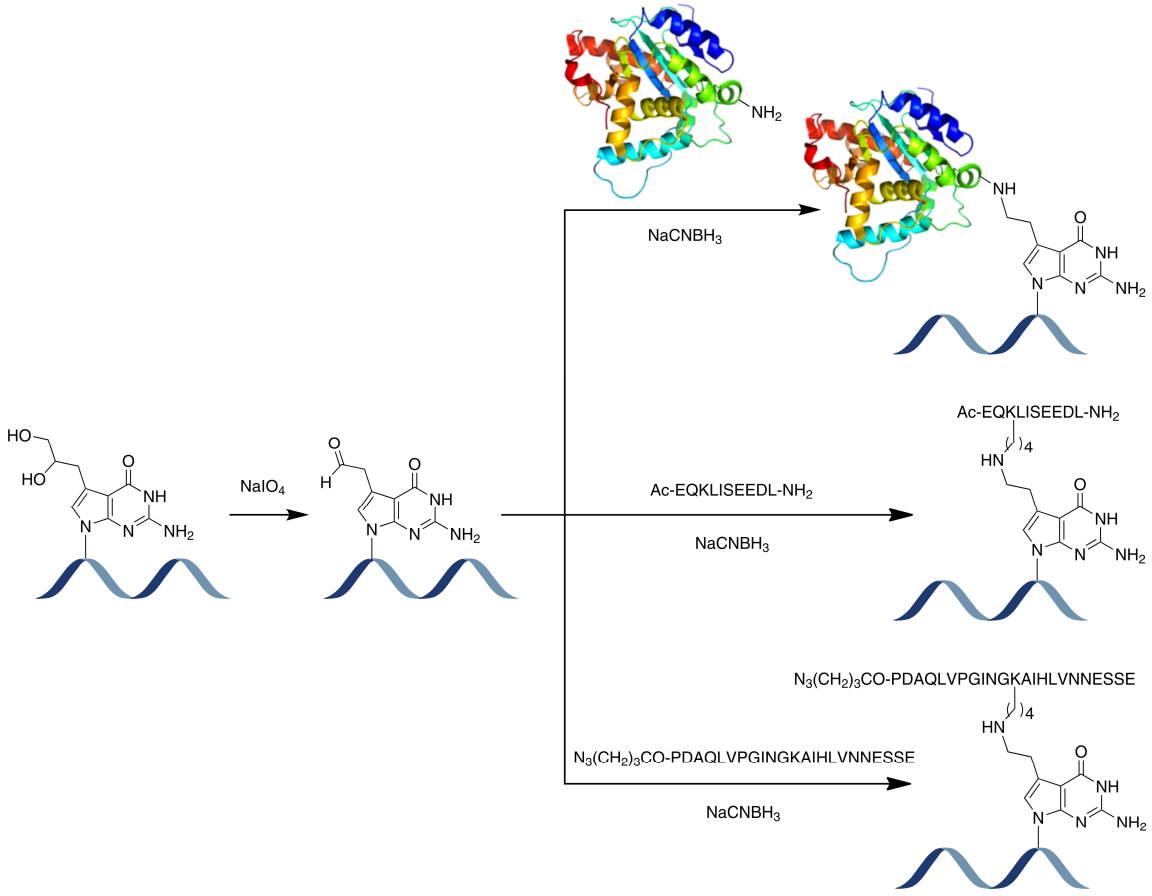
4.2.3 Synthesis and Characterization of DNA-Protein and DNA-Peptide Cross-Links

Site-specific DNA-protein and DNA-peptide cross-links were generated by a post-synthetic reductive amination strategy developed in our laboratory.[354] Briefly, synthetic 20-mer (5'-G TCA CTG GTA ***deaza-DHP-dG***CA AGC ATT G-3') or 39-mer (5'-GAA AGA AG***deaza-DHP-dG*** ACA GAA GAG GGT ACC ATC ATA GAG TCA GTG-3', 1 nmol in 36 μL water) were oxidized in the presence of 50 mM $NaIO_4$ (70 μL)

in 15 mM sodium phosphate buffer (pH 5.4, 70 μL) for 6 h at 4 °C in the dark to unmask the aldehyde moiety on deaza-DHP-dG (**Scheme 4-1**). Excess $NaIO_4$ was quenched with 55 mM Na_2SO_3 (70 μL), and the resulting aldehyde-functionalized DNA was incubated with the protein of interest (AlkB, Histone H4 and 6×His-eGFP, 2–10-fold molar excess) in the presence of 25 mM $NaCNBH_3$ at 37 °C overnight to generate DNA-protein cross-links (**Scheme 4-1**). Site-specific DPCs were isolated using 20% (w/v) denaturing polyacrylamide gels containing 7 M urea followed by gel-elution.

DNA-peptide conjugates were prepared analogously (**Scheme 4-1**), except that 10–100-fold molar excess of the peptide was used, and the pH was adjusted to 7 prior to peptide addition. DNA-peptide conjugates were isolated using 15% (w/v) denaturing polyacrylamide gels containing 7 M urea followed by gel-elution. DNA-protein and DNA-peptide cross-links were desalted by micro bio-spin 6 columns and Sep-Pak C18 SPE, respectively. The purified conjugates were characterized by HPLC-ESI^+-MS/MS as described elsewhere.[354] An aliquot of the purified sample was radiolabeled with γ-^{32}P, and resolved on a 20% (w/v) denaturing polyacrylamide gel containing 7 M urea, followed by visualization using a Typhoon FLA 7000 phosphorimager (GE Healthcare, Pittsburgh, PA) to assess the purity of the conjugates. Depending on the purity of the cross-links obtained, additional gel purifications were carried out to obtain >98% pure conjugates.

Scheme 4-1 Synthesis of DNA-protein and DNA-peptide cross-links by post-synthetic reductive amination.[354]



4.2.4 Preparation of Primer-Template Duplexes

For primer extension assays, the fluorescein labeled primers (5'-CAA ***FAMdT***GC TTG-3' or 5'-CTA TGA ***FAMdT***GG TAC C-3', 50 pmol) were mixed with 2 eq. of template strands containing dG, deaza-DHP-dG or DPC in 10 mM Tris buffer (pH 8) containing 50 mM NaCl. The mixtures were heated at 95 °C for 10 min and allowed to cool slowly overnight to afford the corresponding primer-template duplexes.

For steady-state experiments to determine the kinetics of single nucleotide insertion opposite the cross-links, 9-mer DNA primer (5'-CAA TGC TTG-3', 1 nmol) was radiolabeled with T4 PNK (20 U) and γ-^{32}P ATP (25 μC*i*) in the presence of T4 PNK reaction buffer (total volume = 20 μL) at 37 °C for 1 h. The mixture was heated at 65 °C for 10 min to inactivate the enzyme and passed through Illustra Microspin G25 columns (GE Healthcare, Pittsburgh, PA) to remove excess γ-^{32}P ATP. The 5'-^{32}P-labeled primers (50 pmol) were mixed with 2 eq. of template strands (5'-G TCA CTG GTA ***X***CA AGC ATT G-3') containing either dG or 10-mer peptide in 10 mM Tris buffer (pH 8) containing 50 mM NaCl and annealed as described above.

4.2.5 Primer Extension Assays

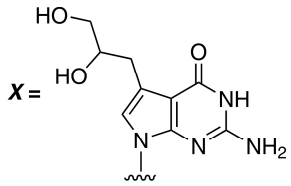
For standing start experiments, the 9-mer FAMdT primer-20-mer template duplexes (0.15 μM, **Scheme 4-2**B) were incubated with human DNA polymerase hPol η, κ or ι (0.30 μM) at room temperature in the presence of a buffered solution containing 50 mM Tris (pH 7.5), 50 mM NaCl, 5 mM DTT, 100 μg/μL BSA, 10% glycerol (v/v) and

5 mM $MgCl_2$. Primer extension reactions were initiated by adding 0.5 mM solution containing all four dNTPs. Aliquots of the reaction mixture (4 μL) were withdrawn at preselected time intervals (0, 5, 30, 60, 90 and 180 min) and quenched by the addition of 95% formamide (v/v) containing 10 mM EDTA (8 μL). Samples were loaded on to a 20% denaturing polyacrylamide gel and ran at 75 W for 3 h. The products were visualized using a Typhoon FLA 7000 (GE Healthcare, Pittsburgh, PA) in the fluorescence imaging mode. The running start experiments were conducted in a similar manner using 13-mer FAMdT primer-39-mer template DNA duplexes (**Scheme 4-2**C).

4.2.6 Single Nucleotide Incorporation Assays

To determine which nucleotides are incorporated opposite the DNA-peptide cross-links, ^{32}P-endlabeled primer-template duplexes containing either dG or 10-mer peptide conjugate (50 nM) at position ***X*** (**Scheme 4-2**D) were incubated with human TLS polymerases (50 nM hPol η and 150 nM hPol κ) in 50 mM Tris-HCl (pH 7.5) buffer containing 50 mM NaCl, 5 mM DTT, 5 mM $MgCl_2$, 100 μg/mL BSA and 10% glycerol (v/v) at room temperature. Reactions were initiated by the addition of individual dNTPs (100 μM) in a final volume of 20 μL. Aliquots (4 μL) were withdrawn at pre-selected time points, and the reaction was quenched by the addition of the quench solution containing 10 mM EDTA, 0.03% bromophenol blue (w/v) and 0.03% xylene cyanol (w/v) in 95% (v/v) formamide (8 μL). The extension products were resolved by 20% (w/v) denaturing PAGE containing 7 M urea and visualized using Typhoon FLA 7000 phosphorimager (GE Healthcare, Pittsburgh, PA).

Scheme 4-2 Sequences of DNA oligomers employed in *in vitro* replication experiments. (A) DNA strands used for DPC synthesis, (B) 9-mer FAMdT primer-20-mer template duplex used in standing start assays, (C) 13-mer FAMdT primer-39-mer template DNA duplex used in running start assays, (D) 9-mer primer-20-mer template duplex used for the single nucleotide incorporation and steady-state kinetic assays.



A. 5'-G TCA CTG GTA ***X***CA AGC ATT G-3' (20-mer deaza-DHP-dG)

5'-GAA AGA AG***X*** ACA GAA GAG GGT ACC ATC ATA GAG TCA GTG-3' (39-mer deaza-DHP-dG)

Y = dG
deaza-DHP-dG
$N_3(CH_2)_3CO$-EQKLISEEDL-NH_2 (10mer peptide)
$N_3(CH_2)_3CO$-PDAQLVPGINGKAIHLVNNESSE (23mer peptide)
AlkB
Histone H4
6xHis-eGFP

B. 5'-CAA ***FAMT***GC TTG -3' (9-mer FAMdT primer)
3'-GTT ACG AAC ***Y***AT GGT CAC TG-5' (20-mer template)

C. 5'-C TAT GA***FAMT*** GGT ACC -3' (13-mer FAMdT primer)
3'-GTG ACT GAG ATA CT A CCA TGG GAG AAG ACA ***Y***GA AGA AAG-5' (39-mer template)

Z = dG
Ac-EQKLISEEDL-NH_2 (10mer peptide)

D. ^{32}P-5'-CAA TGC TTG -3' (9-mer primer)
3'-GTT ACG AAC ***Z***AT GGT CAC TG-5' (20-mer template)

4.2.7 Steady-State Kinetic Analyses

Steady-state kinetics for incorporation of individual nucleotides opposite the unadducted dG and 10-mer peptide cross-link were investigated by performing single nucleotide incorporation assays with 0.5–50 nM hPol η or 3–55 nM hPol κ in the presence of increasing concentrations of individual dNTPs (0–500 μM) and reactions were quenched at preselected time points (0–60 min). The product bands were visualized using Typhoon FLA 7000 phosphorimager (GE Healthcare, Pittsburgh, PA), and quantified by volume analysis using the Image Quant TL 8.0 software (GE Healthcare, Pittsburgh, PA). The steady-state kinetic parameters were calculated by nonlinear regression analysis using one-site hyperbolic fits in Prism 4.0 (GraphPad Software, La Jolla, CA).

4.3 RESULTS

4.3.1 Synthesis and Characterization of DNA-Protein and DNA-Peptide Cross-Links

Model DNA-protein and DNA-peptide cross-links were prepared by covalent linkage of 7-(oxoethyl)-7-deazaguanine within DNA and side chain amino group of Lys or Arg on proteins and peptides under reducing conditions (**Scheme 4-1**).[354] Synthetic 20-mer (5'-G TCA CTG GTA ***deaza-DHP-dG***CA AGC ATT G-3') was previously used for the generation of DPCs in high yields,[354] while the 39-mer (5'-GAA AGA AG***deaza-DHP-dG*** ACA GAA GAG GGT ACC ATC ATA GAG TCA GTG-3') has been employed in replication studies of DNA interstrand cross-links.[389] Biological relevance of histones and DNA repair enzymes in DNA-protein cross-linking[86,158,159,294] warrants the use of histone H4 and *E. coli* repair protein, AlkB, as model proteins for *in vitro* replication studies. We recently reported the replication bypass of 6×His-eGFP protein and the peptides (23-mer and 10-mer) cross-linked to C5 of thymidine.[388] The same polypeptides were used in this study to assess the effect of cross-linking site within DNA on replication bypass of DPCs. Structurally modified DNA strands were isolated using denaturing gel electrophoresis, and the purity of the conjugates was tested by sequencing PAGE to ensure >98% purity. Synthetic DNA-protein and DNA-peptide cross-links were structurally characterized by nanoHPLC-nanspray-MS/MS.[354] In the resulting model cross-links, the Lys and Arg side chains of proteins are conjugated to the 7-deazaguanine

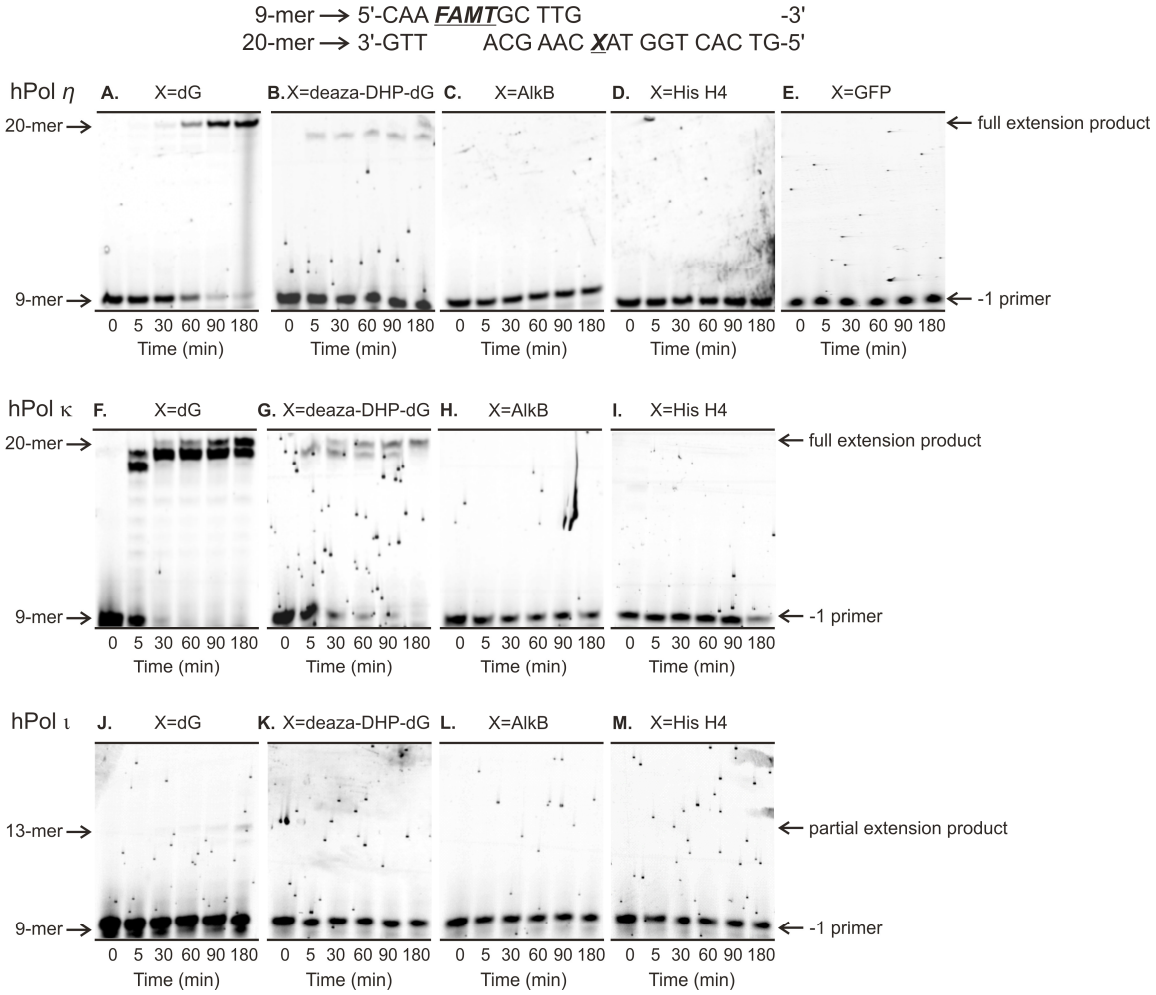
residues of DNA *via* a two carbon linker, creating a structure that is analogous to DPC lesions induced by antitumor nitrogen mustards.[158]

4.3.2 Replication Bypass of DNA-Protein Cross-Links

To examine the influence of model DNA-protein cross-links on DNA replication, template-primer complexes containing site-specific 7-deaza-G cross-links to histone H4 (11.5 kDa), AlkB (22.9 kDa), 6×His-eGFP (28.4 kDa) were subjected to primer extension in the presence of human TLS polymerases η, κ and ι (**Figures 4-1** and **4-2**). Templates containing unconjugated deaza-DHP-dG and native G served as negative controls, and primer extension assays were conducted under both standing start and running start conditions.

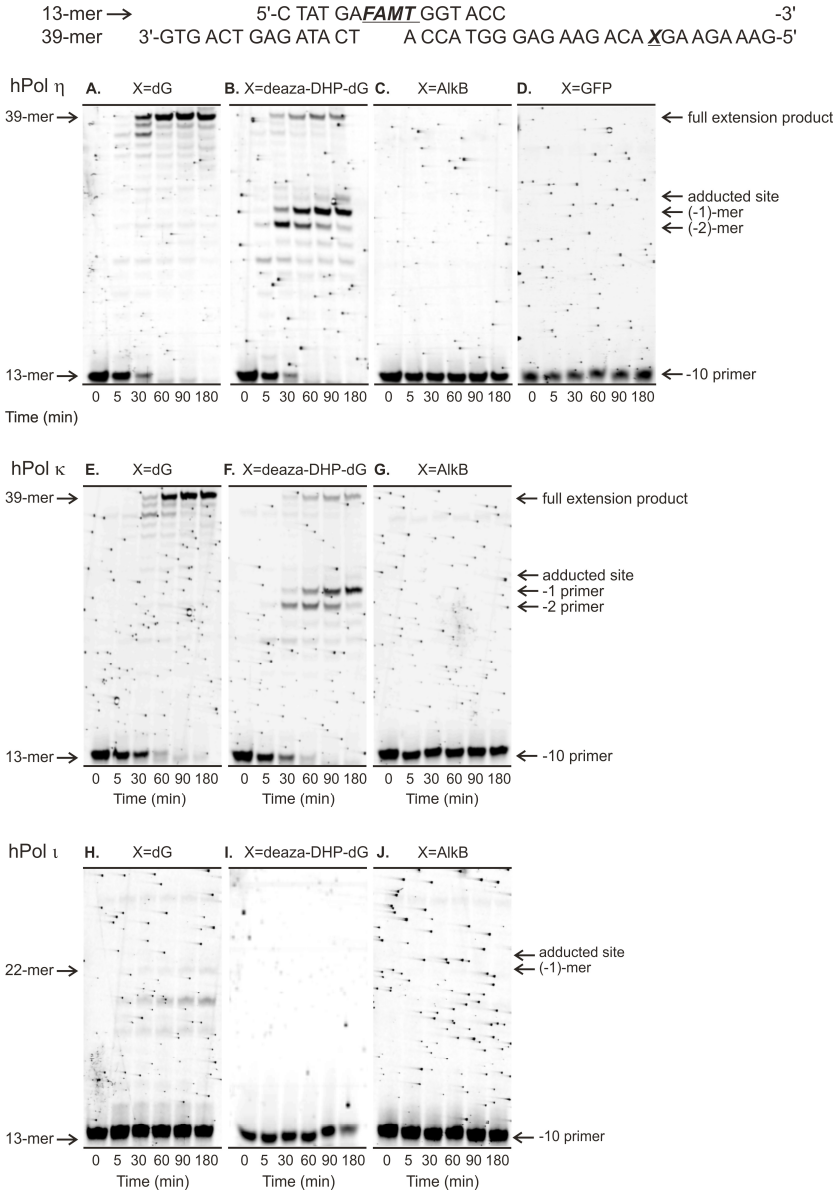
Our standing start assays employed a 9-mer primer (5'-CAA ***FAMdT***GC TTG-3') extending to the penultimate position of the modified site (-1 primer) on the 20-mer template 5'-G TCA CTG GTA ***X***CA AGC ATT G-3', where ***X*** = unmodified dG, deaza-DHP-dG or DPC adduct (**Scheme 4-2**B). A complete extension to a 20-mer product was observed for unmodified and deaza-DHP-dG-containing template strands upon primer extension by hPol η and κ (**Figure 4-1**A, B, F and G), although the extension across from deaza-DHP-dG was relatively inefficient. In contrast, all three DPCs completely blocked DNA synthesis, producing no extension products (**Figure 4-1**C–E, H and I). hPol ι incorporated 1–4 nucleotides opposite unmodified dG template to give 10–13-mer products, however deaza-DHP-dG and DPC lesions completely blocked replication (**Figure 4-1**J–M).

Figure 4-1 Standing start assays for replication bypass of DNA-protein cross-links by hPol η, κ and ι. 9-mer FAMdT primers were annealed with 20-mers containing unmodified dG, deaza-DHP-dG or covalent cross-links to AlkB, histone H4, or 6×His-eGFP. The resulting primer-template duplexes (0.15 μM) were incubated in the presence of hPol η (0.30 μM, A–E), κ (0.30 μM, F–I) or ι (0.30 μM, J–M). Reactions were initiated by the addition of a mixture of dNTPs (500 μM), and quenched at the indicated time points. Extension products were separated by 20% denaturing PAGE, and visualized by fluorescence imaging.



Running start experiments were conducted using a 13-mer primer (5'-CTA TGA ***FAMdT***GG TAC C-3') annealed to a 39-mer template 5'-GAA AGA AG***X*** ACA GAA GAG GGT ACC ATC ATA GAG TCA GTG-3', where ***X*** = unmodified dG, deaza-DHP-dG or DPC. In this primer-template complex, the 3'-end of the primer is placed ten nucleotides upstream from the adducted site (-10 primer) (**Scheme 4-2**B). We found that hPol η and κ copied the templates containing unmodified dG and deaza-DHP-dG to generate full length 39-mer products (**Figure 4-2**A, B, E and F). DNA synthesis using deaza-DHP-dG-containing template was less efficient as compared to the dG control, with pause sites observed at the adducted site and 1–2 nucleotides upstream from the adduct (**Figure 4-2**B and F). As observed in standing start experiments, the DPC lesions completely blocked DNA replication (**Figure 4-2**C, D and G).

Figure 4-2 Running start assays for replication bypass of DNA-protein cross-links by hPol η, κ and ι. 13-mer FAMdT primers were annealed with 39-mers containing unmodified dG, deaza-DHP-dG or covalent cross-links to protein. The resulting primer-template complexes (0.15 μM) were incubated in the presence of hPol η (0.30 μM, A–D), κ (0.30 μM, E–G) or ι (0.30 μM, H–J). Reactions were initiated by the addition of the four dNTPs (500 μM), and quenched at the indicated time points. Extension products were separated on a 20% denaturing PAGE gel, and visualized by fluorescence imaging.



4.3.3 Primer Extension Past DNA-Peptide Cross-links

It has been hypothesized that the protein component of bulky DPC lesions can be proteolytically cleaved to smaller DNA-peptide adducts to facilitate lesion repair and polymerase bypass.[297,330,331] Therefore, we next addressed the ability of human TLS DNA polymerases to bypass smaller lesions containing peptides conjugated to 7-deaza-G in DNA (**Scheme 4-1**). Two biologically relevant peptides were employed: a 10-mer peptide derived from c-Myc protein (EQKLISEEDL)[390,391] and a 23-mer peptide derived from tetanus toxoid (PDAQLVPGINGKAIHLVNNESSE).[392] Both peptides were conjugated to either 20-mer (5'-G TCA CTG GTA ***X***CA AGC ATT G-3') or 39-mer (5'-GAA AGA AG***X*** ACA GAA GAG GGT ACC ATC ATA GAG TCA GTG-3') DNA templates, where ***X*** = peptide cross-link, for primer extension assays (**Scheme 4-2**B and C).

In standing start experiments, 9-mer primers annealed to templates containing unmodified dG, deaza-DHP-dG, and the 10-mer peptide were fully extended by hPol η and κ, albeit with varying efficiencies (**Figure 4-3**A–C and E–G). Interestingly, the 23-mer peptide cross-link completely blocked both polymerases (**Figure 4-3**D and H). In contrast, both 10-mer and 23-mer peptide conjugates were bypassed by hPol η and κ under running start conditions (**Figure 4-4**), although a pronounced pausing was observed 1–2 nucleotides before the adducted site, and extension beyond the modification was very inefficient. In cellular systems, polymerase switching may facilitate the translesion bypass of DNA-peptide cross-links.

Figure 4-3 Standing start assays for replication bypass of DNA-peptide cross-links by hPol η and κ. 9-mer FAMdT primers were annealed with 20-mers containing unmodified dG, deaza-DHP-dG or covalent cross-links to 10-mer peptide or 23-mer peptide. The resulting primer-template duplexes (0.15 μM) were incubated in the presence of hPol η (0.30 μM, A–D) or κ (0.30 μM, E–H). Reactions were initiated by the addition of the four dNTPs (500 μM), and quenched at the indicated time points. Extension products were separated on a 20% denaturing PAGE gel, and visualized by fluorescence imaging.

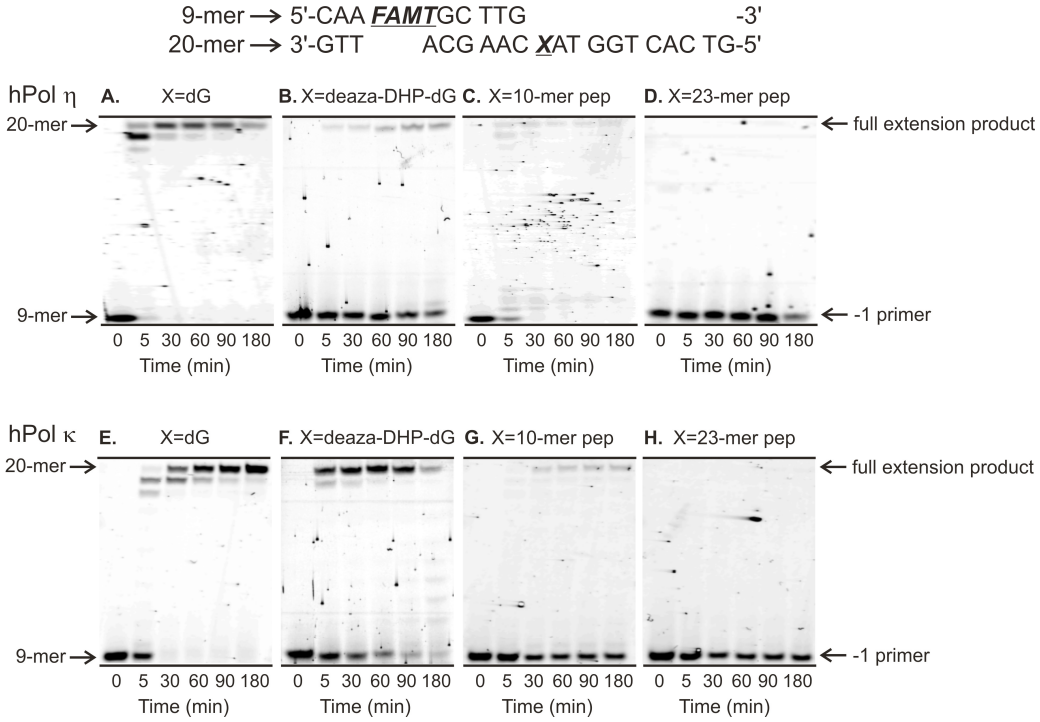
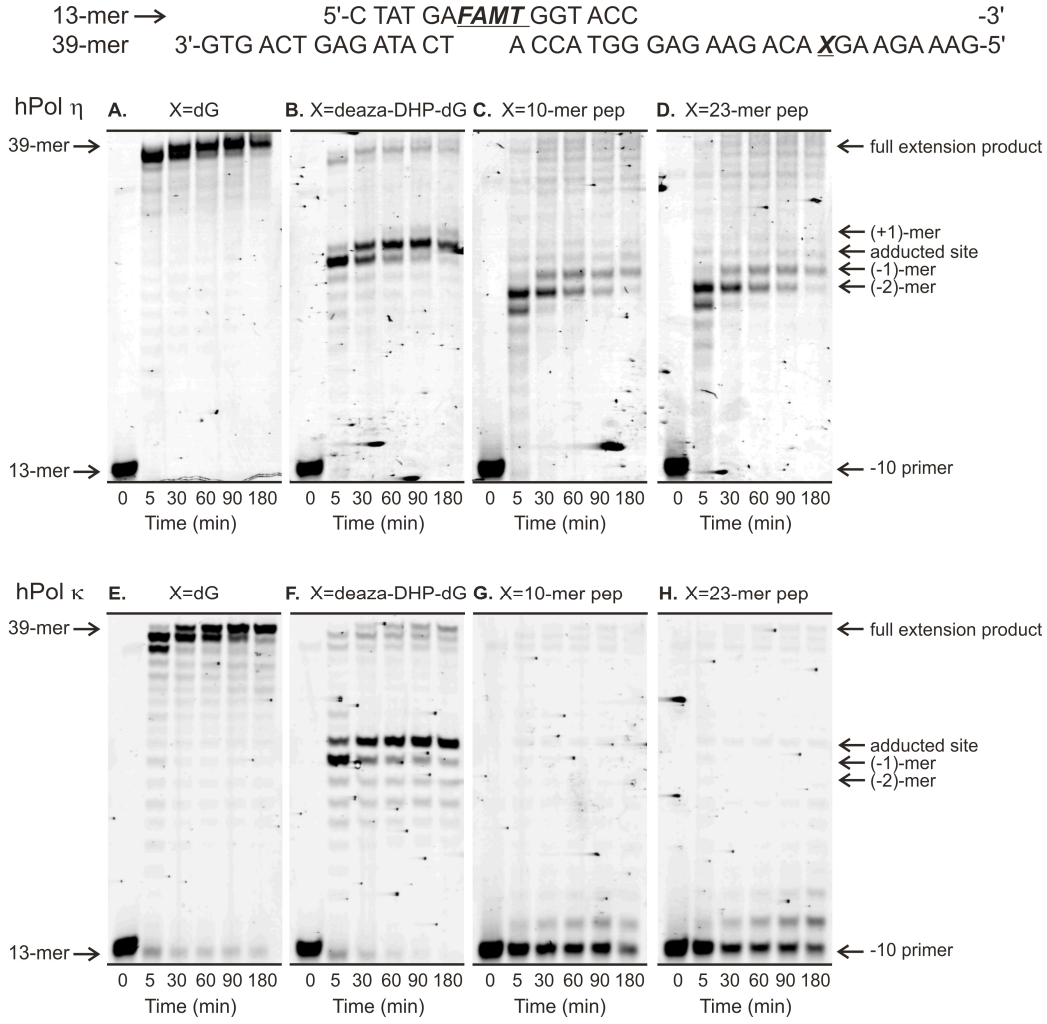


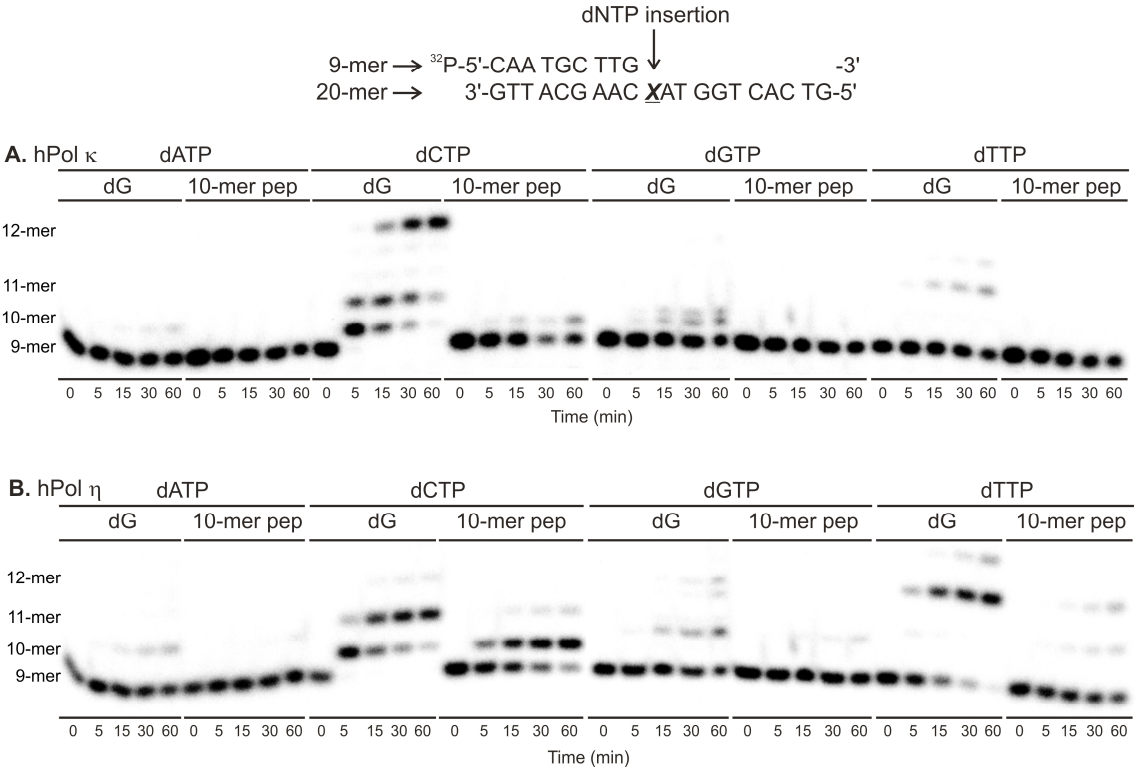
Figure 4-4 Running start assays for replication bypass of DNA-peptide cross-links by hPol η and κ. 9-mer FAMdT primers were annealed with 20-mers containing unmodified dG, deaza-DHP-dG or covalent cross-links to 10-mer peptide or 23-mer peptide. The resulting primer-template duplexes (0.15 μM) were incubated in the presence of hPol η (0.30 μM, A–D) or κ (0.30 μM, E–H). Reactions were initiated by the addition of the four dNTPs (500 μM), and quenched at the indicated time points. Extension products were separated on a 20% denaturing PAGE gel, and visualized by fluorescence imaging.



4.3.4 Single Nucleotide Incorporation Opposite a 10-mer Peptide Cross-Link

Single nucleotide insertion assays were conducted to examine the fidelity of nucleotide insertion opposite the adducted site by TLS polymerases. The 20-mer templates (5'-G TCA CTG GTA ***X***CA AGC ATT G-3') containing unadducted dG or 10-mer peptide cross-link were annealed to 9-mer primers (5'-CAA TGC TTG-3') radiolabeled with ^{32}P (**Scheme 4-2**D). The resulting primer-template complexes were incubated with human bypass polymerase η or κ in the presence of individual dNTPs (100 μM) for 0–60 min. Denaturing PAGE-phosphorimaging analysis has revealed that both polymerases preferentially incorporated the correct base (dC) opposite G, as expected (**Figure 4-5**). Interestingly, single nucleotide insertion by hPol κ opposite the 10-mer peptide cross-link was essentially error-free, even with 3-fold excess of the enzyme with respect to the primer-template duplex (**Figure 4-5**A). Furthermore, hPol η also favored the incorporation of the correct base, dC, although small amount of dT was also incorporated (**Figure 4-5**B). Further, it was evident from these experiments that the nucleotide insertion efficiency opposite small DPCs is higher for hPol η as compared to hPol κ (**Figure 4-5**).

Figure 4-5 Single nucleotide insertion opposite unmodified guanine (dG) and the 10mer peptide cross-linked to *C7* of 7-deaza-guanine (10-mer pep) by TLS polymerase hPol η and κ. 50 nM primer-template duplexes were incubated with hPol η (50 nM, A) and κ (150 nM, B) in the presence 100 μM individual dNTPs. The reactions were quenched at predetermined time points (0–60 min), the products were separated on a 20% denaturing PAGE gel, and visualized by phosphorimaging.



4.3.5 Steady-State Kinetic Analyses of Nucleotide Incorporation Opposite the 10-mer Peptide Cross-Link

Steady-state kinetic studies were performed to determine the catalytic efficiency and to calculate the misinsertion frequency for incorporation of individual dNTPs opposite the 10-mer peptide cross-link by TLS polymerases. Primer-template duplexes containing unmodified dG or 10-mer peptide cross-link were incubated with hPol η or κ in the presence of increasing concentrations of individual dNTPs (0–500 μM), and the reactions were quenched at preselected time points (0–60 min). Polymerase concentration and the time points were selected such that the extent of product formation was <35% of the starting substrate concentration. The specificity constant (k_{cat}/K_m), a measure of the catalytic efficiency of incorporation of each dNTP and the misinsertion frequency (*f*), a quantitative measure of incorporating an incorrect vs. correct dNTP opposite the lesion[1] were calculated by plotting the reaction velocity against dNTP concentration (**Table 4-1**).

The specificity constants (k_{cat}/K_m) for nucleotide insertion by hPol η were 4–150-fold higher as compared to those of hPol κ, suggesting hPol η is more efficient in replication bypass. This was also evident from the qualitative data obtained from single nucleotide insertion assays. Moreover, the specificity constants for the insertion of the correct base, dCTP opposite the adduct were 0.03 and 0.0002 μM^{-1} min^{-1} for hPol η or κ, respectively (**Table 4-1**). These values were 49 and 1850-fold lower than those obtained for dCTP insertion opposite the unmodified dG (1.46 and 0.37 μM^{-1} min^{-1}, **Table 4-1**).

This is expected, since accommodation of the bulky adduct in the active site and nucleotide insertion are unfavorable processes with respect to the unmodified bases.

Nucleotide insertion opposite the 10-mer peptide conjugate by hPol κ was error-free. Although hpol η can incorporate an incorrect base, dT opposite the adduct, overall the nucleotide insertion is mostly error-free. The misinsertion frequency (*f*) for dTTP was 500-fold lower compared to the incorporation of the correct nucleotide, dCTP (**Table 4-1**). Hence, our steady-state kinetic data suggests high fidelity nucleotide insertion opposite the 10-mer peptide cross-linked to *N7* of guanine in DNA.

Table 4-1 Steady-state kinetic parameters for single nucleotide insertion opposite the positive control (dG) and the 10mer peptide cross-linked to *C7* of deazaguanine (10-mer peptide) by human TLS polymerases, hPol η and κ.

Polymerase	***Template***	***Incoming nucleotide***	k_{cat} ***min*$^{-1}$**	K_m ***μM***	k_{cat}/K_m ***μM*$^{-1}$ ***min*$^{-1}$**	***f***
hPol η	dG	dCTP	3.73 ± 0.23	2.55 ± 1.04	1.46	1
		dTTP	0.70 ± 0.08	120.20 ± 38.39	0.006	0.004
	10-mer peptide	dCTP	0.38 ± 0.04	13.11 ± 4.92	0.03	1
		dTTP	0.01 ± 0.001	31.53 ± 11.17	0.00005	0.002
hPol κ	dG	dCTP	0.53 ± 0.04	1.46 ± 1.15	0.37	1
	10-mer peptide	dCTP	0.02 ± 0.004	62.69 ± 23.98	0.0002	1

4.4 DISCUSSION

Cross-linking drugs such as nitrogen mustards, haloethylnitrosoureas, and platinum antitumor agents are commonly used to treat a variety of neoplasms including leukemia, lymphoma, ovarian adenocarcinoma, and breast, lung, and testicular cancer.[269,274,278,286,288] Upon entering the cell nucleus, these drugs form a complex mixture of DNA lesions including nucleobase monoadducts, interstrand and intrastrand DNA-DNA cross-links, and DNA-protein cross-links (DPCs).[263] In contrast to the wealth of information about interstrand DNA-DNA cross-links, which are thought to induce cancer cell death by preventing DNA strand separation,[261,393] little is known about how DPC lesions contribute to the biological effects of *bis*-alkylating agents. Because of their enormous size as compared to "normal" DNA lesions, DPCs are hypothesized to block DNA replication, leading to toxicity.[1,4] However, the ability of human DNA polymerases to replicate DPC-containing DNA has not been systematically studied.

The majority of DNA-protein conjugates induced by antitumor nitrogen mustards such as mechlorethamine, chlorambucil, platinum compounds, chlorooxirane, and diepoxides such as 1,2,3,4-diepoxybutane involve the *N7*-G position of DNA.[85,86,158,159,161,294] However, with the exception of platinum adducts, *N7*-G lesions are hydrolytically labile,[375] making it unfeasible to synthesize site-specific substrates containing such adducts for polymerase bypass assays. We recently developed a reductive amination methodology to prepare hydrolytically stable model DPC substrates that resemble DPCs formed by nitrogen mustards and chlorooxirane (Chapter

3).[354] In the present study, this methodology was employed to prepare DNA-peptide and DNA-protein substrates of increasing size by systematically varying the size of the polypeptide (1.3–28.4 kDa, **Scheme 4-1**). The resulting template–primer complexes were subjected to primer extension experiments to investigate the replication bypass of these bulky lesions by human TLS polymerases.

Lesion bypass is an important mechanism of cellular DNA damage tolerance, which allows for DNA replication past bulky DNA adducts that block replicative polymerases α, δ and ε.[130,131,133-135] Because of the super bulky nature of DPCs, they are postulated to completely block these high fidelity polymerases, leading to stalling of the replication fork.[131,132] A specialized group of polymerases called translesion synthesis (TLS) polymerases can be recruited to blocked replication forks.[130,133,134] Human Y family polymerases hPol η, κ, ι and Rev1, A family polymerase ν, and B family polymerase ζ are TLS polymerases that can trade places with replicative polymerases δ/ε using the PCNA sliding clamp and insert nucleotides opposite bulky nucleobase adducts.[130,131,136-140,142] TLS polymerases are catalytically less efficient than replicative enzymes and are much more error-prone owing to their open and flexible active site that accommodate large DNA adducts[131,142,144,146] and the lack of intrinsic 3'→5' proofreading activity.[129,130,139,143,149,150]

Our *in vitro* DNA replication studies employing model DNA-protein cross-links at the 7-deazaguanine of DNA (**Scheme 4-1**) suggest that these super bulky lesions represent a complete replication block, even in the presence of human lesion bypass

polymerases η, κ and ι (**Figures 4-1** and **4-2**). In both standing and running start experiments, templates containing unmodified dG and deaza-DHP-dG showed full extension with varying efficiency for both hPol η and κ. However, primers annealed with the *C7*-G templates cross-linked to histone H4 (11.4 kDa), AlkB (22.9 kDa) and 6×His-eGFP (28.4 kDa) completely blocked replication, irrespective of the size of the protein. Our results are in agreement with previously reported polymerase blockage by DPCs *in vivo*.[132,326,339] Kreuzer and coworkers found bubble and Y molecule accumulation in wild-type pBR322 plasmids containing MTase (53.5 kDa) cross-linked to 5-azacytosine in *E. coli*, but not in *Eco*RII methylation site mutant.[132] Nakano *et al.* found that transformation efficiency of pGL3-CMV plasmids containing formaldehyde-mediated histone H1 (21.7 kDa) DPCs was <10% compared to undamaged plasmids.[326] In another study, relative transformation efficiency of pMS2 plasmids containing UV-induced T4-pdg (16 kDa) DPCs was found to be <5%, while that of UV-irradiated plasmids was ~50%.[339] These results suggest that DPCs significantly hinder replication *in vivo*, irrespective of the identity of the protein or the cross-linking agent. Observation that transformation efficiency of pGL3-CMV plasmids containing partially digested histone H1 DPCs increased to 58% from <10% for undigested DPC containing plasmids suggest that size of the cross-linking polypeptide plays an important role in replication bypass.[326]

It has been proposed that the protein component of cellular DPCs undergoes proteolytic degradation to peptides, generating smaller lesions that can be bypassed in an error-free or in an error-prone manner.[169,297,324,331] To test this hypothesis, we examined the replication bypass of a 23-mer and a 10-mer peptide cross-linked to *C7*-G. In standing

start assays, 23-mer peptide cross-link completely blocked replication, while the 10-mer peptide cross-link was bypassed with extremely low efficiency (**Figure 4-3**). Replication blockage by *C7*-G adducted 23-mer peptide contradicts previous reports that tetra- and dodecylpeptides cross-linked to major groove of DNA *via* N^6-A were bypassed by Pol ν,[336] while same peptides adducted to N^2-G were also bypassed by hPol κ and Pol IV efficiently.[335] This discrepancy may be due to the structural differences of DPCs such as site of cross-linking on DNA or peptide size employed. In contrast, TLS polymerases were able to replicate past both 10-mer and 23-mer peptide substrates under running start conditions (**Figure 4-4**). Major polymerase pause sites were observed one–two nucleotides prior to the adducted site, despite high replication efficiency observed initially (**Figure 4-4**). Nucleotide insertion as well as the extension past the peptide adducts were extremely inefficient. Further, the bypass efficiency decreased in the the order of dG > deaza-DHP-dG > 10-mer peptide demonstrating the influence of the size of the adduct on replication (**Figure 4-3**). These observations are not surprising, because polymerase loading and nucleotide insertion may be unfavorable at the immediate vicinity of a bulky lesion.

To determine what nucleotides are incorporated opposite the adduct upon bypass of the 10-mer peptide lesions, we conducted single nucleotide insertion assays (**Figure 4-5**). We found that hPol η is catalytically more efficient than hPol κ in bypassing these DPCs. hPol η showed ~80% dCMP insertion within 60 min with DNA:enzyme = 1:1 (**Figure 4-5**B). In contrast, hPol κ only catalyzed the incorporation of ~20% dCMP after 60 min with DNA:enzyme = 1:3 (**Figure 4-5**A). These results were

confirmed by steady-state kinetic studies. All the specificity constants (k_{cat}/K_m), which measure the catalytic efficiency of the enzyme, for hPol η were considerably higher compared to those of hPol κ (**Table 4-1**). Higher catalytic efficiency of hPol η over hPol κ is consistent with previous reports, which employed either exocyclic DNA lesions[112,394,395] or DNA-peptide cross-links.[338]

Fidelity of replication bypass is crucial in the context of cytotoxicity and mutagenicity of DNA lesions. Qualitative data obtained from our single nucleotide insertion assays provided evidence for high fidelity replication bypass of model DNA-peptide cross-links to the major groove of DNA *via* a purine base. Replication bypass of the 10-mer peptide cross-linked to *C7*-G by hPol κ was error free, while hPol η preferentially incorporated the correct base, dC (**Figure 4-5**). According to the misinsertion frequencies (*f*) calculated from steady-state kinetic data, hPol η showed 500-fold preference to insert the correct base, dC opposite the adduct over the mispair, dT (**Table 4-1**). High fidelity observed with hPol ν and hPol κ in bypassing N^6-A and N^2-G adducted peptides, respectively,[335,336] corroborate the high fidelity detected with the Y family polymerases in our study.

The high fidelity of bypass of 10-mer peptide-DNA cross-links observed in the present study is in contrast with the highly error-prone bypass and postlesion synthesis observed when the same 10-mer peptide was cross-linked to *C5* of thymidine (Chapter 6). In the latter study, both hPol η and κ incorporated all four dNTPs opposite the *C5*-T peptide cross-links, while hPol η preferentially incorporated the mispair, dG opposite the

adduct. These pronounced differences in fidelity can be a result of the different cross-linking site on DNA. Unlike 7-deazaguanine DPCs, *C5*-T cross-links appear to interfere with the ability of the DNA polymerases to accurately replicate the adducted nucleobase. However, it is interesting that TLS polymerases distinguish an adducted purine accurately, but not an adducted pyrimidine. Future X-ray crystallographic or NMR structural studies can provide insights into the variability in fidelity of these lesion bypass polymerases to purine vs. pyrimidine DNA adducts.

In summary, we examined the ability of human lesion bypass polymerases to catalyze replication past DNA-protein and DNA-peptide cross-links conjugated to the major groove of DNA *via N7* of guanine. We found that large DNA-protein and DNA-peptide cross-links represent a complete block to human TLS polymerases, while a smaller 10-mer peptide cross-link was bypassed with high fidelity. Our 7-deazaguanine DNA substrates are structurally analogous to major DNA-protein cross-links formed *in vivo* when nitrogen mustards and chlorooxirane-mediate DNA-protein cross-linking.[158,159,376] Hence, our results suggest that DPCs formed to *N7* of guanine in cells get proteolytically degraded to smaller peptide cross-links, which are then bypassed by human TLS polymerases with high fidelity. Sequencing of primer extension products using a liquid chromatography-tandem mass spectrometry methodology is currently in progress. These studies will provide insights into any insertion/deletion mutations caused upon replication bypass of these biologically relevant DNA-peptide cross-links.

5 SYNTHESIS OF SITE-SPECIFIC DNA-PROTEIN CROSS-LINKS CONJUGATED TO THE C5 OF THYMIDINE AND THEIR EFFECTS ON DNA REPLICATION

5.1 INTRODUCTION

DNA-protein cross-linking is non-random, with specific amino acid side chains (typically cysteine, lysine, or arginine) participating in cross-linking.[86,159,294] In addition, acrolein, crotonaldehyde, and 4-hydroxynonenal can form Schiff base cross-links between DNA and the *N*-terminal α-amino groups of proteins.[197] Despite their ubiquitous nature, the biological consequences of DPC formation have not been fully elucidated, probably a result of their inherent structural complexity and the limited availability of structurally defined DPC substrates. It has been hypothesized that covalent DNA-protein conjugates induced by reactive oxygen species may play a role in the etiology of neurodegenerative and cardiovascular diseases due to their deleterious effects on DNA replication, transcription, repair, and chromatin remodeling.[5,35] Indeed, our recent experiments employing epoxide-functionalized protein reagents that selectively induce DPCs have provided the first direct evidence for the ability of DNA-protein cross-links to induce toxicity and mutations in human cells.[304]

Our laboratory has been developing novel methodologies to generate synthetic DPCs structurally analogous to DPC adducts found in cells. We recently reported the use of a reductive amination strategy to create a DPC between an *C7*-deaza-G in DNA and basic lysine or arginine side chains of proteins and peptides (Chapter 3).[354] The resulting model DPC substrates were site-specific within DNA, but involved multiple cross-linking sites within the protein.[354]

In the present work, a bioorthogonal approach employing copper-catalyzed [3+2] Huisgen cycloaddition (click reaction) between azide-functionalized proteins/peptides and alkyne-containing DNA was used to generate structurally defined DPC conjugates. The azide groups were incorporated *via* synthetic methods for short peptides and enzymatically for a larger protein, while alkyne-containing DNA was generated by solid phase synthesis. The resulting cross-links are site specific with regard to both protein and DNA. Synthetic DNA-protein conjugates were subjected to *in vitro* DNA replication experiments in order to evaluate the ability of human DNA polymerases to bypass these bulky lesions.

5.2 MATERIALS AND METHODS

5.2.1 Materials

C8-alkyne-dT-CE phosphoramidite and all other reagents for DNA synthesis were purchased from Glen Research (Sterling, VA). Synthetic DNA oligonucleotides containing native DNA bases and nucleobase modifications were prepared by solid phase synthesis using an ABI 394 DNA synthesizer (Applied Biosystems, CA), purified by HPLC on a Synergi 4u Hydro-RP 80A column, and desalted using NAP-5 columns (GE Healthcare, NJ). T4 polynucleotide kinase was obtained from New England Biolabs (Beverly, MA). γ-^{32}P ATP was purchased from Perkin-Elmer Life Sciences (Waltham, MA). The unlabeled dNTPs were obtained from Omega Bio-Tek (Norcross, GA). 40% 19:1 Acrylamide/bis solution and micro biospin-6 size exclusion columns were purchased from Bio-Rad (Hercules, CA). Ammonium persulfate, CH_3CN, and EDTA were obtained from Fisher (Fair Lawn, NJ). Tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA) was purchased from AnaSpec Inc. (Fremont, CA). NuPAGE Novex 12% Bis-Tris gels (Life Technologies, Grand Island, NY) were run in 1× NuPAGE MOPS SDS running buffer (Life Technologies, Grand Island, NY) and stained with SimplyBlue SafeStain (Life Technologies, Grand Island, NY). Trypsin was obtained from Promega (Madison, WI), and ZipTips for peptide desalting were purchased from Millipore (Billerica, MA). Sep-Pak C18 SPE cartridges were from Waters (Milford, MA). Recombinant hPol η, hPol κ and hPol ι were obtained from Enzymax

(Lexington, KY). All other chemicals and solvents were purchased from Sigma-Aldrich (Milwaukee, WI) and were of the highest grade available.

5.2.2 Oligodeoxyonucleotide Synthesis

DNA 23-mer (5'-AGG GTT TTC CCA G***X***C ACG ACG TT-3') and 18-mer (5'-TCA T***X***G AAT CCT TCC CCC-3'), where ***X*** = 5-(octa-1,7-diynyl)-uracil (C8-alkyne-dT), were prepared by solid phase synthesis on an ABI 394 DNA synthesizer using commercial phosphoramidites (Glen Research, Sterling, VA). Manual coupling was employed for the incorporation of C8-alkyne-dT. Deprotection of the synthesized oligonucleotides was carried out in NH_4OH at room temperature for 2 days. The corresponding unmodified 18-mer (5'-TCA TTG AAT CCT TCC CCC-3') was prepared by standard solid phase methodology. DNA 13-mer (5'-GGG GGA AGG ATT C-3') and 9-mer (5'-GGG GGA AGG-3') were purchased from Integrated DNA Technologies (Coralville, IA). All synthetic DNA oligomers were purified by HPLC on a Synergi Hydro-RP 80A (10 × 250 mm, 4 μm) column eluted at a flow rate of 3 mL/min. HPLC solvents were 150 mM NH_4OAc (A) and 1:1 mixture of 150 mM NH_4OAc and CH_3CN (B). A linear gradient of 10% to 42% B over 60 min was used. HPLC fractions containing full length oligomers were collected, concentrated under vacuum, and desalted by size exclusion *via* NAP-5 columns. All synthetic DNA strands were characterized by HPLC-ESI^--MS and quantified by UV spectrophotometry.

5.2.3 Preparation of Radiolabeled Oligodeoxynucleotides

DNA 23-mers containing C8-alkyne-dT at the 14th position (5'-AGG GTT TTC CCA G***X***C ACG ACG TT-3') or 18-mers containing an C8-alkyne-dT at the 5th position (5'-TCA T***X***G AAT CCT TCC CCC-3', 2 nmol in 17 μL of water) were incubated in 10× PNK buffer (3 μL) in the presence of T4 PNK (20 U) and γ-^{32}P ATP (30 μC*i*) at 37 °C for 1 h. The mixture was heated at 65 °C for 10 min to inactivate the enzyme and passed through Illustra Microspin G25 columns (GE Healthcare, Pittsburgh, PA) to remove excess γ-^{32}P ATP. DNA primers (5'-GGG GGA AGG ATT C-3' and 5'-GGG GGA AGG-3') used in replication assays were also radiolabeled following the same protocol.

5.2.4 Preparation of 6×His-eGFP-N_3

Green fluorescent protein genetically engineered to contain a hexa-histidine tag (6×His-eGFP) was expressed and purified as previously described.[352] 6×His-eGFP was enzymatically prenylated on the cysteine of the *C*-terminal CVIA CaaX box motif using yeast farnesyltransferase. Enzymatic reactions (total volume = 10 mL) were carried out by incubating a solution of the protein (2 μM) and DTT (5 mM, premixed and incubated at room temp. for 1 h) with $MgCl_2$ (10 mM), Tris-HCl (10 μM, pH 7.5), C_{10} dihydroazide (40 μM) and yeast farnesyltransferase (150 nM) at 30 °C for 2 h. The reaction mixture was concentrated to ~500 μL by Centricon centrifugal filters (MWCO 10,000), and the excess azide was removed by size exclusion with a NAP-5 column eluted with PBS

buffer (50 mM NaH_2PO4, 0.1 M NaCl, pH 7.3). Concentration of the prenylated protein was determined by measuring its absorbance at 488 nm (ε_{488} for 6×His-eGFP = 55, 000).

5.2.5 Preparation of Azide-Containing Peptides

Peptide synthesis was carried out using an automated solid-phase peptide synthesizer (PS3, Protein Technologies Inc., Memphis, TN) employing standard Fmoc chemistry and HCTU coupling procedures. The 10-mer peptide $N_3(CH_2)_3CO$-EQKLISEEDL-NH_2 was synthesized on a Rink-amide-MBHA resin (0.1 mmol scale) by coupling each amino acid in DMF containing Fmoc-amino acid (4 eq.), HCTU (4 eq.), 6-Cl-HOBt (4 eq.) and *N*-methylmorpholine (8 eq.) for 20 min. Fmoc deprotection of the peptide was performed with 20% piperidine in DMF for 5 min twice. *N*-terminal 4-azidobutanoic acid was coupled in DMF containing 4-azidobutanoic acid (4 eq.), HCTU (4 eq.), 6-Cl-HOBt (4 eq.) and DIPEA (8 eq.) for 2 h. The peptide was simultaneously cleaved from the resin and deprotected by the treatment with Reagent K (0.5 g phenol, 0.5 mL H_2O, 0.5 mL thioanisole, 0.25 mL 1,2-ethanedithiol in 10 mL trifluoroacetic acid (TFA)) for 2 h. The solution was concentrated to 2 mL by bubbling with N_2 and precipitated in Et_2O. HPLC purification of the peptide was performed using a Luna C18 (10 × 250 mm, 10 μm) column (Phenomenex, Torrance, CA) eluted with a gradient of aqueous 0.1% TFA (A) and 0.1% TFA in CH_3CN (B). Solvent composition was maintained at 0% B for 5 min, followed by a linear gradient of 0 to 70% B over 50 min at a flow rate of 5 mL/min. Calculated for $C_{55}H_{93}N_{16}O_{21}$, $[M+H]^+$ = 1313.6696, found $[M+H]^+$ = 1313.6436 from ESI^+-MS.

Synthesis of 23-mer peptide $N_3(CH_2)_3CO$-PDAQLVPGINGKAIHLVNNESSE began on a preloaded Fmoc-Glu(OtBu)-Wang resin (0.10 mmol) and the peptide chain was elongated using HCTU/*N*-Methylmorpholine-catalyzed, single coupling steps with protected amino acids (4 eq.) and HCTU (4 eq.) for 30 min. Following complete chain elongation, the peptide's *N*-terminus was deprotected with 10% piperidine in DMF (v/v) and the presence of the resulting free amine was confirmed by ninhydrin analysis. The resin containing the peptide was washed with CH_2Cl_2, dried *in vacuo* overnight, weighed, and divided into two portions for further synthesis on a reduced scale. Using 50.0 μmol of peptide, the free amino terminus was acylated with the 4-azidobutanoic acid (13.5 mg, 100 μmol, 2 eq.) catalyzed by DIC (13 mg, 100 μmol, 2 eq.) in the presence of DIEA (8.6 μL, 5.0 μmol, 0.1 eq.) in DMF (5 mL) for 10 h. After acylation was judged complete by ninhydrin analysis, the resin bound peptide was washed thoroughly with CH_2Cl_2 and dried *in vacuo* for 4 h. The peptide was cleaved from the resin along with simultaneous side chain deprotection by treatment with Reagent K at room temperature for 2 h. The released peptide was collected and combined with TFA washes of the resin before precipitation of the peptide in chilled Et_2O (100 mL). The crude solid peptide was collected by centrifugation, the supernatant was removed, and the resulting pellet was washed 2 times with cold Et_2O (50 mL) repeating the centrifugation and supernatant removal steps each time. The crude peptide was purified using a semi-preparative C18 RP-HPLC column with detection at 220 nm and eluted with a gradient of solvent A (H_2O/0.1% TFA, v/v) and solvent B (CH_3CN/0.1% TFA, v/v). The crude peptide (150 mg) was dissolved in a DMF/H_2O solution (1:5 v/v, 25 mL), applied to the column

equilibrated in Solvent A, and eluted using a linear gradient of 0–70% solvent B over 1.5 h at a flow rate of 5 mL/min. Fractions were analyzed for purity by an analytical C18 RP-HPLC column employing a linear gradient of 0–100% solvent B over 60 min at a flow rate of 1 mL/min and detected at 220 nm. Fractions containing peptide product of at least 90% purity were pooled and concentrated by lyophilization to yield 35 mg (25% yield) of a white solid. A small amount (<1mg) of the resulting purified peptide was dissolved in 10 µl of 0.1% TFA/CH_3CN and diluted 1:50 in a mixture of CH_3CN/H_2O (1:1 v/v) prior to MS analysis. MS analysis was performed using a 50 µL injection and collecting 3000 scans. Calculated for $C_{109}H_{178}N_{34}O_{38}$, $[M+2H]^{2+}$ = 1258.1345, found $[M+2H]^{2+}$ = 1258.1406 from ESI^+-MS.

5.2.6 Copper-Catalyzed Cycloaddition Reaction between 6×His-eGFP-N_3 and Alkyne-Containing DNA

HPLC pure DNA oligodeoxynucleotides (5'-AGG GTT TTC CCA G***X***C ACG ACG TT-3' or 5'-TCA T***X***G AAT CCT TCC CCC-3', 1 nmol), where ***X*** = C8-alkyne-dT, were mixed with 6×His-eGFP-N_3 (6 nmol), 2 µL of Tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl) methyl]amine (TBTA, 5 mM stock in DMSO:tBuOH = 1:4), 20 µL of tris(2-carboxy-ethyl)phosphine (TCEP, 5 mM stock in H_2O), and 20 µL of $CuSO_4$ (5 mM stock in H_2O) in 50 mM phosphate buffer (pH 7.5), in a final reaction volume of 100 µL. The reaction was allowed to proceed for 1.5–2 h at room temperature upon mixing with a rotatory shaker. Following desalting on Micro biospin-6 columns, aliquots of the reaction mixtures were withdrawn and resolved by 12% SDS-PAGE. To visualize DPC formation,

NuPAGE Novex 12% Bis-Tris gels (Life Technologies, Grand Island, NY) were run at a constant voltage of 130 V for 1 h in 1× NuPAGE MOPS SDS running buffer. The reaction mixtures obtained from DNA-protein cross-linking reactions were reconstituted in NuPAGE SDS sample buffer and heated at 70 °C for 10 min prior to loading on the gel. The unreacted protein and DNA-protein conjugates were visualized by staining with SimplyBlue SafeStain. Proteinase K digestion (6 units, at 37 °C for 48 h) was conducted to confirm the presence of protein in slowly moving DNA bands. The reaction yields were quantified by ImageJ software.

5.2.7 Cross-Linking Reaction between Azide-Functionalized Peptides and Alkyne-Containing DNA

DNA oligodeoxynucleotides (5'-AGG GTT TTC CCA G***X***C ACG ACG TT-3' or 5'-TCA T***X***G AAT CCT TCC CCC-3', 15 pmol), where ***X*** = C8-alkyne-dT, were mixed with 1.5 nmol of azide-functionalized peptide 10-mer $N_3(CH_2)_3CO$-EQKLISEEDL-NH_2 or the 23-mer $N_3(CH_2)_3CO$-PDAQLVPGINGKAIHLVNNESSE, 1 μL of TBTA (5 mM stock in DMSO:t-BuOH 1:4), 10 μL of TCEP (5 mM stock in H_2O), 10 μL of $CuSO_4$ (5 mM stock in H_2O), and 50 mM phosphate buffer (pH 7.5) in a final reaction volume of 50 μL. The reaction was mixed using a rotatory shaker at room temperature. After allowing the reaction to proceed for 1.5–2 h, it was quenched by adding 0.5 mM EDTA. An aliquot was radiolabeled using γ-^{32}P ATP as described previously, loaded onto a 20% denaturing PAGE gel, and ran at 15 W for 1.5 h. The products were visualized using a Typhoon FLA 7000 phosphorimager (GE Healthcare, Pittsburgh, PA).

5.2.8 Gel Electrophoresis Purification of DNA-Protein Cross-Links

DNA-protein cross-linking reaction mixtures were desalted by Micro biospin-6 size exclusion columns (BioRad, Hercules, CA), while DNA-peptide reaction mixtures were desalted using Sep-Pak C18 SPE cartridges (Waters, Milford, MA). The resulting solutions were loaded onto 15% or 20% (w/v) denaturing PAGE gels containing 7 M urea. DPC bands were visualized by SimplyBlue SafeStain, excised, and the DPCs were extracted using the flextube gel elution system (IBI Scientific, IA) following the manufacturer's instructions. DNA-peptide conjugates were extracted from the gel by using a freeze-thaw method. Gel pieces were suspended in 1× TE buffer, and subjected to 10 cycles of subsequent freezing in dry ice-ethanol and allowed to thaw at room temp. The gel bands were then incubated at 37 °C for 2 days. Following gel purification, DNA-protein cross-links were desalted using Micro biospin-6 columns, while DNA-peptide conjugates were desalted by SPE on Sep-Pak C18 cartridges.

5.2.9 Mass Spectrometry Analysis of DNA-Protein Cross-Links

DNA-protein cross-links containing 6×His-eGFP protein conjugated to synthetic 23-mer oligodeoxynucleotide (5'-AGG GTT TTC CCA G***X***C ACG ACG TT-3') at position ***X*** were purified by 12% SDS-PAGE and stained with SimplyBlue SafeStain. Gel bands were cut into slices and subjected to reduction with 300 mM DTT (10 μL), followed by alkylation with iodoacetamide (10 μL in 25 mM NH_4HCO_3, pH 7.9). Gel pieces were dehydrated with CH_3CN, dried under vacuum, reconstituted in 25 mM NH_4HCO_3 (pH 7.9, 75 μL), and incubated with PDE I (120 mU) at 37 °C overnight to

digest the DNA portion of the cross-link. The resulting 6×His-eGFP-nucleotide conjugates were subjected to tryptic digestion using MS grade Trypsin Gold at 37 °C for 20 h and desalted using C18 ZipTips (Millipore, Billerica, MA). Samples were dissolved in 0.1% acetic acid (25 μL), and 5–8 μL of this solution was used for MS analysis.

NanoLC-nanospray-MS/MS analysis was conducted using an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, Waltham, MA) in line with a NanoLC-Ultra 2D HPLC system (Eksigent, Dublin, CA). Chromatography was performed using a hand packed Luna C18 capillary column (75 μm ID, 10 cm packed bed, 15 μm orifice, 5 μm particle size). The HPLC mobile phases used were 0.1% formic acid in H_2O (A) and 0.1% formic acid in CH_3CN (B). Peptide mixtures (5 μL) were injected using a 5 μL loop and loaded onto the column with a 1 μL/min flow of 2% B for 5.5 min, at which point the injection valve was switched to the load position, and the flow was reduced to 0.3 μL/min. The following linear gradient profile was then used: 2% to 70% B over 60 min, then to 95% B over 1 min, kept at 95% B for a further 5 min, and decreased to 2% B in 1 min. Finally, the flow rate was increased to 1 μL/min and kept at 2% B for 4 min. Mass spectrometry analyses were performed using an FTMS mass analyzer with a resolution of 60,000 and a scan range of 300–2000. Peptide MS/MS spectra were collected using data-dependent scanning in which one full scan mass spectrum was followed by 8 MS/MS spectra using an isolation width of 2.5 *m/z*, 35% normalized CID collision energy, 1 repeat count and 30 s repeat duration with an exclusion mass width of 5 ppm. Spectral data were analyzed using Proteome Discoverer 1.3 software (Thermo Scientific, 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 the protein FASTA database. Search parameters included trypsin specificity and up to 2 missed cleavage sites.

5.2.10 Mass Spectrometry Characterization of Synthetic DNA-Peptide Cross-Links

DNA-peptide conjugates, 5'-AGG GTT TTC CCA G***X***C ACG ACG TT-3' containing a covalent cross-link from *C5* position of dT to the *N*-terminus of the 10-mer peptide EQKLISEEDL, were generated by copper-catalyzed click chemistry as described above and isolated using 15% or 20% (w/v) denaturing polyacrylamide gels containing 7 M urea. Following elution from the gel using the freeze-thaw method, the cross-links were desalted by SPE on Sep-Pak C18 cartridges. The DNA component of the conjugate was digested to nucleosides with PDE I (120 mU), PDE II (105 mU), DNase (35 U), and alkaline phosphatase (22 U) at 37 °C overnight in 10 mM Tris-HCl/15 mM $MgCl_2$ (pH 7) buffer. The resulting dT-peptide conjugates were dried under vacuum and desalted with C18 ZipTips. Samples were dissolved in 0.1% acetic acid (25 μL), and 5–8 μL of this solution was used for MS analysis. NanoLC-nanospray-MS/MS analysis was conducted using an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, Waltham, MA) as described above for DNA-protein cross-links, with the exception that HPLC was conducted using a gradient of 0.1% formic acid in H_2O (A) and CH_3CN (B). Mass spectrometry analysis of DNA-peptide cross-links was performed using an FTMS mass analyzer with a resolution of 30,000 and a scan range of 300–2000 in the full scan mode

using an isolation width of 1.0 *m/z*, 35% normalized CID collision energy. Peptide MS/MS spectra were collected using an isolation width of 3.0 *m/z*, 40% normalized HCD collision energy with a resolution of 7500 and a scan range of 50–2000 *m/z*.

5.2.11 Polymerase Bypass Assay

Oligodeoxynucleotide primers (5'-GGG GGA AGG ATT C-3' and 5'-GGG GGA AGG-3', 100 pmol) were radiolabeled in the presence of T4 PNK (20 unit) and γ-^{32}P ATP (30 μCi) at 37 °C for 60 min in 1× PNK buffer (total volume = 20 μL). The solutions were heated at 65 °C for 10 min to inactivate the enzyme and passed through Illustra Microspin G25 columns (GE Healthcare, Pittsburgh, PA) to remove excess γ-^{32}P ATP. 5'-^{32}P-labeled primers (50 pmol) were mixed with 2 eq. of HPLC-pure template strands (5'-TCA T***X***G AAT CCT TCC CCC-3' where ***X*** = the click reaction generated covalent cross-link from the *C5* position of dT to the *C*-terminus Cys of 6×His-eGFP, and *N*-terminus of 23-mer peptide (PDAQLVPGINGKAIHLVNNESSE) or 10-mer peptide (EQKLISEEDL)) in 10 mM Tris buffer (pH 7) containing 50 mM NaCl. Control template strands contained unmodified dT at position ***X***. The strands were annealed by heating at 90 °C for 10 min and cooling slowly overnight to afford the desired radiolabeled template-primer duplexes.

Primer-template duplexes (40 nM in the final reaction volume of 40 μL) were incubated with human recombinant DNA polymerases (final concentrations: 160 nM hPol η, 400 nM hPol κ, 80 nM hPol ι) at 37 °C in the presence of a buffered solution containing 50 mM Tris (pH 7.5), 50 mM NaCl, 5 mM DTT, 100 μg/μL BSA,

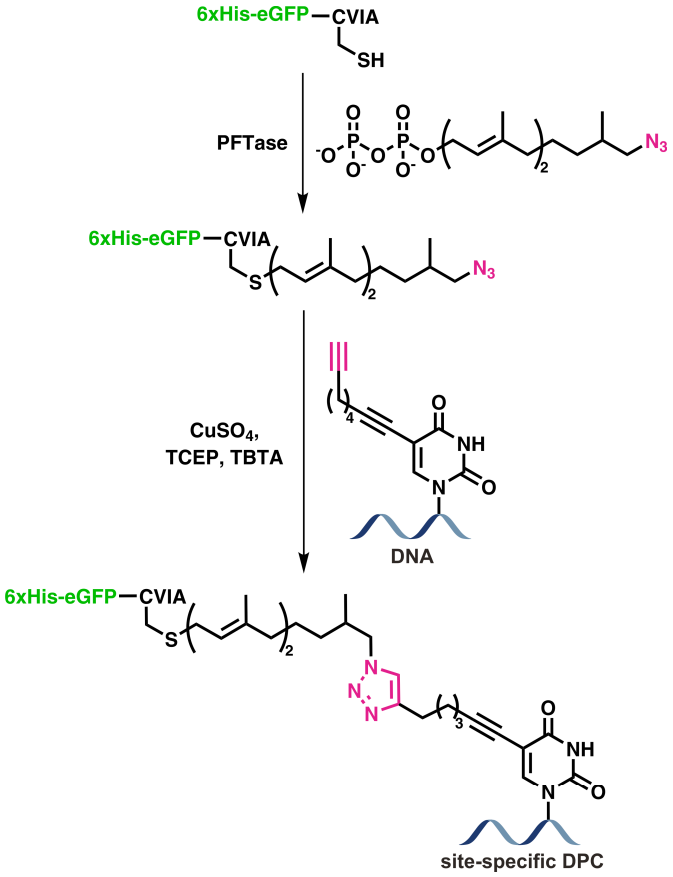
10% glycerol (v/v) and 5 mM $MgCl_2$. Primer extension reactions were initiated by adding 0.5 mM solutions of all four dNTPs. Aliquots of the reaction mixtures (4 μL) were withdrawn at preselected time intervals (0–180 min) and quenched by the addition of 18 μL of a solution containing 95% formamide (v/v), 10 mM EDTA, 0.03% bromophenol blue (w/v) and 0.03% xylene cyanol (w/v). Samples were loaded on to a 20% denaturing polyacrylamide gel containing 7 M urea and run at 80 W for 2.5 h. The extension products were visualized using a Typhoon FLA 7000 instrument in the phosophorimaging mode.

5.3 RESULTS

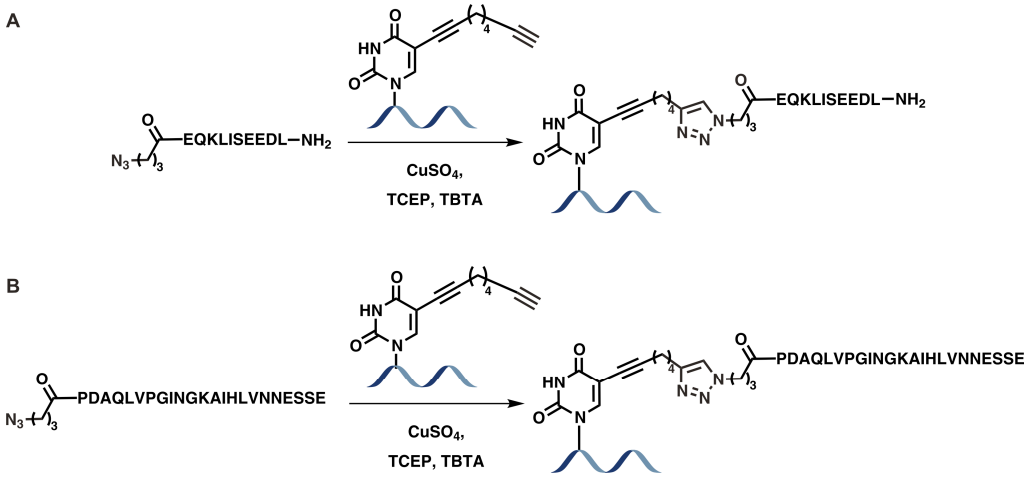
5.3.1 Site-Specific DNA–Protein Cross-Linking Using Alkyne−Azide Cycloaddition (Click) Reaction

Site-specific DPCs were generated *via* 1,3-dipolar cycloaddition between azide-containing proteins/peptides and alkyne-functionalized oligodeoxynucleotides in the presence of copper ([3+2] Huisgen cycloaddition) to give a 1,2,3-triazole (**Schemes 5-1** and **5-2**).[352] To prepare azide-functionalized green fluorescent protein (6×His-eGFP-N_3), a previously described eGFP construct bearing an *N*-terminal His-tag and a *C*-terminal CVIA sequence was employed.[352] The latter sequence allows the cysteine residue within CVIA to be enzymatically prenylated by protein farnesyltransferase (PFTase) using an azide-containing farnesyl diphosphate substrate analogue (**Scheme 5-1**).[352,396,397] Synthetic 10-mer and 23-mer peptides were prepared *via* solid phase peptide synthesis and appended with an *N*-terminal 4-azidobutanoic acid group for subsequent Cu-catalyzed click reaction (**Scheme 5-2**). Synthetic DNA oligomers containing 5-(octa-1,7-diynyl)-uracil (C8-alkyne-dT) were prepared by solid phase synthesis starting with commercial phosphoramidites (Glen Research, Sterling, VA). The resulting biomolecules were purified by HPLC and characterized by mass spectrometry.

Scheme 5-1 Generation of site-specific DNA-protein conjugates by copper-catalyzed [3+2] Huisgen cycloaddition. Click reaction between an alkyne group from 5-(octa-1,7-diynyl)-uracil in DNA and an azide group within modified green fluorescent protein (6×His-eGFP-N_3) in the presence of Cu^I generates triazole-linked DPCs.



Scheme 5-2 Synthesis of site-specific DNA-peptide cross-links by copper-catalyzed azide-alkyne cycloaddition reaction. Synthetic 10-mer and 23-mer peptides were prepared *via* solid phase peptide synthesis and appended with an *N*-terminal 4-azidobutanoic acid group, and reacted with the terminal alkyne of 5-(octa-1,7-diynyl)-uracil in DNA to form site-specific DNA-peptide cross-links.



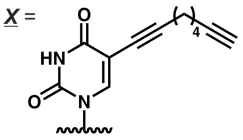
Our initial bioconjugation experiments were conducted using 6×His-eGFP-N_3 protein and a ^{32}P-endlabeled 23-mer oligodeoxynucleotide (5'-AGG GTT TTC CCA G***X***C ACG ACG TT-3') containing site-specific C8-alkyne-dT at ***X*** (**Scheme 5-3**A). Following cycloaddition reaction in the presence of Cu^I, denaturing SDS-PAGE of the reaction mixture revealed the appearance of a new slowly moving band (Lane 2 in **Figure 5-1**), which was not present in the DNA control (Lane 1 in **Figure 5-1**). The high molecular weight band disappeared when the reaction mixture was incubated with proteinase K, confirming that it corresponds to a covalent DNA-protein conjugate (Lane 3 in **Figure 5-1**). The cross-linking yield was estimated as ~70% based on densitometry analysis (Lane 2 in **Figure 5-1**).

Scheme 5-3 Sequences of DNA oligomers used for conjugation reactions with proteins and peptides (A) and DNA substrates employed in standing start (B) and running start primer extension experiments (C).

A 5'-AGG GTT TTC CCA G***X***C ACG ACG TT-3' (23-mer template)

5'-TCA T***X***G AAT CCT TCC CCC-3' (18-mer template)

X =



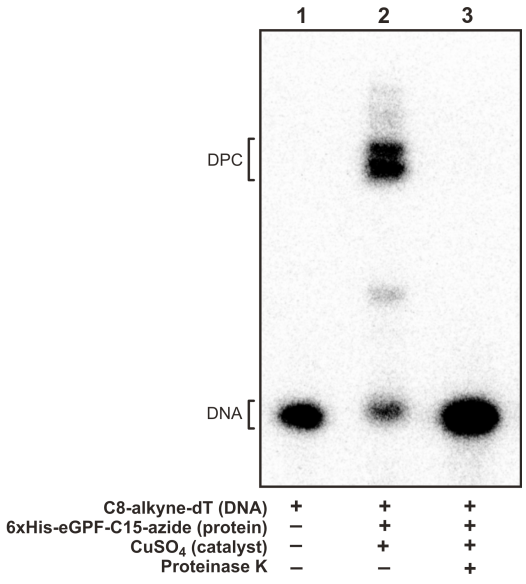
B 5'-TCA T***X***G AAT CCT TCC CCC-3' (18-mer template)
3'- C TTA GGA AGG GGG-5' (13-mer primer)

X = dT (positive control)
EQKLISEEDL (10 mer peptide)
PDAQLVPGINGKAIHLVNNESSE (23 mer peptide)
6xHis eGFP

C 5'-TCA T***X***G AAT CCT TCC CCC-3' (18-mer template)
3'- GGA AGG GGG-5' (9-mer primer)

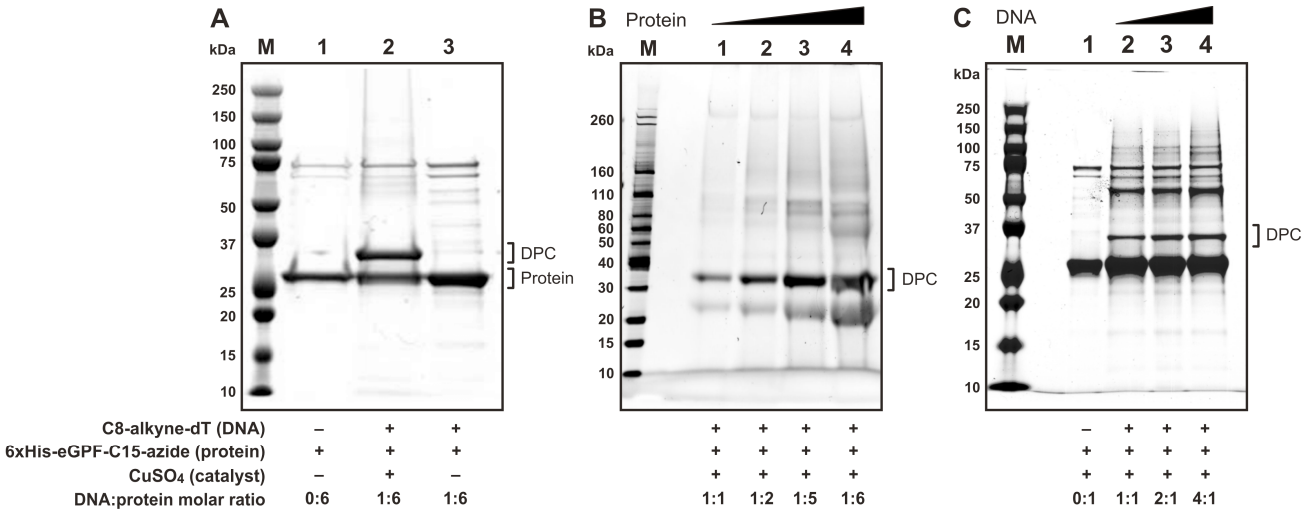
X = dT (positive control)
EQKLISEEDL (10 mer peptide)
PDAQLVPGINGKAIHLVNNESSE (23 mer peptide)
6xHis eGFP

Figure 5-1 SDS-PAGE analysis of site-specific DNA-protein cross-links generated by Cu-catalyzed azide-alkyne cycloaddition. DPCs were generated by using 6×His-eGFP-N_3 protein and ^{32}P-endlabeled DNA 23-mer (5'-AGG GTT TTC CCA G***C8-alkyne-dT***C ACG ACG TT-3', where C8-alkyne-dT is 5-(octa-1,7-diynyl)-uracil). Lane 1: alkyne containing DNA; Lane 2: reaction mixture following cycloaddition between C8-alkyne-dT containing DNA and 6×His-eGFP-N_3 protein; Lane 3: proteinase K digested reaction from lane 2.



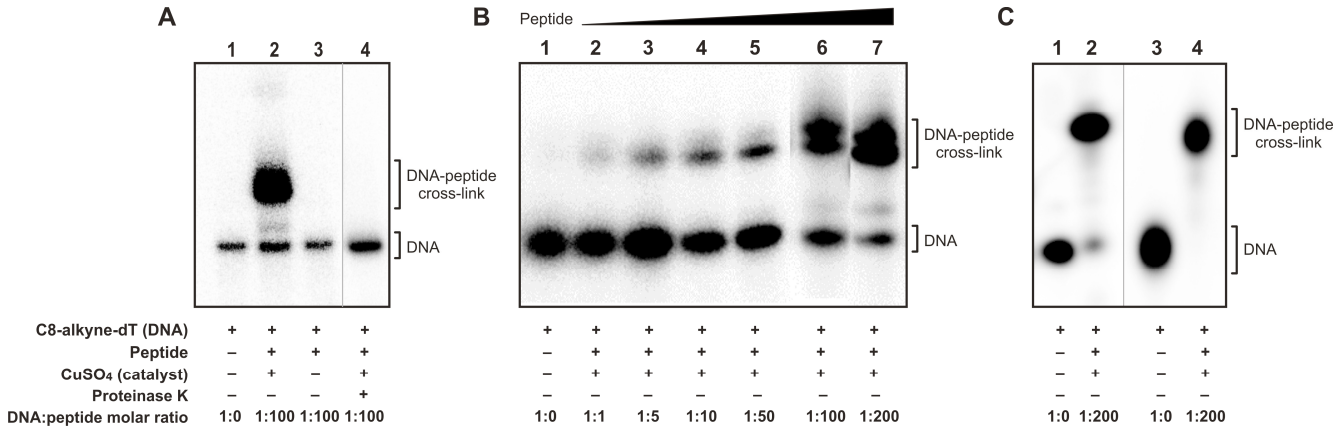
In a separate experiment, unlabeled DNA 23-mer (5'-AGG GTT TTC CCA G***X***C ACG ACG TT-3') was conjugated to 6×His-eGFP-N_3, and the reaction mixture was separated by SDS-PAGE, followed by protein visualization by SimplyBlue SafeStain (**Figure 5-2**A). A new band at ~35 kDa was observed upon analysis of the reaction mixtures (Lane 2 in **Figure 5-2**A), which is consistent with the conjugate of 23-mer oligodeoxynucleotide (7.1 kDa) and 6×His-eGFP-N_3 (28.4 kDa). This band was not observed in protein only control (Lane 1 in **Figure 5-2**A) or in control reactions conducted in the absence of Cu (Lane 3 in **Figure 5-2**A). To examine the influence of DNA-polypeptide molar ratios on the efficiency of DPC formation, the cycloaddition reaction was repeated in the presence of increasing molar equivalents of either 6×His-eGFP-N_3 or DNA, followed by gel electrophoretic analysis (**Figure 5-2**B and C). We found that the DPC yields improved with increasing protein concentration, reaching a maximum yield of DPCs when a 6-fold molar excess of 6×His-eGFP-N_3 was employed (Lane 4 in **Figure 5-2**B). We also observed that the DPC yields improved with increasing DNA concentration (**Figure 5-2**C). These results indicate that site-specific DNA-protein cross-links can be generated in good yields using copper mediated 1,3-dipolar cycloaddition between azide-containing proteins and alkyne-functionalized DNA.

Figure 5-2 SDS-PAGE analysis of site-specific DNA-protein cross-links generated by Cu-catalyzed azide-alkyne cycloaddition (CuAAC). (A) DPCs generated by using 6×His-eGFP-N_3 protein and unlabeled DNA 23-mer containing C8-alkyne-dT were separated by 12% SDS-PAGE and proteins were visualized *via* SimplyBlue staining. Lane M: protein marker; Lane 1: 6×His-eGFP-N_3; Lane 2: reaction mixture following CuAAC between 6×His-eGFP-N_3 protein and alkyne containing DNA; Lane 3: Reaction mixture following cycloaddition in the absence of Cu. (B) The yields of cycloaddition-induced DPCs increased with increasing DNA: protein molar ratios. The reaction was conducted as in (A), but the molar ratio of DNA:6×His-eGFP-N_3 was varied between 1:1 and 1:6. (C) The reaction was conducted as in (A) by increasing the molar equivalents of DNA to protein from 1:1 to 4:1.



DNA-peptide cross-links were similarly generated by cycloaddition reactions between C8-alkyne-dT containing DNA (7.1 kDa) and synthetic azide-containing peptides ($N_3(CH_2)_3CO$-Glu-Gln-Lys-Leu-Ile-Ser-Glu-Glu-Asp-Leu-NH_2, 1.3 kDa or $N_3(CH_2)_3CO$-Pro-Asp-Ala-Gln-Leu-Val-Pro-Gly-Ile-Asn-Gly-Lys-Ala-Ile-His-Leu-Val-Asn-Asn-Glu-Ser-Ser-Glu, 2.5 kDa). As shown for the 10-mer peptide reaction, the presence of the DNA-peptide conjugates (8.4 kDa) was detected by denaturing PAGE (**Figure 5-3**A). A new, low mobility band (Lane 2 in **Figure 5-3**A) corresponding to oligonucleotide-polypeptide conjugate was only found when the reaction was conducted in the presence of Cu (Lane 3 in **Figure 5-3**A), and disappeared upon incubation with proteinase K (Lane 4 in **Figure 5-3**A). A 200-fold molar excess of peptide to DNA was required to achieve optimal yields of DNA-peptide conjugates (81.5% yield, **Figure 5-3**B). The cycloaddition reaction with 23-mer peptide was conducted analogously (78% yield, **Figure 5-3**C).

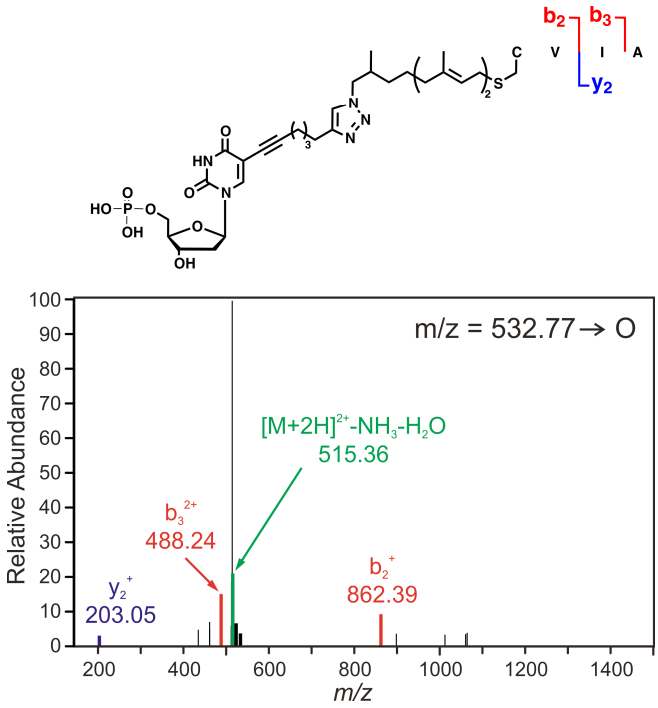
Figure 5-3 Generation of site-specific DNA-peptide cross-links by Cu-catalyzed azide-alkyne cycloaddition. (A) Denaturing PAGE analysis of DNA-peptide conjugates generated using 10-mer peptide ($N_3(CH_2)_3$COEQKLISEEDLNH_2) and ^{32}P-23-mer DNA containing C8-alkyne-dT. The reaction mixtures were resolved on a 20% (w/v) denaturing PAGE gel, and visualized by phosphorimaging analysis. Lane 1: C8-alkyne-dT containing 23-mer; Lane 2: reaction mixture following CuAAC between alkyne containing DNA and peptide-N_3; Lane 3: the same reaction as in Lane 2 in the absence of Cu; Lane 4: proteinase K digested reaction from Lane 2. (B) The DPC yield increased with increasing amount of the peptide. DPCs were generated as in (A) by increasing the molar equivalents of DNA to peptide from 1:1 to 1:200. (C) Denaturing PAGE analysis of CuAAC of 23-mer peptide ($N_3(CH_2)_3$CO-PDAQLVPGINGKAIHLVNNESSE) and 5'-^{32}P-23-mer C8-alkyne-dT. Lane 1: alkyne containing DNA; Lane 2: alkyne containing DNA and peptide-N_3 in the presence of Cu; Lane 3: alkyne containing DNA; Lane 4: gel purified DNA-peptide conjugate from lane 2.



5.3.2 Mass Spectrometry Characterization of DNA–Protein and DNA–Peptide Conjugates

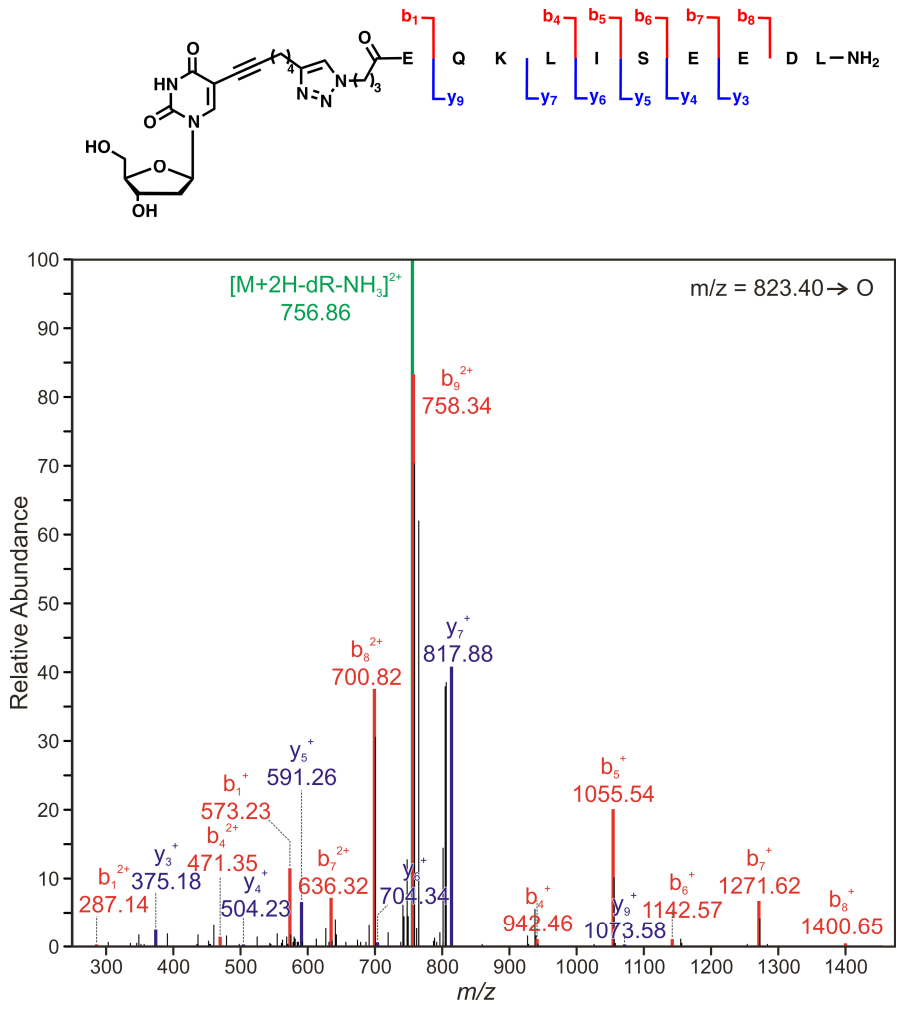
To confirm the formation of covalent DNA-protein and DNA-peptide cross-links, the purified conjugates were characterized by tandem mass spectrometry. In order to characterize the conjugates between 23-mer DNA oligomer and 6×His-eGFP-N_3 protein (**Scheme 5-1**), DNA was digested to nucleotides, while the protein was cleaved to peptides with trypsin. Following SDS-PAGE purification, gel bands containing DPCs were excised and subjected to in-gel digestion with phosphodiesterase I (PDE I) and trypsin, and the resulting peptide-nucleotide conjugates were analyzed by nanoHPLC-nanospray-HRMS/MS using an Orbitrap Velos mass spectrometer. The mass spectral data were processed using Thermo Proteome Discoverer 1.3 software (ThermoScientific, San Jose, CA) to identify the cross-linking site(s). A doubly charged ion at *m/z* 532.77 was observed corresponding to the tetrapeptide CVIA containing a covalent cross-link to dUMP (theoretical mass = 1064.52, **Figure 5-4**). MS/MS fragmentation of *m/z* 532.77 ions under CID conditions gave rise to a series of *b*- and *y*-fragments, including a singly charged b_2 ion at *m/z* 862.39 and a doubly charged b_3 fragment ion at *m/z* 488.24 (**Figure 5-4**). Since the cysteine residue within the sequence CVIA is known to be the site of enzymatic prenylation, these observations are consistent with the predicted site of modification. Taken together, these results are consistent with cycloaddition reaction taking place at the specific cysteine residue of the protein containing the azido modification.

Figure 5-4 Mass spectrometry characterization of DNA-protein conjugates. NanoLC-nanospray-MS/MS spectrum of eGFP tryptic peptide, CVIA, cross-linked to 5-(octa-1,7-diynyl)-2'-deoxyuridine monophosphate. DPCs were generated by Cu-catalyzed cycloaddition between 6×His-eGFP-N_3 and C8-alkyne-dT containing DNA 23-mer, and DPCs were isolated by 12% SDS-PAGE as shown in **Figure 5-1**. DNA component of the DPCs was digested with phosphodiesterase I, and the resulting protein-nucleotide conjugate (*m/z* 532.77, doubly charged) was subjected to tryptic digestion followed by MS/MS analysis on an Orbitrap Velos mass spectrometer.



In order to simplify MS analysis of DNA-peptide cross-links (**Scheme 5-2**A), the DNA component of the cross-link was digested to nucleosides. NanoHPLC-nanospray-MS/MS analysis allowed for the detection of doubly charged peptide species at *m/z* 823.40, which corresponds to the decapeptide EQKLISEEDL containing a triazole cross-link to deoxyuridine (**Figure 5-5**). The doubly charged peptide was subjected to HCD fragmentation within an Orbitrap Velos instrument, and the resulting fragments were analyzed in the accurate mass mode. Both *b*- and *y*-series fragment ions were detected (**Figure 5-5**), and the MS/MS fragmentation under HCD conditions was consistent with the predicted conjugate structure (**Scheme 5-2**A).

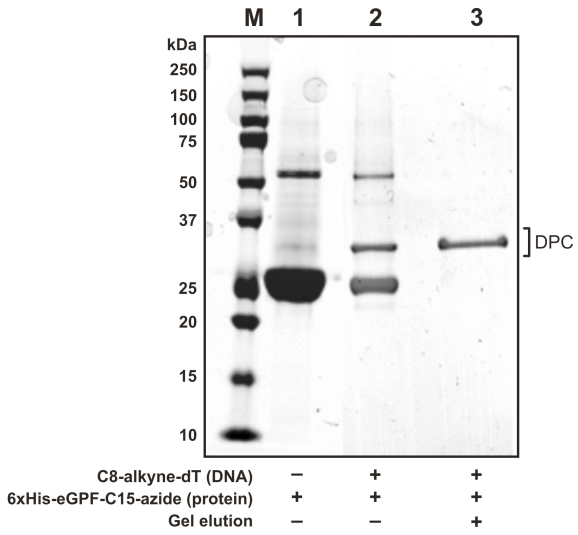
Figure 5-5 Mass spectrometry characterization of DNA-peptide cross-links. NanoLC-nanospray-MS/MS characterization of DNA-peptide conjugates generated using 10-mer peptide ($N_3(CH_2)_3$COEQKLISEEDLNH_2) and C8-alkyne-dT containing DNA 23-mer. Following gel purification as shown in **Figure 5-3**, the DNA component of the cross-link was digested with phosphodiesterases and alkaline phosphatase, and the resulting peptide-nucleoside conjugate (*m/z* 823.40, doubly charged) was sequenced by nanoLC-nanospray-MS/MS on an Orbitrap Velos mass spectrometer.



5.3.3 Polymerase Bypass of Synthetic DNA–Protein and DNA–Peptide Conjugates

The model DPC-containing DNA substrates were subjected to several rounds of purification prior to their use in biochemical assays. The reaction mixtures were initially desalted to remove reagents and salts. DPCs were isolated by SDS-PAGE (DNA-protein conjugates) or 15% or 20% (w/v) denaturing polyacrylamide gel containing 7 M urea (DNA-peptide conjugates), and extracted from the gel using a gel elution kit or a freeze-thaw method. DPC purity was confirmed by analysis of purified material *via* either SDS-PAGE or denaturing PAGE (Lane 3 in **Figure 5-6**). To confirm the removal of excess protein, purified DPCs were fractionated by SDS-PAGE and stained with SimplyBlue SafeStain. Purified DPCs were radiolabeled using γ^{32}P-ATP, and the absence of unreacted oligonucleotides was verified by denaturing PAGE, followed by phosphorimaging. Depending on the purity of DPC substrates obtained from initial isolation, additional gel purifications were carried out. Only conjugates whose purity was greater than 96% were employed in DNA polymerase assays.

Figure 5-6 Purification of DNA-protein cross-links by gel electrophoresis–electroelution. DPCs were generated by CuAAC of 6×His-eGFP-N_3 and 23-mer oligonucleotide (5'-AGG GTT TTC CCA G***C8-alkyne-dT***C ACG ACG TT-3'). The reaction mixture was desalted by micro biospin-6 columns, loaded onto a 12% (w/v) SDS-PAGE gel, and the gel band containing the DPC was excised and applied to electroelution. Lane M: protein marker; Lane 1: 6×His-eGFP-N_3; Lane 2: a CuAAC reaction mixture containing 6×His-eGFP-N_3 and alkyne containing DNA in the presence of Cu; Lane 3: a purified sample *via* electroelution method.

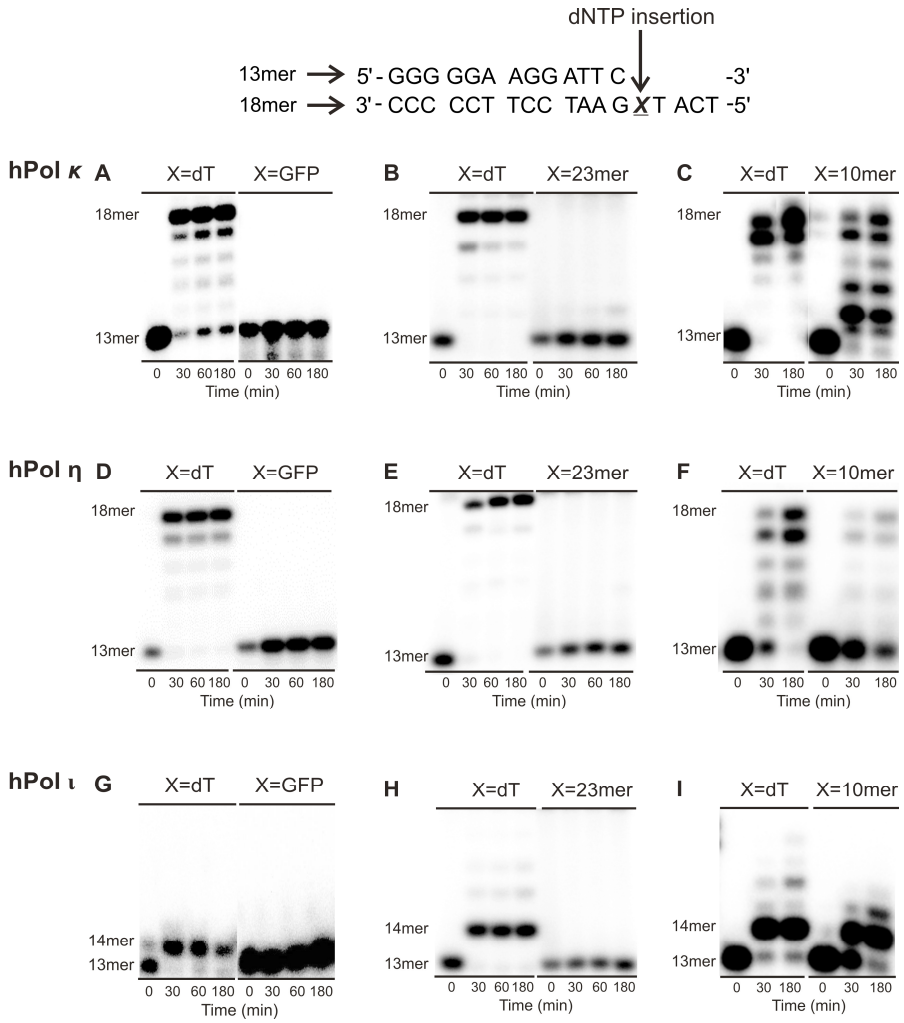


To elucidate the influence of DNA-protein and DNA-peptide cross-links on DNA replication, template-primer complexes containing site-specific cross-links to 6×His-eGFP-N_3 protein, 23-mer peptide (PDAQLVPGINGKAIHLVNNESSE), 10-mer peptide (EQKLISEEDL), and unmodified dT (negative control) were subjected to primer extension in the presence of human translesion synthesis (TLS) polymerases κ, η, and ι. Two types of experiments were conducted: standing start, with the primer extending to the -1 position from the DPC lesion (**Scheme 5-3**B), and running start, with the primer ending four nucleotides upstream from the adduct site on the 18-mer template 5'-TCA T***X***G AAT CCT TCC CCC-3', where ***X*** = unmodified dT or synthetic DPC lesion (**Scheme 5-3**C).

In standing start experiments with control template (***X*** = dT, **Figure 5-7**), both hPol κ (a 10:1 molar ratio of polymerase to primer-template duplex) and hPol η (a 4:1 ratio of polymerase to primer-template) completely extended the primer opposite the control template to form 18-mer products (***X*** = dT, **Figure 5-7**). hPol ι generated mainly a single nucleotide addition product, probably due to its known low processivity as compared to other Y-family polymerases (***X*** = dT, **Figure 5-7**).[398-401] The presence of 6×His-eGFP-dT at position ***X*** completely blocked primer extension by all three human lesion bypass polymerases (**Figure 5-7**A, D and G). Similar results were obtained for the 23-mer peptide conjugate (**Figure 5-7**B, E and H). In contrast, all three polymerases were capable of bypassing the smaller DPC containing a 10-mer peptide, albeit with differing efficiency (**Figure 5-7**C, F and I). Extension products of hPol κ included the complete 18-mer and multiple incomplete extension products (**Figure 5-7**C). Interestingly,

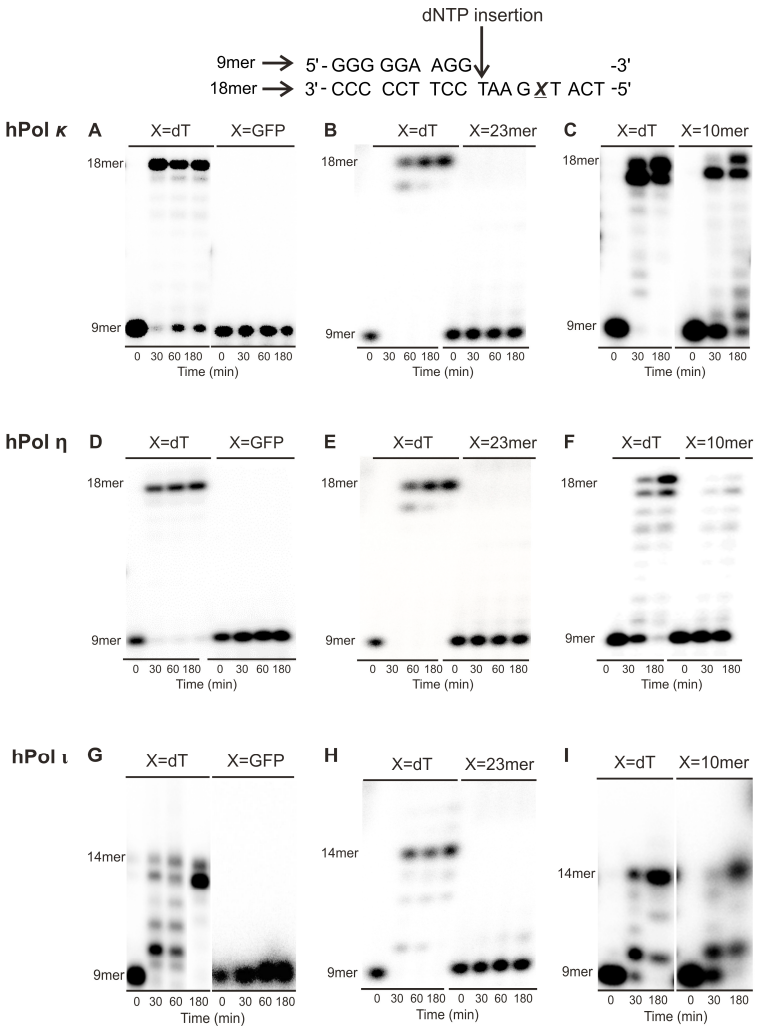
nucleotide incorporation opposite the lesion was more efficient than the addition of subsequent nucleotides, resulting in accumulation of the +1 product (**Figure 5-7**C). In the case of hPol ι, the efficiency of primer extension was significantly lower than the substrate bearing a native dT, but nearly complete conversion of a 13-mer to a 14-mer product was observed in 180 min (**Figure 5-7**I).

Figure 5-7 Replication bypass of DNA-peptide and DNA-protein conjugates of increased size adduct by human lesion bypass polymerases under standing start conditions. The ^{32}P-labeled 13-mer primers were annealed with 18-mer template containing unmodified dT or covalent cross-links to 6×His-eGFP-N_3, 23-mer peptide, or 10-mer peptide. The resulting primer-template complexes (40 nM) were incubated in the presence of hPol κ (400 nM, A–C), hPol η (160 nM, D–F) or hPol ι (80 nM, G–I). The polymerase reactions were started by the addition of the four dNTPs (500 μM) and quenched at the indicated time points. The extension products were separated by 20% (w/v) denaturing PAGE and visualized by phosphorimaging analysis.



For running start experiments, the 18-mer template (5'-TCA T***X***G AAT CCT TCC CCC-3', where ***X*** = dT or DPC adduct containing 10-mer peptide (EQKLISEEDL), 23-mer peptide (PDAQLVPGINGKAIHLVNNNESSE), or 6×His-eGFP-N_3 was annealed to a 9-mer (-4) primer (**Scheme 5-3**C). Complete primer extension by hpol η and hpol κ was observed for the control substrate (***X*** = dT, **Figure 5-8**), while hpol ι produced a +1 (14-mer) product (***X*** = dT, **Figure 5-8**). As was the case for our standing start experiments, hpol κ, η and ι were completely blocked by the cross-links containing 6×His-eGFP-dT and the 23-mer peptide (**Figure 5-8**A, B, D, E, G and H), whereas the presence of a 10-mer cross-link at position ***X*** has led to varied amounts of extended products with hpol κ, η and ι (**Figure 5-8**C, F and I). Low amounts of fully extended products (18-mers) were observed in the experiment with hpol η, suggesting that bypass of DNA-peptide conjugates by hpol η is inefficient (**Figure 5-8**F). In contrast, hpol ι has shown a robust primer extension activity (**Figure 5-8**I), suggesting that it may coordinate with other human polymerases to allow for efficient bypass of small DNA-peptide cross-links *via* polymerase switching.[402]

Figure 5-8 Replication bypass of DNA-peptide and DNA-protein conjugates of increased size adduct by human lesion bypass polymerases under running start conditions. The ^{32}P-endlabeled 9-mer primers were annealed to the 18-mer templates containing unmodified dT, 6×His-eGFP-N_3, 23-mer peptide, or 10mer peptide. The resulting primer-template complexes (40 nM) were incubated at 37 °C in the presence of hPol κ (400 nM, A–C), hPol η (160 nM, D–F) or hPol ι (80 nM, G–I). Reactions were started by the addition of all four dNTPs (500 μM) and quenched at indicated time points. The extension products were resolved by 20% (w/v) denaturing PAGE and visualized by phosphorimaging analysis.



5.4 DISCUSSION

DNA-protein cross-links (DPCs) are among the most abundant and the least understood DNA lesions present in the human genome. Despite their propensity to form in tissues and their likely involvement in human disease, little is known about their cellular effects. However, because of the structural complexity of DPC lesions and the difficulty of generating site specific, chemically defined DPC substrates, there is very limited information, and no consensus, on how cells respond to this class of DNA lesions. This lack of insight hinders our ability to fully understand the molecular basis of the therapeutic and adverse effects associated with a major class of anticancer agents and may limit insight into a fundamental cause of age-related disorders.

As mentioned in Chapter 3, a major limitation in the field is the paucity of DNA substrates containing site-specific, homogeneous, and structurally defined DNA-protein conjugates. Currently available methodologies have several limitations such as poor reaction efficiency and low yields,[164] limited choices of protein reagents (e.g. specific DNA modifying proteins),[331] and insufficient site specificity with respect to the cross-linking site within the protein.[164,246,354] Recently, we have developed a methodology for site specific protein labeling using protein farnesyltransferase (PFTase).[352] In this method, synthetic substrate analogs containing bioorthogonal functional groups including azides, alkynes and aldehydes are enzymatically transferred to proteins that are engineered to contain a *C*-terminal CAAX-box.[397] Protein farnesyltransferase (PFTase) is used to label a substrate protein containing a *C*-terminal tetrapeptide tag with an azide-modified isoprenoid diphosphate. The Cu-catalyzed alkyne azide cycloaddition reaction,

commonly known as the click reaction, has been used extensively for preparing modified forms of DNA.[403,404] We have previously used the PFTase method described above to prepare azide-modified proteins that were subsequently linked to the 5'-ends of alkyne-functionalized oligodeoxynucleotides *via* the Cu-catalyzed click reaction[352] or the Cu-free variation,[353] but internal DNA-protein cross-links have not been previously prepared. We elected to use the Cu-catalyzed reaction in the present study since it generates a less bulky linkage between the protein and DNA. To generate DNA-peptide conjugates, synthetic peptides were prepared *via* solid phase peptide synthesis and appended with an *N*-terminal 4-azidobutanoic acid group for subsequent Cu-catalyzed click reaction.

The new approach was successfully applied to generate site-specific DPCs to a 28.4 kDa protein, a 23-mer peptide, and a 10-mer peptide (**Schemes 5-1** and **5-2**, and **Figures 5-1**–**5-6**). Our optimized reaction conditions and purification strategy generates structurally defined DNA-protein and DNA-peptide conjugates in high yield and with excellent purity (**Figures 5-1**, **5-3** and **5-6**). We have recently employed this strategy to engineer plasmid molecules containing site-specific DPCs for *in vivo* replication studies.

The model DNA-protein and DNA-peptide conjugates generated by click reaction (**Schemes 5-1** and **5-2**) resemble DNA-protein cross-links induced by *bis*-alkylating agents[158] and reactive α,β-unsaturated carbonyls.[197] Many *bis*-electrophiles including nitrogen mustards, platinum compounds, and diepoxides form DPCs by alkylating cysteine thiols within proteins.[85,86,158,159,294] On the other hand, acrolein, crotonaldehyde, and 4-hydroxynonenal form Schiff base cross-links between DNA and the *N*-terminal α-amine of the peptide.[197] Although the linker length within our model DNA-protein

conjugates is longer than that observed for cross-links generated physiologically, we anticipate that the linker length will play a relatively minor role in determining the route of lesion processing. It is more likely that the nature of the protein/peptide and the attachment site within DNA will determine the cellular fate of DNA-protein cross-links.

Our observation of complete polymerase blockage by DNA-protein conjugates (Panels A, D and G of **Figures 5-7** and **5-8**) is consistent with an earlier finding of Kuo and collaborators, who reported that 5-azacytidine-induced methyltransferase-DNA adducts block DNA replication *in vivo*.[132] In contrast, our finding that *C5*–thymine cross-links to a 23-mer peptide block human lesion bypass polymerases κ and η (Panels B, E and H of **Figures 5-7** and **5-8**) contradict earlier reports that pol κ efficiently bypassed γ-HOPdG mediated DNA-peptide cross-links connected to the N^2 position of guanine in DNA,[335] while pol ν was able to catalyze replication past γ-HOPdA mediated DNA-peptide cross-links to the N^6 position of adenine.[336] This may be due to structural differences between the DPCs examined in these studies and also the differences in peptide size, since the previous reports were limited to peptide 4-mers and 12-mers[335,336] and did not examine the effects of larger peptide lesions on DNA replication. Indeed, our results indicate that smaller cross-links to a peptide 10-mer can be bypassed by pol κ and pol η (Panels C, F and I of **Figures 5-7** and **5-8**).

In summary, site-specific cross-links between DNA oligomers and polypeptides of increasing size were generated using copper-catalyzed [3+2] Huisgen cycloaddition (click reaction) between an alkyne group from C8-alkyne-dT in DNA and an azide group within engineered proteins/polypeptides in high yield and with excellent purity.

Polymerase bypass experiments conducted with model DPC substrates incorporating 10-mer peptide, 23-mer peptide, and a 28.4 kDa protein have shown that while the two larger lesions blocked all human polymerases tested, the DPC to a 10-mer peptide can be bypassed by polymerases η, κ, and ι. These results suggest that large DPCs generated in cells may require proteolytic processing in order to be tolerated. Our ongoing studies will identify the proteolytic mechanisms involved and elucidate the effects of proteasomal inhibitors on toxicity of common antitumor drugs that are known to form DPCs.

6 ERROR-PRONE TRANSLESION SYNTHESIS PAST DNA-PEPTIDE CROSS-LINKS CONJUGATED TO THE MAJOR GROOVE OF DNA *VIA* C5 OF THYMIDINE

6.1 INTRODUCTION

Previous investigations have revealed that the ability of translesion synthesis (TLS) polymerases to bypass DNA-peptide cross-links is dependent on the lesion size, the attachment site within DNA, and polymerase identity.[132,164,219,326,327,334-338] Lloyd and coworkers reported that human polymerase, hPol κ and its *E. coli* orthologue, Pol IV were able to catalyze error-free primer extension past DNA templates containing tetra- and dodecapeptides cross-linked to the N^2 position of guanine *via* a trimethylene linker.[335] In a different study, A family polymerase ν was capable of bypassing both 4-mer and 12-mer peptides cross-linked to N^6 of adenine *via* the same linkage, with only the correct base (dT) inserted opposite the adduct,[336] while the analogous adducts at N^2 of guanine completely blocked replication.[336] Yamanaka *et al.* proposed that small major groove DPC adducts have sufficient conformational flexibility to be accommodated within the active site of Pol ν without disturbing primer-template-enzyme interactions, while the corresponding minor groove adducts block replication.[336] More recently, Guengerich and coworkers reported that human TLS polymerases κ and ι as well as the

bacterial polymerases, Pol T7 and DPO4 were capable of replicating DNA containing *S*-[4-(N^6-deoxyadenosinyl)-2,3-dihydroxybutyl]glutathione adducts (N^6-dA-$(OH)_2$butyl-GSH).[338] However, none of the previous studies, except for the one reported by Guengerich and coworkers,[338] have sequenced the primer extension products. Despite the accurate nucleotide insertion opposite the lesion, TLS polymerases can cause insertion/deletion mutations.[394,405-407] To our knowledge, no systematic studies have been performed on replication bypass of DPCs cross-linked to pyrimidine bases on DNA, despite their significance *in vivo*. Cytosine methylation by DNA methyltransferases involves the covalent cross-linking of an active site Cys to *C6* of cytosine, and if irreversibly trapped, can have adverse effects on epigenetic mechanisms and gene expression.[408,409] Free radicals and radiation sources can also induce protein cross-linking to pyrimidine bases on DNA.[182,211,410-412]

We have recently developed a new methodology using copper-catalyzed azide-alkyne cycloaddition to generate site-specific DNA-protein and DNA-peptide conjugates cross-linked to *C5* of thymidine in DNA (Chapter 5).[388] Our initial polymerase bypass studies with substrates containing site-specific *C5*-dT cross-links to polypeptides of increasing size (10-mer, 23-mer, and 28.4 kDa protein) have revealed that large polypeptides completely block replication, whereas 10-mer peptide cross-links were bypassed by hPol η and κ.[388] This suggests that large DPC lesions may undergo proteolytic degradation to peptides, which are subsequently bypassed by translesion DNA polymerases.

In the present work, we examined the kinetics and the fidelity of translesion synthesis past model DNA-peptide cross-links containing a 10-mer peptide ($N_3(CH_2)_3CO$-EQKLISEEDL-NH_2) connected to the *C5* position of thymidine in DNA[388] using a combination of gel electrophoresis and mass spectrometry-based methodologies. Our results provide evidence for a highly error-prone nature of replication past major groove DNA-peptide adducts by hPol η and κ, which give rise to large numbers of base substitutions and frameshift mutations.

6.2 MATERIALS AND METHODS

6.2.1 Materials

C8-alkyne-dT-CE phosphoramidite, protected 2'-deoxyribonucleoside-3'-phosphoramidites (dA-CE, Ac-dC-CE, dmf-dG-CE, dT-CE), Ac-dC-CPG ABI columns, and all other reagents required for automated DNA synthesis were purchased from Glen Research (Sterling, VA). Recombinant human polymerases hPol η and κ were either purchased from Enzymax (Lexington, KY) or expressed and purified as previously described.[384,385] T4 polynucleotide kinase (T4-PNK) and *E. coli* uracil DNA glycosylase (UDG) were obtained from New England Biolabs (Beverly, MA), while γ-^{32}P ATP was purchased from Perkin-Elmer Life Sciences (Boston, MA). 40% 19:1 Acrylamide/bis solution and micro bio-spin 6 columns were purchased from Bio-Rad (Hercules, CA). The unlabeled dNTPs were obtained from Omega Bio-Tek (Norcross, GA). Illustra NAP-5 desalting columns and Sep-Pak C18 SPE cartridges were purchased from GE Healthcare (Pittsburg, PA) and Waters (Milford, MA), respectively. All other chemicals and solvents were purchased from Sigma-Aldrich (Milwaukee, WI) and were of the highest grade available.

6.2.2 Synthesis and Characterization of Oligodeoxynucleotides

Synthetic 18-mer oligodeoxynucleotides (5'-TCA T***X***G AAT CCT TCC CCC-3') containing native thymidine at position ***X*** were synthesized by automated solid phase synthesis using an ABI 394 DNA synthesizer (Applied Biosystems, CA). 18-mer strands

containing a 5-(octa-1,7-diynyl)-uracil (C8-alkyne-dT) modification were prepared by solid phase synthesis. The modified nucleotide was added using an offline manual coupling protocol. Biotinylated 23-mer primer (Biotin-5'-$(T)_{10}$ GGG GGA AGG A**U**T C-3') and 13mer primer (5'-GGG GGA AGG ATT C-3') were purchased from Integrated DNA Technologies (Carolville, IA). All oligodeoxynucleotides were purified by semi-preparative HPLC, desalted by Illustra NAP-5 columns, characterized by HPLC-ESI^--MS, and quantified by UV spectrophotometry as previously described.[388]

6.2.3 Synthesis and Characterization of DNA-Peptide Conjugates

Synthetic DNA 18-mer (5'-TCA T***X***G AAT CCT TCC CCC-3') containing C8-dT-alkyne at position ***X*** was cross-linked to the 10-mer azide-containing peptide ($N_3(CH_2)_3$CO-Glu-Gln-Lys-Leu-Ile-Ser-Glu-Glu-Asp-Leu-NH_2) *via* copper catalyzed [3+2] Huisgen cycloaddition (click reaction).[388] Site-specific DNA-peptide conjugates were isolated using 20% (w/v) denaturing polyacrylamide gels containing 7 M urea followed by gel-elution and desalting *via* Sep-Pak C18 SPE. To assess the purity of the isolated conjugates, an aliquot of the purified sample was radiolabeled with γ-^{32}P ATP, and resolved on a 20% (w/v) denaturing polyacrylamide gel containing 7 M urea, followed by visualization using a Typhoon FLA 7000 phosphorimager (GE Healthcare, Pittsburgh, PA). A separate aliquot was subjected to alkaline phosphatase/ phosphodiesterase digestion, and the resulting nucleoside-peptide conjugates were characterized by nanoLC-nanospray-MS/MS.[388]

6.2.4 Preparation of Primer-Template Duplexes

For single nucleotide insertion assays, 13-mer DNA primers (5'-GGG GGA AGG ATT C-3', 1 nmol) were radiolabeled by incubating with T4 PNK (20 U) and γ-^{32}P ATP (20 μCi) in the presence of T4 PNK reaction buffer (total volume = 20 μL) at 37 °C for 1 h. The mixture was heated at 65 °C for 10 min in order to inactivate the enzyme and passed through Illustra Microspin G25 columns (GE Healthcare, Pittsburgh, PA) to remove excess γ-^{32}P ATP. The 5'-^{32}P-labeled primers (50 pmol) were mixed with 2 eq. of template strands (5'-TCA T***X***G AAT CCT TCC CCC-3') containing either 2'-deoxythymidine (dT) or 10-mer peptide cross-link (dT-peptide) in 10 mM Tris buffer (pH 8) containing 50 mM NaCl. The primer and template strands were annealed by heating at 95 °C for 10 min and cooled slowly overnight to afford the corresponding end-labeled primer-template complexes.

Primer-template duplexes used for the capillary HPLC-ESI$^-$-MS/MS sequencing experiments were generated analogously by annealing biotinylated 23-mer primer (Biotin-5'-$(T)_{10}$ GGG GGA AGG AUT C-3', 100 pmol) to template strands (5'-TCA T***X***G AAT CCT TCC CCC-3') containing either unmodified dT or dT-peptide cross-link at position ***X*** (2 eq.).

6.2.5 Single Nucleotide Incorporation Assays

Single nucleotide insertion assays were conducted in order to determine which nucleotides are incorporated opposite the DNA-peptide cross-links, ***X***. ^{32}P-end-labeled

primer-template duplexes containing either dT or dT-peptide (40 nM) were incubated at 37 °C with human translesion synthesis polymerases (20 nM hPol η or 200 nM hPol κ) in 50 mM Tris-HCl (pH 7.5) buffer containing 50 mM NaCl, 5 mM DTT, 5 mM $MgCl_2$, 100 μg/mL BSA and 10% glycerol (v/v). Reactions were initiated by the addition of individual dNTPs (50 μM for hPol η and 100 μM for hPol κ) in a final volume of 20 μL. Aliquots (4 μL) were withdrawn at pre-selected time points, and the reaction was quenched by the addition of 8 μL of the quench solution containing 10 mM EDTA, 0.03% bromophenol blue (w/v) and 0.03% xylene cyanol (w/v) in 95% (v/v) formamide. The extension products were resolved by 20% (w/v) denaturing PAGE containing 7 M urea and visualized using a Typhoon FLA 7000 phosphorimager (GE Healthcare, Pittsburgh, PA).

6.2.6 Steady-State Kinetic Analyses

Steady-state kinetics for incorporation of individual nucleotides opposite the unadducted dT and dT-peptide cross-link was investigated by performing single nucleotide incorporation assays. Experiments were conducted with 0.25–1.25 nM hPol η or 2.5–100 nM hPol κ in the presence of increasing concentrations of individual dNTPs (0–800 μM) and for specified time periods (0–180 min). The extension products and the unextended 13-mer primers were visualized after electrophoretic fractionation using a Typhoon FLA 7000 phosphorimager (GE Healthcare, Pittsburgh, PA), and quantified by volume analysis using the Image Quant TL 8.0 software (GE Healthcare, Pittsburgh, PA).

The steady-state kinetic parameters were calculated by nonlinear regression analysis using one-site hyperbolic fits in Prism 4.0 (GraphPad Software, La Jolla, CA).

6.2.7 Sequencing and Quantitation of Primer Extension Products by Liquid Chromatography-Tandem Mass Spectrometry

Primer extension assays were conducted by incubating the biotinylated 23-mer primer-18-mer template duplexes (100–150 pmol) with hpol κ or η (40–60 pmol) in 50 mM Tris-HCl (pH 7.5) buffer containing 50 mM NaCl, 5 mM DTT, 5 mM $MgCl_2$, 100 μg/mL BSA and 1 mM each of the four dNTPs at 37 °C for 6 h. Following the polymerase reaction, 400 μL of 20 mM sodium phosphate (pH 7.0) buffer containing 150 mM NaCl were added to the reaction. Streptavidin sepharose high performance beads (0.2 mL, GE Healthcare, Pittsburgh, PA) were prepared by centrifugation and washing with 500 μL of 20 mM sodium phosphate (pH 7.0) buffer containing 150 mM NaCl. The beads were added to the polymerase reaction, and the resulting suspension was incubated at room temperature for 2 h with tapping every 10 min to promote mixing. The liquid was removed by centrifugation, and the beads were washed three times with water (300 μL). A solution of Tris-HCl buffer (50 mM, pH 7.5) containing uracil DNA glycosylase (UDG, 20 units), 1 mM EDTA, and 1 mM DTT (500 μL) was added to the streptavidin-coated bead suspension, and the mixture was incubated for 4 h at 37 °C. The liquid was removed, and the beads were washed with water (3× 300 μL). An aqueous solution of 250 mM piperidine (400 μL) was added to the streptavidin-coated beads, and the mixture was heated at 95 °C for 60 min. The liquid from the piperidine cleavage was

decanted, and the beads were washed with water (3× 200 μL). The piperidine cleavage fraction was combined with the water washes and dried *in vacuo*, and the residue was reconstituted in 25 μL of water containing 14-mer internal standard (5'-pCTT CAC GAG CCC CC-3', 40 pmol).

Capillary HPLC-ESI$^-$-MS/MS analyses were conducted on an Agilent 1100 HPLC system (Agilent Technologies, Wilmington, DE) coupled to a Thermo LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Waltham, MA). The instrument was operated in the negative ion ESI MS/MS mode. Primer extension products were resolved on an Agilent Zorbax 300SB-C18 (0.5 × 150 mm, 5 μm) column using a gradient of 15 mM ammonium acetate (buffer A) and acetonitrile (buffer B). The column was eluted at a flow rate of 15 μL/min. The solvent composition was linearly changed from 1 to 10% B in 24 min, further to 75% B in 1 min, held at 75% B for 3 min, and brought back to 1% B in 2 min.

Electrospray ionization conditions were as follows: ESI source voltage, 3.5 kV; source current, 6.7 A; auxiliary gas flow rate setting, 0; sweep gas flow rate setting, 0; sheath gas flow setting, 30; capillary temperature, 275 °C; and S-lens RF level, 50%. The most abundant ions from the ESI$^-$-FTMS spectra were selected and subjected to collision-induced dissociation (CID) analysis using a linear ion trap. The MS/MS conditions were as follows: normalized collision energy, 35%; activation Q, 0.250; activation time, 10 ms; product ion scan range, *m/z* 300–2000. Relative quantitation of primer extension products was done by comparing peak areas corresponding to each product in extracted ion chromatograms with the peak area of the internal standard. Product sequences were

confirmed by comparing the MS/MS fragments to expected CID fragmentation patterns of oligodeoxynucleotides obtained using the Mongo Oligo mass calculator version 2.06 (The RNA Institute, College of Arts and Sciences, State University of New York at Albany).

6.3 RESULTS

6.3.1 Synthesis of Primer-Template Duplexes Containing Site-Specific DNA-Peptide Conjugates

Synthetic DNA strands containing site specific DNA-peptide cross-links were generated by copper-catalyzed [3+2] Huisgen cycloaddition between the 5-(octa-1,7-diynyl)-uracil (C8-alkyne-dT) base (***X***) positioned within DNA 18-mer (5'-TCA T***X***G AAT CCT TCC CCC-3') and the *N*-terminal azide moiety appended to a 10-mer c-Myc peptide ($N_3(CH_2)_3CO$-EQKLISEEDL-NH_2) (**Scheme 6-1**).[388] DNA-peptide cross-links (**Scheme 6-1**) were isolated by denaturing PAGE, and their purity was confirmed by sequencing PAGE (**Figure 6-1**). Only DNA templates that were at least 99% pure were employed for primer extension assays. Purified DNA-peptide cross-links were characterized by nanoLC-nanospray-MS/MS analysis to confirm the structure of the conjugate[388] and quantified by UV spectrophotometry assuming $\varepsilon_{260\ nm}$ (DNA-peptide cross-link) = $\varepsilon_{260\ nm}$ (DNA) + $\varepsilon_{260\ nm}$ (peptide) = 220.08 $mM^{-1}\ cm^{-1}$.

Scheme 6-1 Synthesis of DNA-peptide cross-links by copper-catalyzed [3+2] Huisgen cycloaddition (click reaction).[388]

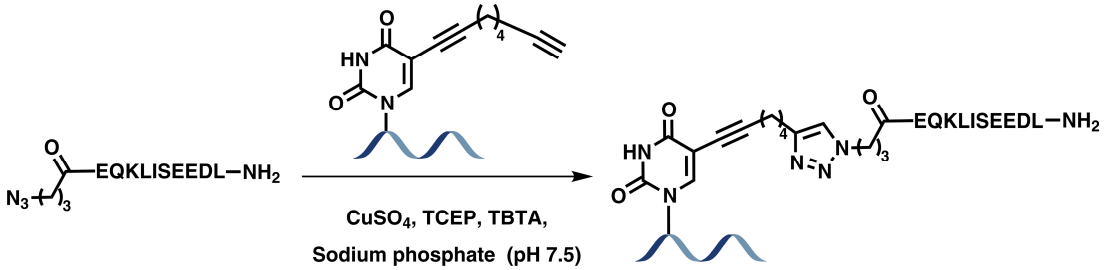
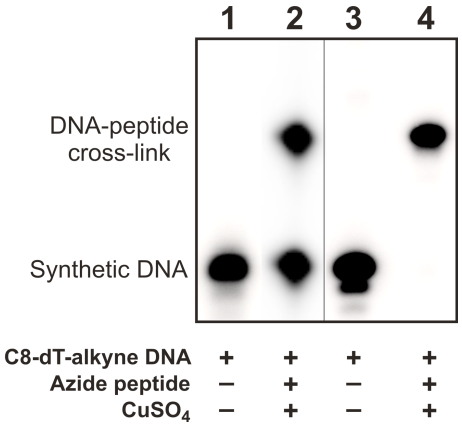


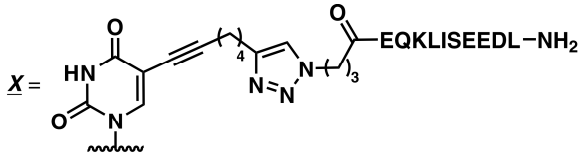
Figure 6-1 Denaturing PAGE analysis of DNA-peptide cross-links generated by click chemistry. Lane 1: oligodeoxynucleotide containing C8-alkyne-dT; Lane 2: oligodeoxynucleotide containing C8-alkyne-dT and peptide azide in the presence of Cu catalyst; Lane 3: oligodeoxynucleotide containing C8-alkyne-dT; Lane 4: gel purified DNA-peptide cross-links from lane 2.



6.3.2 Single Nucleotide Incorporation Assays

Lesion-containing DNA 18-mers (5'-TCA T***X***G AAT CCT TCC CCC-3') were annealed to a ^{32}P-end-labeled 13-mer primer, 5'-GGG GGA AGG ATT C-3'. In the resulting primer-template complexes, the 3'-primer terminus is positioned immediately prior to the lesion site (-1 primer, **Scheme 6-2**A). Initial single nucleotide incorporation assays were conducted to determine what nucleotides can be inserted opposite the adducted site by human translesion synthesis polymerases, hPol η and κ. Those enzymes were selected for the current study because these polymerases are known to catalyze replication past other DNA-peptide cross-links.[335,338] The resulting primer-template duplexes were incubated with recombinant polymerases in the presence of individual dNTPs for 0–90 min. Denaturing PAGE followed by phosphorimaging analysis revealed that any of the four dNTPs can be incorporated, indicative of error-prone replication of DNA containing dT-peptide conjugates (**Figure 6-2**). Under our experimental conditions (substrate:enzyme = 2:1), hPol η showed full incorporation opposite the DPC lesion within 5 min (**Figure 6-2**A). In contrast, hPol κ (substrate:enzyme = 1:5) only showed <60% incorporation of all dNTPs even after 90 min (**Figure 6-2**B). These initial experiments suggested that hPol η has a higher efficiency of replication past DPCs as compared to hPol κ.

Scheme 6-2 Sequences of DNA oligomers employed in *in vitro* replication studies. (A) 13-mer primer-18-mer template duplexes used for the single nucleotide incorporation and steady-state kinetic assays. (B) Biotinylated 23-mer primer-18-mer template duplexes employed for the liquid chromatography-tandem mass spectrometry analyses of primer extension products.



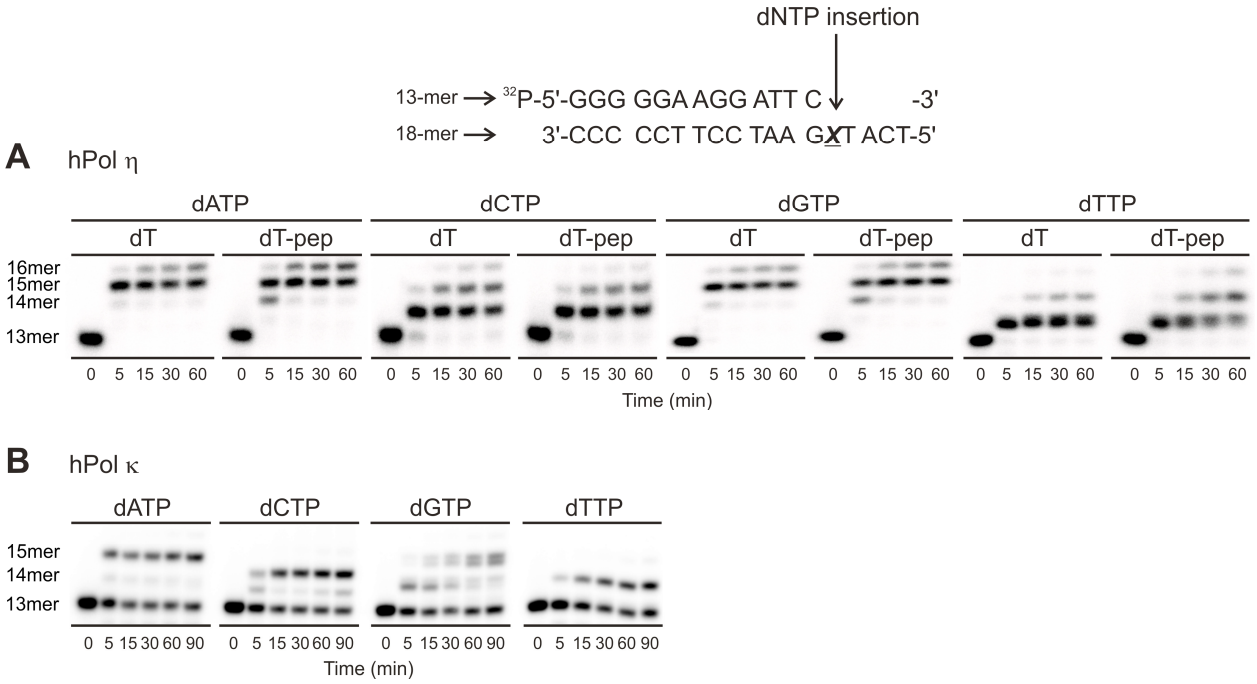
A ^{32}P-5'-GGG GGA AGG ATT C -3' (5'-end-labeled 13-mer primer)
3'-CCC CCT TCC TAA G***T***T ACT -5' (18-mer template with positive control)

^{32}P-5'-GGG GGA AGG ATT C -3' (5'-end-labeled 13-mer primer)
3'-CCC CCT TCC TAA G***X***T ACT -5' (18-mer template with dT-peptide)

B Biotin-5'-$(T)_{10}$GGG GGA AGG A***U***T C -3' (Biotinylated 23-mer primer with uridine)
3'-CCC CCT TCC TAA G***T***T ACT-5' (18-mer template with positive control)

Biotin-5'-$(T)_{10}$GGG GGA AGG A***U***T C -3' (Biotinylated 23-mer primer with uridine)
3'-CCC CCT TCC TAA G***X***T ACT-5' (18-mer template with dT-peptide)

Figure 6-2 Single nucleotide insertion opposite DNA-peptide cross-links by human TLS polymerases. (A) 40 nM primer-template duplexes containing unmodified thymidine (dT control) and the 10-mer peptide cross-linked to *C5* of thymidine (dT-peptide) were incubated at 37 °C with 20 nM hPol η in the presence 50 μM individual dNTPs. The reactions were quenched at predetermined time points (0–60 min) and analyzed by 20% denaturing PAGE. (B) 40 nM primer-template duplexes containing dT-peptide were incubated at 37 °C with 200 nM hPol κ in the presence 100 μM individual dNTPs. The reactions were quenched at predetermined time points (0–90 min) and analyzed by 20% denaturing PAGE.



6.3.3 Steady-State Kinetic Analyses

To determine the catalytic efficiency and the misinsertion frequency of individual dNTPs by hPol η and κ, steady-state kinetic assays were conducted. Primer-template complexes containing unmodified dT (positive control) or dT-peptide conjugate (***X*** in **Scheme 6-2**A) were incubated with hPol η or hPol κ in the presence of increasing concentrations of individual dNTPs (0–800 μM), and the reactions were quenched at specified time points (0–180 min). Conditions were chosen such that the maximum percentage of products was ≤35% of the starting substrate concentration. Kinetic parameters (k_{cat} and K_m in **Table 6-1**) were calculated by plotting reaction velocity against concentrations of individual dNTP.[369] The specificity constant (k_{cat}/K_m) is a measure of the catalytic efficiency of incorporation of each dNTP, while the misinsertion frequency (*f*) provides a quantitative measure of incorporating an incorrect vs. correct dNTP opposite the lesion, hence the fidelity of the polymerase for a specific lesion.[1]

Table 6-1 Steady-state kinetic parameters for single nucleotide insertion opposite the positive control (dT) and the 10-mer peptide cross-linked to *C5* of thymidine (dT-peptide) by human TLS polymerases, hPol κ and η.

Polymerase	***Template***	***Incoming nucleotide***	k_{cat} min^{-1}	K_m μM	k_{cat}/K_m $\mu M^{-1}\ min^{-1}$	f ‡‡
hPol κ	dT	dATP	2.75 ± 0.18	5.11 ± 1.96	0.54	1
		dCTP	0.19 ± 0.02	145.6 ± 48.55	0.0013	0.002
		dGTP	1.93 ± 0.22	94.84 ± 31.27	0.02	0.04
		dTTP	0.51 ± 0.07	371.8 ± 98.51	0.0014	0.003
	dT-peptide	dATP	0.67 ± 0.04	16.61 ± 6.51	0.04	1
		dCTP	0.07 ± 0.007	90.81 ± 42.90	0.0007	0.02
		dGTP	0.16 ± 0.008	152.5 ± 25.61	0.001	0.03
		dTTP	0.03 ± 0.003	361.8 ± 84.11	0.00008	0.002
hPol η	dT	dATP	9.44 ± 0.26	2.85 ± 0.70	3.31	1
		dCTP	2.26 ± 0.52	72.60 ± 54.38	0.031	0.009
		dGTP	9.00 ± 1.67	13.36 ± 8.22	0.67	0.2
		dTTP	1.04 ± 0.16	58.82 ± 27.89	0.018	0.005
	dT-peptide	dATP	2.81 ± 0.37	31.16 ± 22.91	0.09	1
		dCTP	2.67 ± 0.36	65.85 ± 29.08	0.041	0.45
		dGTP	13.44 ± 1.91	33.42 ± 11.60	0.402	4.5
		dTTP	1.21 ± 0.14	32.60 ± 16.55	0.037	0.41

‡‡ f = *misinsertion frequency* = $(k_{cat}/K_m)_{incorrect\ dNTP}/(k_{cat}/K_m)_{correct\ dNTP}$

According to our steady-state kinetics data (**Table 6-1**), the specificity constants (k_{cat}/K_m) for the incorporation of the correct base (dA) opposite the dT-peptide conjugate were 0.04 and 0.09 μM^{-1} min^{-1} for hPol κ and η, respectively. These values were 14–37-fold lower than those obtained for dATP insertion opposite the unmodified template (0.54 and 3.31 μM^{-1} min^{-1}). This is not unexpected, since it may be difficult to accommodate the bulky DPC adduct in the polymerase active site, and this process is likely to be both kinetically and thermodynamically disfavored. The specificity constants for the incorporation of the other three nucleotides by hPol κ were also 2–20-fold lower for the adducted template compared to the unmodified template (**Table 6-1**). hPol η also showed higher specificity constants for the incorporation of incorrect bases (dCTP and dTTP) opposite the adducted template as compared to hPol κ (1.3- and 2-fold, respectively). The calculated k_{cat}/K_m values for hPol η were larger than those for hPol κ (**Table 6-1**) suggesting that hPol η is more efficient in the replication bypass of the dT-peptide cross-link as compared to hPol κ.

The misinsertion frequencies (*f*) calculated from kinetic data (**Table 6-1**) suggest that both hPol κ and η bypass dT-peptide conjugates in an error-prone manner. For hPol κ, the frequency of insertion of dCTP opposite the adducted template was 10-fold higher as compared to the unmodified template, while *f* was only 0.7-fold for incorporation of dGTP and dTTP. In contrast, hPol η showed 50-fold, 23-fold and 82-fold higher misinsertion frequencies for dCTP, dGTP and dTTP, respectively. Further, the order of nucleotide insertion opposite the adducted template was A > G > C > T for hPol κ, with 33-fold, 50-fold and 500-fold preference for the incorporation of dATP over dGTP, dCTP and dTTP, respectively. The order of misinsertion opposite the adducted base by hPol η was G > A > C > T with 4.5-fold, 10-fold and 11-fold

preference for the incorporation of dGTP over dATP, dCTP and dTTP. These data suggest that replication bypass of the DNA-peptide conjugates by hPol η is much more error-prone as compared to bypass of the same adduct by hPol κ.

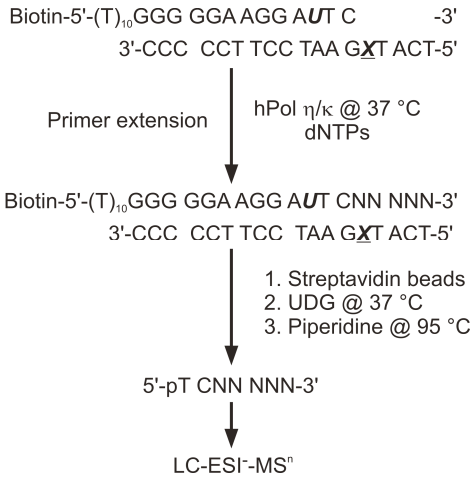
6.3.4 Sequencing and Quantitation of Primer Extension Products by Liquid Chromatography-Tandem Mass Spectrometry

A mass spectrometry-based approach (**Scheme 6-3**), similar to the one developed by Christov *et al.*,[413] was employed to sequence the products of translesion synthesis. This methodology allows for the detection of insertion and/or deletion products and potential errors at the adducted site and the +1 position from the adduct.[407] A biotinylated primer (Biotin-5'-$(T)_{10}$ GGG GGA AGG A**U**T C-3') was annealed to the 18-mer template containing the site-specific DPC at position ***X*** (**Scheme 6-2**B). The biotin group on the 5'-end of the oligodeoxynucleotide was used to purify the extension products *via* affinity capture prior to LC-MS analysis.[413] The third nucleotide from the 3'-end of the template was replaced with a uridine to allow for selective cleavage of the extended primer using uracil DNA glycosylase (UDG) and hot piperidine. This strategy generates short oligodeoxynucleotides (7-mers, if fully extended, or 5- or 6-mers, if deletion mutations occured) that are more amenable for sequencing by collision induced dissociation-based tandem mass spectrometry.[407,414]

As was the case for gel-based assays, primer extension reactions were carried out with unmodified and dT-peptide containing templates (**Scheme 6-2**B). The extended primers were captured on streptavidin high performance sepharose beads, and the beads were incubated with UDG. The 3'-end fragments containing the extension products of interest were cleaved by

treatment with hot piperidine (**Scheme 6-3**), and the eluates were dried *in vacuo* and analyzed by HPLC-ESI$^-$-FTMS and MS/MS on an Orbitrap Velos mass spectrometer. The samples were initially analyzed in the full scan mode to detect all extension products. HPLC-ESI$^-$-FTMS peak areas in extracted ion chromatograms (**Figures 6-4** and **6-6**) were used to determine the relative quantities of each extension product as compared to an internal standard. Each sample was subsequently analyzed in the HPLC-ESI$^-$-MS/MS mode to allow for nucleotide sequence determination (**Figures 6-3**, **6-5**, **6-7** and **6-8**).

Scheme 6-3 Streptavidin capture in combination with capillary HPLC-ESI$^-$-MS/MS method for sequencing and quantitation of primer extension products.



Capillary HPLC-ESI$^-$-FTMS analysis of the hPol κ-catalyzed reactions with unmodified template revealed one main peak corresponding to the error-free full extension product (5'-pT C***A***A TGA-3'; *m/z* 1090.68; $[M-2H]^{2-}$, **Table 6-2**, **Figures 6-3**A and **6-4**A). In contrast, HPLC-ESI$^-$-FTMS of the corresponding primer extension reactions conducted with dT-peptide conjugate revealed eight different extension products (**Table 6-2** and **Figure 6-4**B). The main peak (89% of total products) corresponded to a single base deletion product (5'-pT C_A TGA-3'; *m/z* 934.15; $[M-2H]^{2-}$), and only 2% of the primer extension products yielded the error-free replication product (5'-pT C***A***A TGA-3'; *m/z* 1090.68; $[M-2H]^{2-}$). Other major MS peaks at *m/z* 942.15, 777.62, 1086.17 and 1098.68 corresponded to doubly charged ions of 5'-pT C_G TGA-3' (one-base deletion, 5%), 5'-pT C_ _ TGA-3' (two-base deletion, 2%), 5'-pT C***T***A TGA-3' (transversion mutation, 1%) and 5'-pT C***G***A TGA-3' (transition mutation, 1%). The exact sequences were determined using the MS/MS spectra obtained by collision-induced dissociation of the corresponding doubly charged ions (**Figures 6-3**B and **6-5**).

Table 6-2 Summary of extension products formed by hPol κ and η as identified by liquid chromatography-tandem mass spectrometry.

Biotin-5'-$(T)_{10}$ GGG GGA AGG A***U***T C -3'
3'-CCC CCT TCC TAA G***X***T ACT-5'

Polymerase	***X***	***Extension product***	***Percent product***	***Base opposite X***	***Comment***
hPol κ	dT	T C***A***A TGA	87	A	error-free
		T C***T***A TGA	9	T	substitution opposite adduct
		T C***G***A TGA	3	G	substitution opposite adduct
		T C***A*** _ TGA	1	A	one-base deletion
		T C***C***A TGA	<1	C	substitution opposite adduct
		T C _ G TGA	<1	G	one-base deletion
	dT-peptide	T C***A*** _ TGA	89	A	one-base deletion
		T C _ G TGA	5	G	one-base deletion
		T C***A***A TGA	2	A	error-free
		T C_ _ TGA	2	-	two-base deletion
		T C***T***A TGA	1	T	substitution opposite adduct
		T C***G***A TGA	1	G	substitution opposite adduct
		T C _ T TGA	<1	T	one-base deletion
		T C _ C TGA	<1	C	one-base deletion
hPol η	dT	T C***A***G TGA	36	A	substitution opposite +1-mer
		T C***A***C TGA	36	A	substitution opposite +1-mer
		T C***A***A TGA	28	A	error-free
	dT-peptide	T C***A***G TGA	23	A	substitution opposite +1-mer
		T C***A***A TGA	19	A	error-free
		T C _ G TGA	16	G	one-base deletion
		T C***A***C TGA	15	A	substitution opposite +1-mer
		T C***A*** _ TGA	15	A	one-base deletion
		T C _ C TGA	12	C	one-base deletion
		T C_ _ TGA	<1	-	two-base deletion

Figure 6-3 Collision-induced dissociation (CID) mass spectra of the major extension products of *in vitro* polymerase bypass by hPol κ. (A) MS/MS spectrum of error-free product, 5'-pT CAA TGA-3', observed by replication of the unadducted template. (B) Tandem mass spectrum of one-base deletion product, 5'-pT C_A TGA-3', detected when the dT-peptide containing template was bypassed.

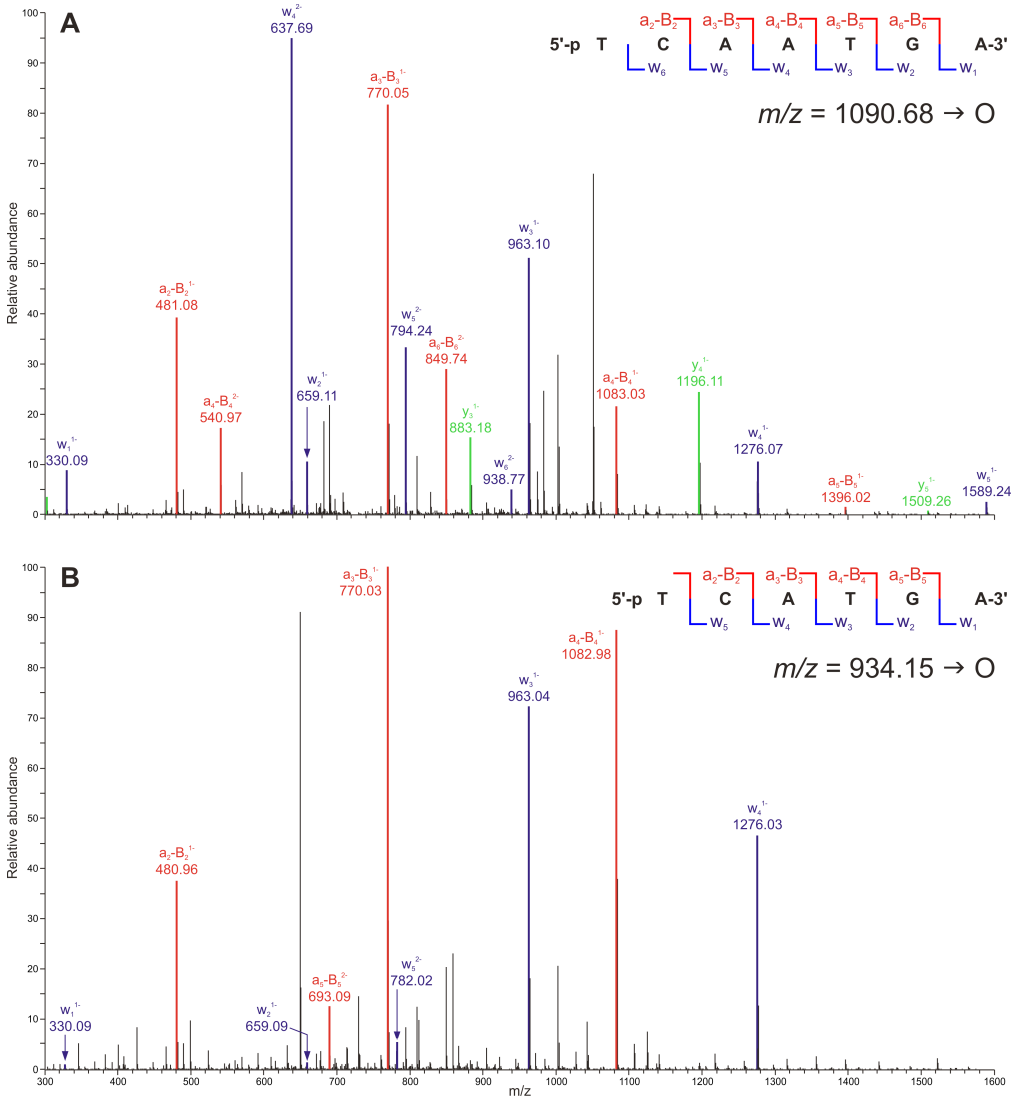


Figure 6-4 HPLC-ESI$^-$-FTMS analysis of primer extension by hPol κ. (A) Extracted ion chromatograms of replication products observed upon primer extension of the unmodified template. The major extension product, 5'-pT CAA TGA-3' detected (87%) was the error-free full extension product. (B) Extracted ion chromatograms of bypass products observed upon primer extension of the template containing 10-mer peptide conjugate. Multiple deletion mutations and point mutations were observed. The major extension product, 5'-pT C_A TGA-3' detected (89%) was a frameshift mutation.

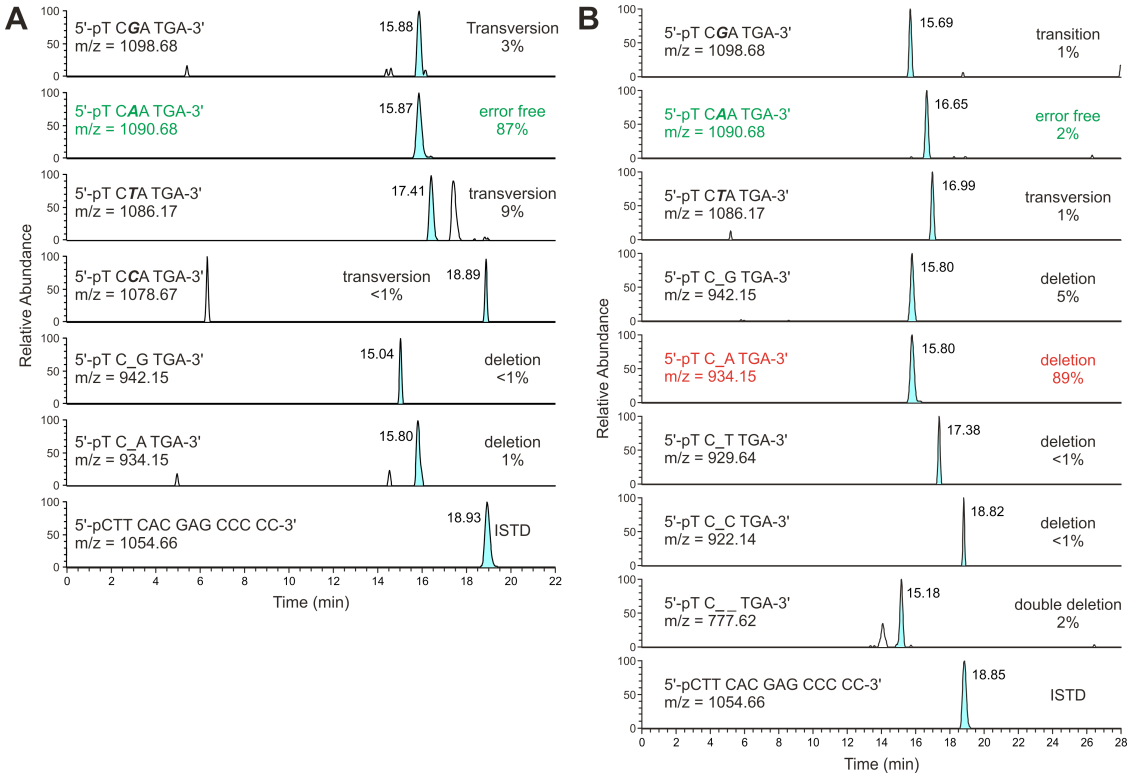
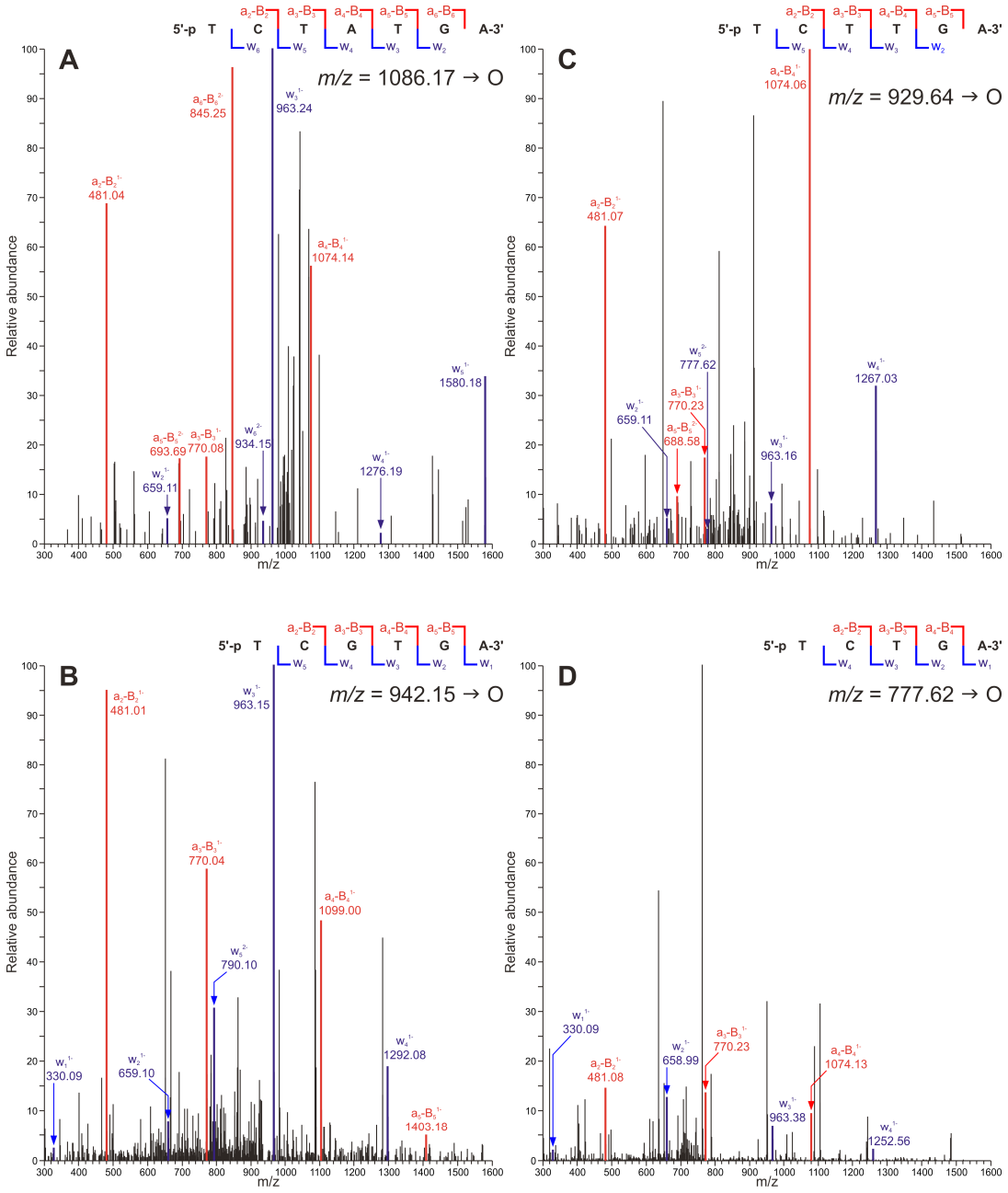


Figure 6-5 Representative MS/MS spectra of extension products observed following *in vitro* replication past 10-mer peptide cross-linked to *C5* of thymidine by hPol κ. CID spectra of extension products: (A) 5'-pT CTA TGA-3', a substitution mutation, (B) 5'-pT C_G TGA-3' and (C) 5'-pT C_T TGA-3', one-base deletion followed by a substitution, and (D) 5'-pT C_ _ TGA-3', double deletion.



HPLC-ESI$^-$-FTMS analysis of the primer extension of the undamaged template by hPol η revealed three main peaks present in roughly equal amounts at *m/z* 1090.68 (28%), 1098.68 (36%) and 1078.67 (36%), which correspond to the doubly charged ions of 5'-pT C***A***A TGA-3', 5'-pT C***A***G TGA-3' and 5'-pT C***A***C TGA-3', respectively (**Table 6-2** and **Figure 6-6**A). MS/MS sequencing confirmed that hPol η inserts the correct base, dA opposite the unmodified T, but insertion of the succeeding nucleotide is error-prone (**Figure 6-7**). When the primer extension reaction corresponding to the adducted template was analyzed, seven different products were observed (**Table 6-2** and **Figure 6-6**B). 19% of the total extension products corresponded to the error-free full extension product (5'-pT C***A***A TGA-3'; *m/z* 1090.68; $[M-2H]^{2-}$), while 23% and 15% were G and C substitutions at the +1 position (5'-pT C***A***G TGA-3'; *m/z* 1098.68; $[M-2H]^{2-}$ and 5'-pT C***A***C TGA-3'; *m/z* 1078.67; $[M-2H]^{2-}$, respectively, **Table 6-2**, **Figures 6-6**B and **6-8**). Other major MS peaks at *m/z* 942.15, 934.15 and 922.14 corresponded to the doubly charged ions of 5'-pT C_G TGA-3' (one-base deletion, 16%), 5'-pT C_A TGA-3' (one-base deletion, 15%) and 5'-pT C_C TGA-3' (one-base deletion, 12%, **Table 6-2**, **Figures 6-6**B and **6-8**). These results suggest that replication bypass of the dT-peptide cross-link by TLS polymerases is highly error-prone and that the extension products vary dramatically with the polymerase involved in replication of the adducted template. The deletion mutations observed are consistent with the fact that these TLS polymerases can skip the adducted site during replication and continue the polymerization reaction one base after the bulky adducted site.[394,405-407]

Figure 6-6 HPLC-ESI$^-$-FTMS analysis of primer extension by hPol η. (A) Extracted ion chromatograms of bypass products observed upon primer extension of the unmodified template. Three major peaks of roughly equal peak areas were observed. (B) Extracted ion chromatograms of replication products observed upon primer extension of the template containing 10-mer peptide conjugate. Multiple substitution and deletion products were observed. Error-free extension product, 5'-pT CAA TGA-3', constituted only 19% of total products.

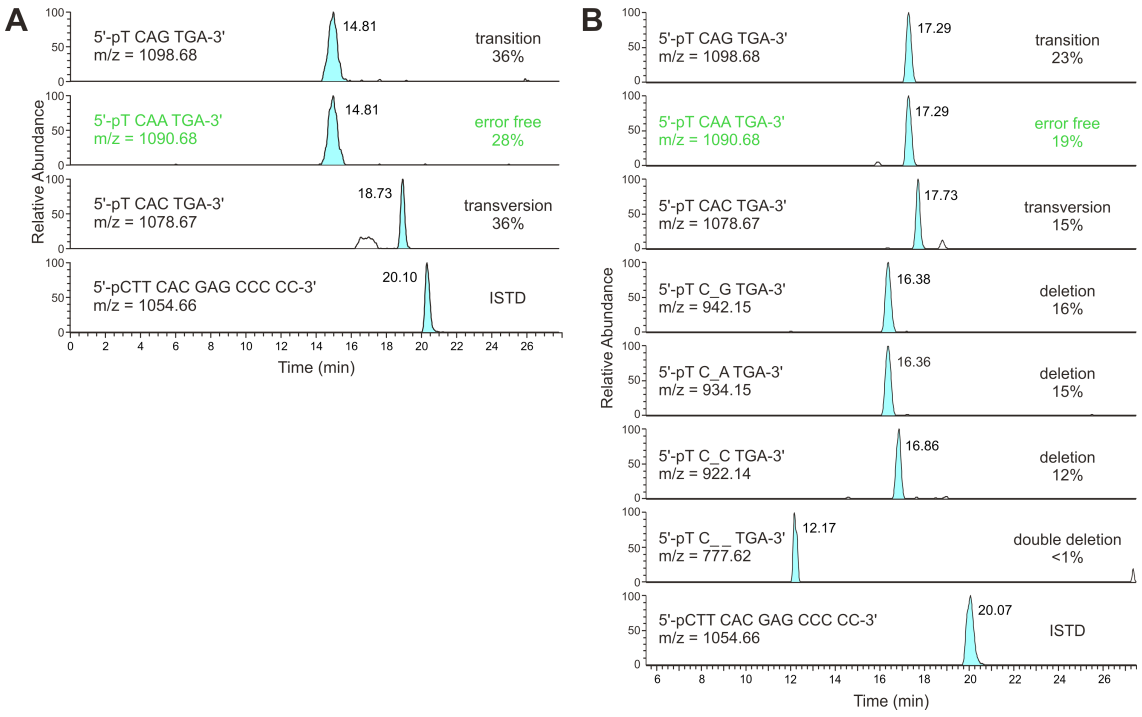


Figure 6-7 Representative MS/MS spectra of extension products observed following *in vitro* replication past unmodified template by hPol η. CID spectra of extension products: (A) 5'-pT CAA TGA-3', error-free product, and (B) 5'-pT CAG TGA-3' and (C) 5'-pT CAC TGA-3', substitutions opposite +1-mer.

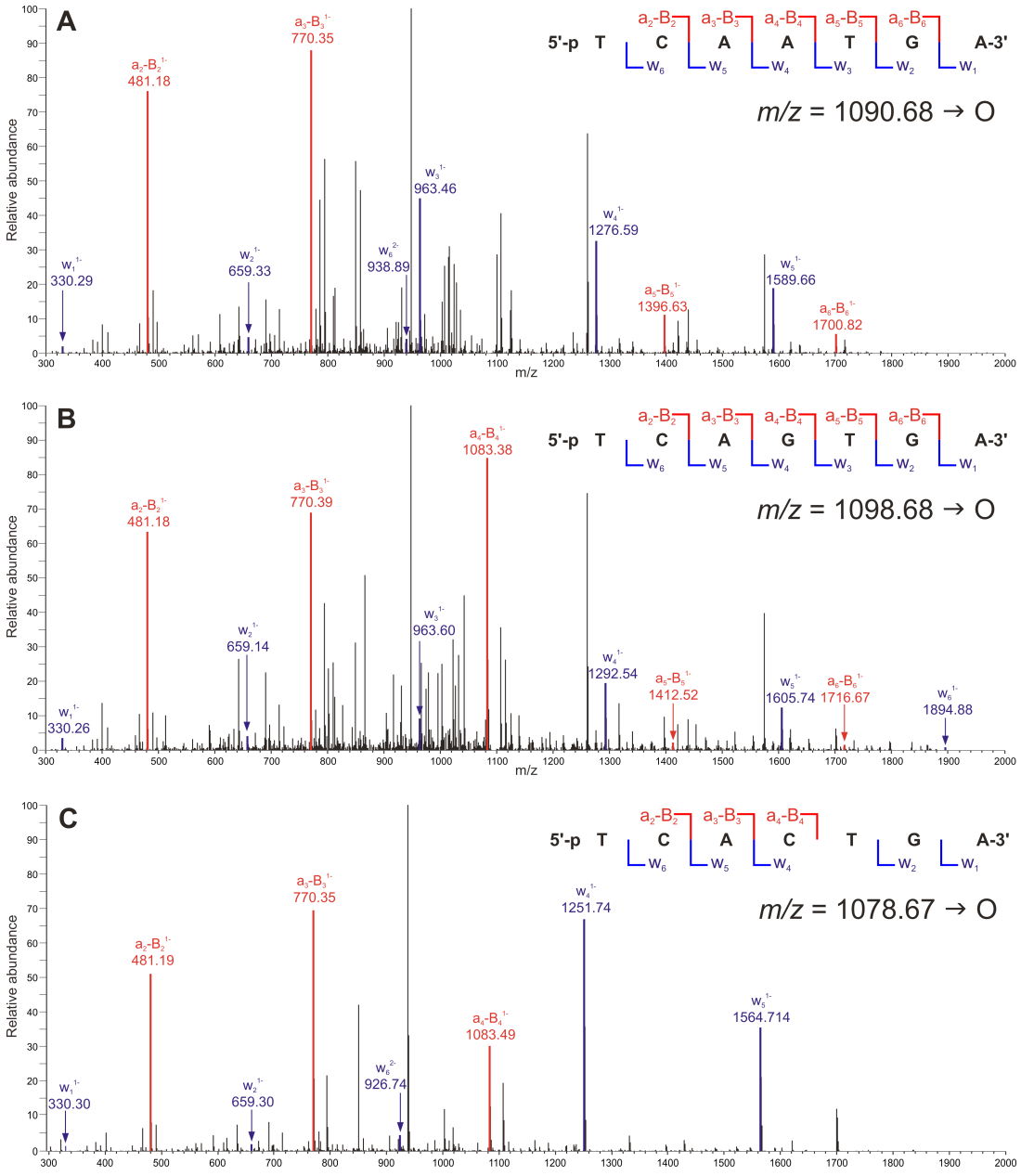
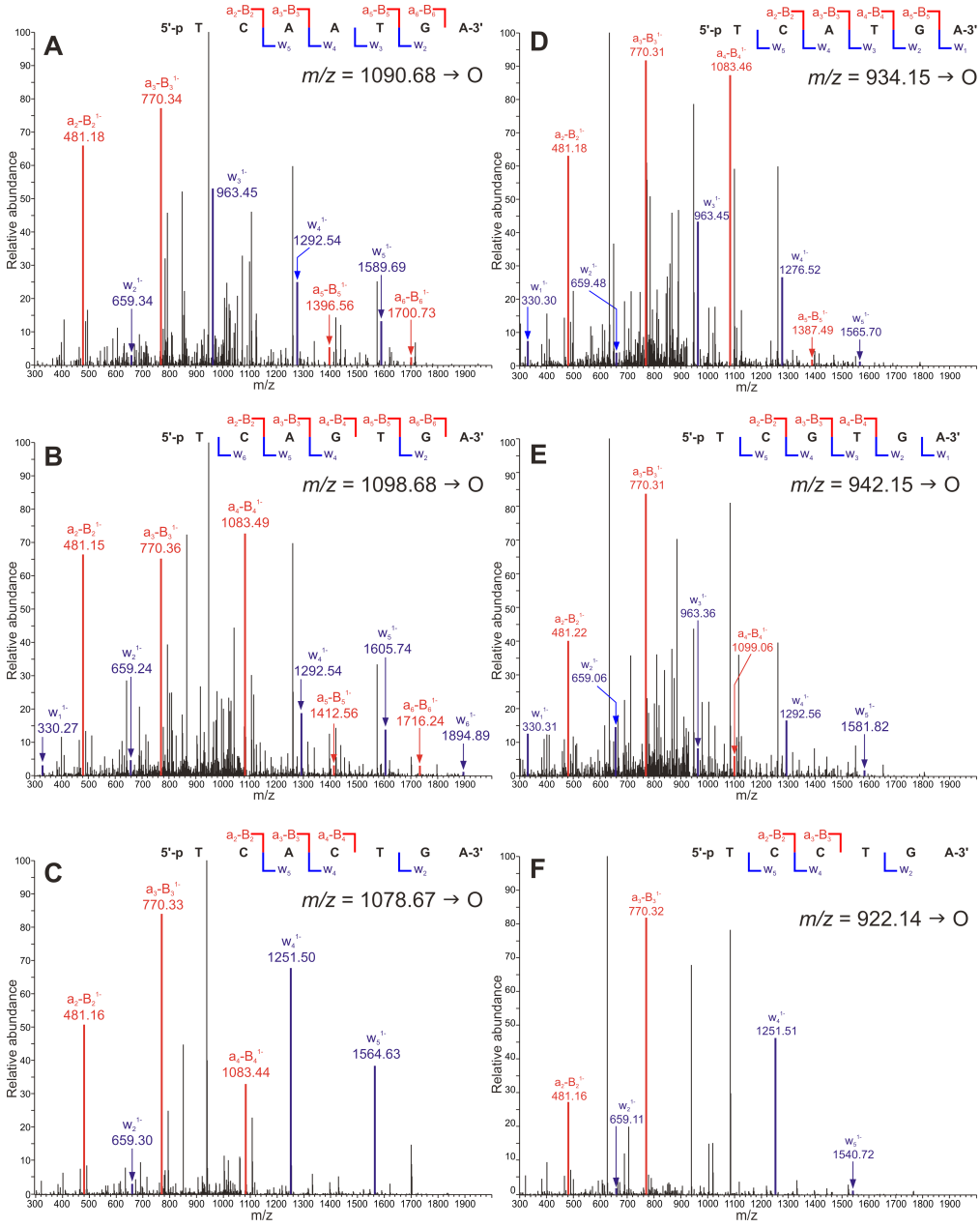


Figure 6-8 Representative MS/MS spectra of extension products observed following *in vitro* replication past 10-mer peptide cross-linked to *C5*-T by hPol η. CID spectra of extension products: (A) 5'-pT CAA TGA-3', error-free product, (B) 5'-pT CAG TGA-3' and (C) 5'-pT CAC TGA-3', substitutions opposite +1-mer, (D) 5'-pT C_A TGA-3', one-base deletion, and (E) 5'-pT C_G TGA-3' and (F) 5'-pT C_C TGA-3', one-base deletion followed by a substitution.



6.4 DISCUSSION

Because of their large size compared with other DNA lesions, DPCs are likely to block DNA replication. It has been proposed that proteins cross-linked to DNA can be proteolytically processed to generate less bulky DNA-peptide cross-links, which can be subsequently bypassed by TLS polymerases in an error-free or error-prone manner. Several studies provide evidence in support of this hypothesis. It is reported that DNA-peptide cross-links are preferable substrates over DNA-protein cross-links in initiating DNA repair.[169,324,325,330,331] Further, proteasomal inhibitors have been reported to slow down the repair of formaldehyde-induced DPCs in cells[297] and intracellular repair of DPC-containing plasmids.[331] In addition, several *in vitro* replication studies using site-specific substrates containing small peptides cross-linked within the major or minor groove of DNA have provided direct evidence for the ability of lesion bypass polymerases to catalyze nucleotide incorporation opposite the DNA-peptide lesions.[335-337] However, all of the previous studies were conducted with model DPCs containing peptides conjugated to various positions within purine nucleobases of DNA (the N^6 of adenine and the N^2 of guanine), and no information is available on polymerase bypass of the corresponding lesions involving pyrimidine residues in DNA. Such lesions can be generated upon exposure to endogenous free radicals, UV light and ionizing radiation.[182,211,410-412] Further, irreversible trapping of DNA methyltransferases on cytosines in genomic DNA can form DPCs that can adversely affect gene expression.[408,409]

Our results reported herein (**Tables 6-1** and **6-2**) provide the first evidence for error-prone bypass of peptides conjugated to the major groove of DNA through a pyrimidine nucleobase by human lesion bypass polymerases. Our model lesions generated *via* click chemistry contained a 10-mer peptide (Glu-Gln-Lys-Leu-Ile-Ser-Glu-Glu-Asp-Leu-NH_2) cross-linked to the *C5* position of thymidine (**Scheme 6-1**). We found that both hPol κ and η can incorporate any of the four nucleotides opposite the cross-link (**Figure 6-2**) suggesting the possibility of both transition and transversion mutations. Furthermore, according to MS quantitation, 96% and 81% of the replication products of DNA-peptide cross-links by hPol κ and η were erroneous products (**Table 6-2**).

Low fidelity DNA replication by hPol η and κ is well documented.[150,415-417] Both enzymes catalyze error-prone bypass of 8-oxo-dG,[418,419] abasic site,[418-420] 1,N^6-HMHP-dA,[112] 1,N^6-etheno-dA[394] and N^2,3-etheno-dG.[395] hPol η also replicate past bulky benzo[*a*]pyrene-N^2-dG[418] in an error-prone manner, while hPol κ erroneously bypasses 2-acetylaminofluorene-dG.[419,420] With respect to DNA-peptide cross-links, these results are consistent with highly miscoding potential observed with hPol κ and η in bypassing the DNA-peptide conjugate, N^6-dA-$(OH)_2$butyl-GSH.[338] Extremely low accuracy observed with hPol κ in bypassing the dT-peptide cross-link in this work is of great interest in that hPol κ is known to exclusively incorporate the correct base, dC opposite peptides of comparable size (12-mer) cross-linked to N^2 of dG.[335] This discrepancy can arise from that dT-peptide lesion is cross-linked to the major groove of DNA, while N^2-dG-peptide is to the minor groove,[335] because it is reported that site of cross-linking within DNA is important for the ability of polymerase to bypass DNA adducts.[336]

Efficiency of replication bypass of adducted DNA is as important as the fidelity of the replication. From single nucleotide insertion assays, we found that hPol η is more efficient at bypassing DNA-peptide cross-links as compared to hPol κ (**Figure 6-2**). These results were confirmed by quantitative data obtained from steady-state kinetics (**Table 6-1**). Catalytic efficiencies (k_{cat}/K_m) of dNTP incorporation opposite the dT-peptide conjugate by hPol η were 2.3–463-fold higher than that of hPol κ. Further, misinsertion frequencies (*f*) for hPol η were 22.5–205-fold higher than those of hPol κ. Taken together, these data suggests that hPol η is more efficient than hPol κ in replication bypass in the context of the sequence and the adduct studied in this work. The order of bypass efficiency, hPol η > hPol κ is in agreement with the previously reported values for 1,N^6-etheno-dA,[394] 1,N^6-HMHP-dA,[112] N^2,3-etheno-dG[395] and N^6-dA-$(OH)_2$butyl-GSH.[338]

Additional information was provided from capillary HPLC-ESI^--MS/MS sequencing of extension products (**Table 6-2**). We found that hPol κ replicated the unadducted template with good accuracy, where 87% of the products correspond to the error-free full extension product (**Table 6-2**, **Figures 6-3**A and **6-4**A). In contrast, replication bypass of the dT-peptide conjugate was highly error-prone, inducing a large number of one and two-nucleotide deletions (96% of total products, see **Table 6-2**, **Figures 6-3**B, **6-4**B and **6-5**). It has been previously reported that hPol κ shows particularly high single base insertion/deletion rates in repetitive sequences,[136,421] which is applicable in this case (5'-TCA ***TT***G AAT CCT TCC CCC-3').

HPLC-ESI$^-$-MS/MS sequencing of extension products of hPol η using the unmodified template has revealed that the replication was surprisingly error prone (**Table 6-2**). Specifically, hPol η incorporated the correct base (dA) opposite the first T in 5'-TCA ***TT***G AAT CCT TCC CCC-3', but the subsequent nucleotide insertion opposite the second T was highly error-prone (**Table 6-2**, **Figures 6-6**A and **6-7**). The relatively low fidelity of replication of undamaged DNA by hPol η is well documented.[136,150,415,417] It has been reported that hPol η has the highest error rates amongst Y-family polymerases.[136] The highest substitution rates for hPol η were observed for templates containing undamaged dT and d(TT), and dG and dC were preferred in such instances,[415,417] as observed with our template (**Table 6-2**). The presence of dT-peptide lesion at the first T led to a large number of one-base deletions (43% of the total replication products; **Table 6-2**, **Figures 6-6**B and **6-8**). In addition, hPol η-catalyzed replication gave rise to significant amounts of transition and transversion mutations (15% and 23%, respectively; **Table 6-2**, **Figures 6-6**B and **6-8**).

These results are consistent with the low fidelity and error-prone replication of TLS polymerases due to their large and flexible active sites and lack of 3'→5' exonuclease activity.[149] Mechanisms for the induction of deletions by TLS polymerases have been proposed,[394,405,406] while X-ray crystallographic intermediates of these extension products have also been detected.[407] The adducted base or adducted base and the base adjacent to it can misalign rendering a slipped or extrahelical conformation.[394,407] Alternatively, slippage event can occur subsequent to nucleotide insertion opposite the adduct.[394,406,407] Extension of the misaligned primer yields one- and two-base

deletions,[394,406,407] while realignment after nucleotide insertion and further extension cause substitution mutations.[394,407] These prominent mutations (96% of total products for hPol κ) may also take place upon replication of other DNA-peptide conjugates, but would not be detected by electrophoresis methodologies employed in previous studies.[335-337] Furthermore, standard single nucleotide insertion assays opposite the damaged base cannot detect mutations occurring downstream from the adduct, as observed in the present study using HPLC-ESI$^-$-MS/MS sequencing (**Table 6-2**).

In summary, we investigated the replication bypass of 10-mer peptides cross-linked to major grove of DNA *via* a pyrimidine base (*C5* of T) using human lesion bypass polymerases η and κ. Our gel electrophoresis and HPLC-ESI$^-$-MS/MS results provide evidence for highly error-prone replication past these bulky adducts. Specifically, large numbers of -1 and -2 deletions were observed for the replication catalyzed by hPol κ (**Table 6-2**). Low fidelity and deletion products observed in bypass of DNA-peptide cross-links can eventually contribute to mutagenesis of DPCs.[304] To our knowledge, this is the first systematic study on replication fidelity of DPCs involving pyrimidine residues of DNA. In future studies, it would be interesting to determine how these peptide conjugates are accommodated in the active site of TLS polymerases. Furthermore, X-ray crystallographic or NMR structural studies of ternary complexes of these adducts in the presence of TLS polymerases and dNTPs can provide details on why and how this error-prone replication occurs, especially the reasons for the large amounts of deletion mutations observed.

7 SUMMARY AND CONCLUSIONS

DNA damage is a continuous process by which genomic DNA is modified by various endogenous and exogenous physicochemical agents, and can potentially threaten the integrity of genetic information and decrease cell viability.[1,4,5] DNA adducts can disrupt DNA transactions triggering apoptosis,[422] unless rapidly repaired.[8,9] Alternatively, the persisting lesions can be replicated in an error-prone manner,[2,8,9] ultimately causing adverse biological effects.[30-39] Mitigation or exacerbation of these deleterious effects is determined by the cellular dynamics of DNA repair and replication. Hence, systematic investigation of repair and replication of DNA adducts is essential to fully understand the biological consequences of DNA damage.

1,3-Butadiene (BD) is a known human carcinogen and mutagen that has a widespread occurrence in the environment, resulting in a high potential for human exposure.[25,45-50,63-65] Epidemiological and toxicological studies in humans and inhalation studies in animals have provided evidence for increased risk of cancer incidence upon exposure to BD.[46,64,65] Further, chronic exposure to BD has been associated with respiratory and cardiovascular diseases.[51-62]

The adverse biological effects of BD are attributed to DNA damage mediated by its epoxide metabolites. Highly reactive epoxides: 3,4-epoxybut-1-ene (EB), 3,4-epoxybutan-1,2-diol (EBD) and 1,2,3,4-diepoxybutane (DEB) are formed upon

metabolic activation of BD by cytochrome P450 monooxygenases (**Scheme 1-1**).[66,67] Unless detoxified by epoxide hydrolase (EH) or conjugation with glutathione,[68] these metabolites can react with multiple sites within DNA and proteins, forming covalent adducts (**Scheme 1-1**).[73-77] Among these, DEB is considered the ultimate carcinogenic metabolite of BD because of its high genotoxicity, which is attributed to its ability to *bis*-alkylate biomolecules to form exocyclic adducts (**Chart 1-5**),[78,79] DNA-DNA cross-links (**Chart 1-4**) [80-84] and DNA-protein cross-links.[85,86]

The mutagenic and carcinogenic properties of BD are well documeted.[70,90,358-360] Nonetheless, the mechanisms of cellular repair of BD-induced DNA lesions have not been identified. Since the dynamics of repair and replication of DNA adducts as well as the ability of these lesions to cause mispairing can significantly alter the biological effects of DNA damage, identifying the repair pathways of BD-DNA adducts is crucially important.

Despite the preferential alkylation of guanine nucleobases by BD-derived epoxides,[357] BD induces a large number of A $\rightarrow$ T mutations, suggesting the significant contribution of adenine lesions to its biological effects.[90-93] In the first part of this thesis, we examined the repair mechanisms responsible for the removal of three novel BD-induced 2'-deoxyadenosine (dA) lesions in human cells (Chapter 2). Synthetic oligodeoxynucleotides containing site- and stereospecific *S*-N^6-(2-hydroxy-3-buten-1-yl)-2'-deoxyadenosine (*S*-N^6-HB-dA) monoadduct and the exocycles, *R,S*-1,N^6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2'-deoxyadenosine (*R,S*-1,N^6-HMHP-dA) and *R,R*-N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine *(R,R*-N^6,N^6-DHB-dA) using a post-

oligomerization strategy (**Scheme 2-3**).[362] The adducted DNA strands were characterized by capillary HPLC-ESI$^-$-MS, ESI$^+$-MS^3 and HPLC-UV analysis of enzymatic digests and MALDI-TOF-MS of controlled exonuclease digests (**Figure 2-1** and **Table 2-1**).

Gel electrophoresis-based *in vitro* repair assays employing radiolabeled dsDNA substrates and nuclear extracts from human fibrosarcoma (HT1080) cells have revealed that all three BD-dA adducts were recognized and cleaved by the base excision repair (BER) mechanism (**Figures 2-2** and **2-3**). First order rate constants revealed that the efficiency of repair decreased in the order of (*S*)-N^6-HB-dA > (*R*,*R*)-N^6,N^6-DHB-dA > (*R*,*S*)-1,N^6-HMHP-dA, while BD-dA adducts were recognized at comparable rates to known BER substrates, 5F-dU and 8-oxo-dG (**Table 2-2** and **Figure 2-4**). We observed 50–75% inhibition in repair of the adducts in the presence of a known BER inhibitor, methoxyamine, confirming that BER is indeed responsible for repair of these adducts (**Figure 2-4**). To further confirm the involvement of BER, we developed an HPLC-ESI$^-$-MS/MS method to characterize the excision products. These experiments have detected the expected precursor ions ([M-H]$^-$) and fragmentation patterns for the expected oligonucleotide products in two sequence contexts (**Figure 2-6**). In an effort to identify the DNA glycosylases involved in removal of BD-dA lesions, we employed 12 human and bacterial glycosylases (AAG, Mpg, edited and unedited NEIL1, OGG1, EndoIII, Fpg, Mug, MutY, MutYH, Nei and UDG). However, none of the enzymes were able to recognize the adducts of interest (**Figures 2-7** and **2-8**), suggesting that an alternate glycosylase is responsible for the observed repair. To our knowledge, this is the first

direct evidence for the active repair of biologically relevant BD-induced DNA adducts in human cell-free extracts.

Covalent entrapment of cellular proteins on genomic DNA upon exposure to various endogenous and exogenous chemicals and physical agents *in vivo* results in the production of highly heterogeneous and unusually bulky DNA lesions, DNA-protein cross-links (DPCs).[4,5] The complexity and diversity of these highly abundant, yet understudied DNA adducts mainly arise from the partaking of a wide range of proteins of varying sizes, properties and functions,[85,158-160] and participation of numerous sites on DNA and multiple amino acids on proteins in cross-linking (**Chart 1-6**).[4,85,158,161-168] Endogenous and exogenous *bis*-electrophiles are the most common type of chemicals that participates in DPC generation in cells.[85,86,189,190,195-210]

DPCs can compromise genetic integrity and cell viability due to their unusual bulkiness as compared to "standard" DNA lesions.[4,85,158-168] The cytotoxic effects of a number of DNA damaging agents including formaldehyde have been attributed to DPC formation.[39,132,216,220,234,251,259,260,299-302] The ability of DPCs to interfere with crucial DNA-protein interactions in cellular processes such as DNA replication, transcription and repair is thought to lead to cytotoxic and mutagenic outcomes.[4,5] These in turn can contribute to various diseases including cancer,[30-33] cardiovascular disease[34-36] and age-related neurodegeneration.[35-39]

The biological consequences of DPCs are poorly understood because of the tendency of many DPC-inducing agents to induce other types of DNA adducts,[262] and

their highly heterogenic nature. In this regard, novel methodologies are required to induce and characterize DPCs selectively in cells and to generate site-specific, biologically relevant DPCs to study repair and replication *in vitro*. Previously reported synthetic methodologies have several limitations such as poor reaction efficiency and low yields,[164] limited choice of proteins,[331] and insufficient site specificity with respect to the cross-linking site within the protein.[164,246,354] In the second part of this thesis, we have developed two novel methodologies to generate hydrolytically stable, sequence specific model DPCs and investigated their effects on DNA replication.

The *N7* position of guanine is the most nucleophilic site within DNA and is therefore the most common site for nucleobase adduct formation.[357] It is also the most likely site to be involved in DNA-protein cross-linking in the presence of *bis*-electrophiles such as environmental carcinogens and antitumor drugs (**Schemes 1-5** and **1-6**).[85,86,158,159,161,294] Yet, no methods exist in the literature to generate *N7*-G conjugated DPCs because of the tendency of *N7* alkylated guanines to undergo spontaneous depurination.[375] Herein, we developed a new synthetic methodology to generate site-specific, hydrolytically stable model DPCs that are structural mimics of *N7* conjugated DPCs such as those formed *in vivo* (Chapter 3). We developed a strategy based on post-synthetic reductive amination to cross-link Lys and Arg residues within proteins and peptides to the *C7* position of aldehyde-functionalized 7-deaza-guanine *via* an ethylene linkage (**Scheme 3-2**). The resulting model DPC substrates are direct analogs of DPCs generated by antitumor nitrogen mustards and chlorooxirane.[158,159,161] We tested the applicability of our method to generate DPCs using a range of proteins and peptides

(**Table 3-1** and **Figure 3-4**). In general, we observed that DNA binding proteins produced DPCs in high yields compared to proteins that do not have an affinity to DNA. We employed denaturing gel electrophoresis and a nanoLC-nanospray-MS/MS method to characterize the model DPCs. DPCs were visualized as low mobility bands on denaturing gels by phosphorimaging of radiolabeled DNA or protein staining (**Figures 3-1** and **3-2**). Mass spectrometry characterization of the peptide-nucleoside conjugates resulting from the enzymatic digestion of the DNA component and tryptic digestion of the protein revealed that Lys and Arg residues on proteins are involved in cross-linking reactions (**Table 3-2**, and **Figures 3-5** and **3-6**). We further employed a capillary HPLC-ESI$^+$-MS^2 method and a nanoLC-nanospray-MS^3 method to confirm the exact chemical structures of the resulting conjugates. MS/MS spectra of 7-deaza-7-(2,3-dihydroxy-propan-1-yl)-2'-deoxyguanosine (deaza-DHP-dG) conjugates involving free Arg and Lys (**Figure 3-7**) and two peptide conjugates (involving substance P and Angiotensin I) (**Figure 3-8**) have confirmed that side chain amino groups of Lys/Arg residues participate in cross-linking reactions. DPCs generated using this methodology can be employed in repair and replication studies to fill the gaps in our understanding of the biological consequences of DPCs formed in cells.

Model DNA-protein and DNA-peptide cross-links conjugated to the major groove of DNA *via* the *C7* position of 7-deazaguanine were used as templates in primer extension studies to evaluate the effects of DPCs on DNA replication. Gel electrophoresis-based primer extension assays have revealed that large DNA-protein cross-links (11.4 kDa histone H4, 22.9 kDa AlkB and 28.4 kDa 6xHis-eGFP) represent a

complete block to human TLS polymerases η and κ under both standing start and running start conditions (**Figures 4-1** and **4-2**). Although a 23-mer peptide cross-linked to *C7*-G blocked replication in standing start experiments, a smaller 10-mer peptide cross-link was bypassed (**Figure 4-3**). These results suggest that DPCs formed to *N7* of guanine in cells can get proteolytically degraded to smaller DNA-peptide cross-links to facilitate polymerase bypass. The efficiency and fidelity of bypass of the decapeptide cross-linked to DNA was evaluated using steady-state kinetic experiments (**Table 4-1**). We found that hPol η incorporated the correct base, dC, opposite the peptide cross-link with high efficiency and 500-fold preference over dT. hPol κ catalyzed error-free bypass of the peptide cross-link, despite low efficiency. The order of bypass efficiency, hPol η > hPol κ is in agreement with the previous reports for other bulky lesions.[112,338,394,395]

In Chapters 5 and 6, we described a novel method to generate DPCs conjugated to the major groove of DNA *via* pyrimidine nucleobases and examined their effects on replication bypass. Site-specific DPC substrates were prepared by copper-catalyzed [3+2] Huisgen cycloaddition between an alkyne-functionalized 5-(octa-1,7-diynyl)-uracil in DNA and an azide group within 6×His-eGFP, and 23-mer and 10-mer peptides derived from tetanus toxoid[423] and c-Myc oncogene,[424] respectively (**Schemes 5-1** and **5-2**). Gel electrophoresis and mass spectrometry-based techniques described above were used to characterize DNA-protein and DNA-peptide conjugates (**Figures 5-1**–**5-5**). The results of *in vitro* primer extension assays were similar to those observed with *C7*-G cross-linked polypeptides. DNA primers annealed to templates containing large 6×His-eGFP and a 23-mer cross-links completely blocked lesion bypass polymerases η and κ (**Figures 5-7**

and **5-8**). In contrary, decapeptide conjugated to *C5* of thymidine was bypassed, further confirming our hypothesis that cross-linked protein in cellular DPCs gets proteolytically processed to generate smaller peptide prior to replication (**Figures 5-7** and **5-8**). Complete replication block by proteins conjugated to pyrimidines has been reported.[132] However, our observation of replication block in the presence of 23-mer peptide contradicts previous reports that tetra- and dodecapeptides cross-linked to N^2-G and N^6-A were bypassed by hPol κ and ν, respectively.[335,336] This supports the notion that the cross-linking site within DNA can influence the biological consequences of DPC lesions.

We further examined the fidelity of replication bypass of the decapeptide cross-linked to major grove of DNA *via* pyrimidine bases (Chapter 6). Steady-state kinetic studies were performed using the purified conjugates to determine the catalytic efficiency (k_{cat}/K_m) and misinsertion frequency (*f*) opposite the adducted site. We found that hPol η is catalytically more efficient than hPol κ, but nucleotide insertion opposite the adduct was more error-prone (**Table 6-1**). Catalytic efficiencies of dNTP incorporation opposite the dT-peptide conjugate by hPol η were 2.3–463-fold higher than that of hPol κ. Further, misinsertion frequencies for hPol η were 22.5–205-fold higher than those of hPol κ. Although both enzymes can insert any of the four dNMPs opposite the adduct, hPol η misincorporated dGMP with 4.5-fold preference over the correct nucleotide, dAMP. A methodology involving affinity capture in combination with liquid chromatography-tandem mass spectrometry was developed to quantify and sequence the replication products (**Scheme 6-3**). Capillary HPLC-ESI$^-$-HRMS quantitation and MS/MS sequencing of primer extension products of the 10-mer peptide conjugated to *C5* of

thymidine provided evidence for the highly error-prone replication past these bulky adducts by human TLS polymerases η and κ (**Table 6-2** and **Figures 6-3**–**6-8**). Moreover, the majority of the products of hPol κ-catalyzed primer extension (96%) corresponded to one- and two-base deletions (**Figure 6-4**). In addition to the effect of the bulky adduct, the local sequence context (5'-TCA ***TT***G AAT CCT TCC CCC-3') may be responsible for these deletions, since hPol κ can cause insertion/deletion events in repetitive sequences.[136] Overall, our results are consistent with the low fidelity and error-prone replication of TLS polymerases due to their large and flexible active sites and lack of 3' → 5' exonuclease activity.[149] X-ray crystallographic evidence of deletion products resulting from translesion synthesis of 1,N^2-εdG further corroborate our findings.[407] These prominent mutations (96% of total products for hPol κ) may also take place upon replication of other DNA-peptide conjugates, but would not be detected by electrophoresis methodologies employed in previous studies.[335-337] Furthermore, the mutations observed downstream from the adduct, such as base substitutions at +1 position, cannot be detected by standard gel electrophoresis-based assays and were only revealed due to the use of HPLC-ESI$^-$-MS/MS methodology (**Table 6-2**). Extremely low polymerase fidelity and deletion products observed in bypass of these DNA-peptide cross-links can eventually contribute to mutagenesis of DPCs.[304] To our knowledge, this is the first systematic study on replication fidelity of peptides covalently attached to pyrimidines on DNA, and it has provided valuable insights into deleterious effects of such bulky DNA lesions formed *in vivo*.

In summary, the present study employed a combination of nucleic acid chemistry and organic synthetic tools to prepare stereo- and site-specific DNA oligonucleotides containing BD-dA lesions and site-specific, hydrolytically stable DPCs that are structural mimics of biologically relevant DPCs formed to purine and pyrimidine nucleobases in the major groove of DNA. Biochemical tools such as gel electrophoresis and bioanalytical methodologies including liquid chromatography-tandem mass spectrometry were used to examine the repair of BD-dA lesions and to investigate the effects of DPCs on DNA replication. In the first part of this thesis, gel electrophoresis-based assays provided evidence that all three BD-dA adducts studied in the current work, *S*-N^6-HB-dA, *R,S*-1,N^6-HMHP-dA and *R,R*-N^6,N^6-DHB-dA, can be recognized by cellular BER pathway and cleaved efficiently in human cells (Chapter 2). Repair inhibition assays and 5' excision products characterized by capillary HPLC-ESI$^-$-MS/MS further confirmed the involvement of BER in removal of these lesions.

In the second part of this thesis, we have developed two synthetic strategies to generate site-specific DPCs using reductive amination (Chapter 3) and copper-catalyzed [3+2] Huisgen cycloaddition reaction (Chapter 5). Replication bypass of synthetic DPCs prepared by both strategies revealed that large proteins and polypeptides conjugated to the major groove of DNA completely block replication irrespective of the protein identity and site-of cross-linking within DNA (Chapter 4 and 5). However, human TLS polymerases were able to replicate past small decapeptides conjugated to DNA with varying efficiencies suggesting the proteolytic degradation of the cross-linked protein *in vivo* prior to repair or replication (Chapter 4 and 5). Furthermore, according to the

results from single nucleotide insertion assays and steady state kinetic studies, we found that the human TLS polymerases catalyze bypass of a decapeptide cross-linked to *C7* of guanine with high fidelity (Chapter 4). In contrast, bypass of a decapeptide conjugated to *C5* of thymidine was highly error-prone (Chapter 6). Our Capillary HPLC-ESI$^-$-HRMS and MS/MS experiments further revealed that replication past peptides conjugated to *C5* of thymidine causes a large number of base substitution and deletion mutations (Chapter 6). Taken together, our data suggest that efficiency and fidelity of replication bypass of DNA-polypeptide cross-links are dependent on size of the cross-linked polypeptide, cross-linking site within DNA and the identity of the polymerase.

8 FUTURE DIRECTIONS

8.1 IDENTIFICATION OF DNA GLYCOSYLASES INVOLVED IN REPAIR OF N^6-2'-DEOXYADENOSINE LESIONS OF 1,3-BUTADIENE

As described in Chapter 2, we have shown that base excision repair (BER) is one of the repair pathways involved in the removal of *S*-N^6-HB-dA, *R,S*-1,N^6-HMHP-dA and *R,R*-N^6,N^6-DHB-dA adducts in human cells. However, none of the human and bacterial glycosylases tested (human AAG, Mpg, edited and unedited NEIL1 and OGG1, *E. coli* EndoIII, Fpg, Mug, MutY, MutYH and Nei, and *A. fulgidus* UDG) were able to recognize these adducts (**Figures 2.3** and **2.4**). The use of mispaired DNA duplexes, if significant changes in repair rates were observed for mispaired DNA duplexes with nuclear extracts, can be a useful strategy to identify the BER glycosylases responsible for the cleavage of these adducts. These experiments are currently in progress (in collaboration with Dr. Sheila David, University of California, Davis).

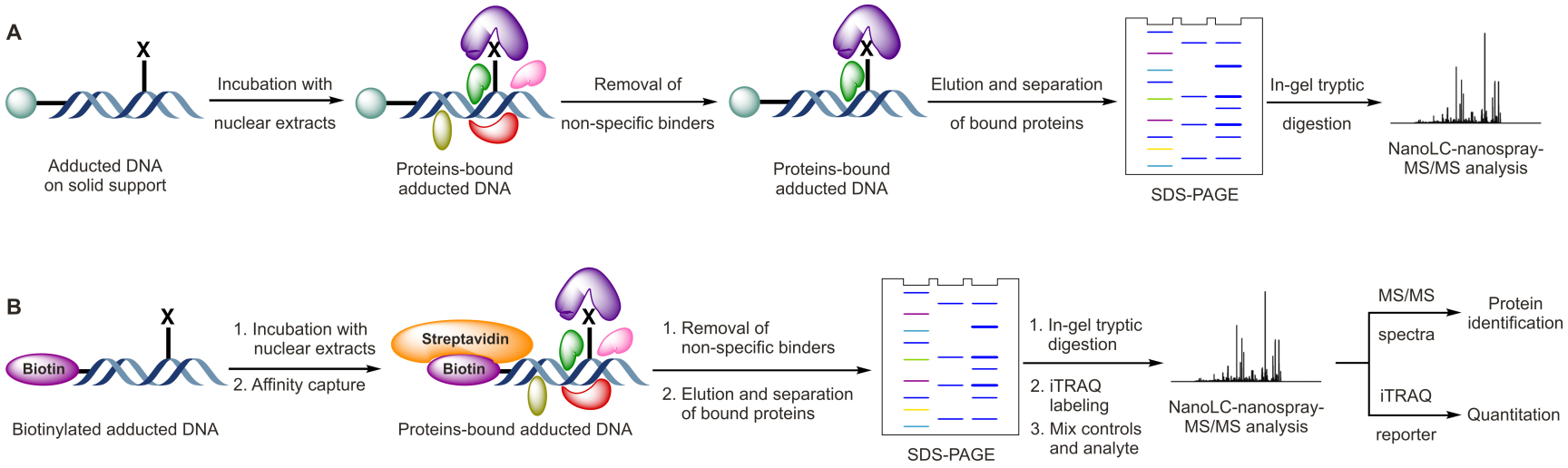
To identify any additional/novel DNA glycosylases that may be involved in repair of BD-dA lesions, a mass spectrometry-based proteomics strategy can be employed (**Scheme 8-1**). The adducts-containing DNA oligomers and unadducted controls will be

prepared on solid support. Adducted and control DNA duplexes will be preincubated with methoxyamine to prevent APE cleavage of abasic sites generated by BER enzymes upon removal of the adduct. Following incubation with nuclear protein extracts from human cells in the presence of a large excess of carrier DNA to minimize non-specific binding, DNA-bound proteins will be eluted using a salt gradient, SDS, and heat. Salts and low concentrations of SDS will remove any non-specific binders, while heating with SDS will release strongly-bound proteins. Fractionation of DNA-bound proteins by SDS-PAGE and sensitive nanoLC-nanospray-MS/MS detection on an LTQ Orbitrap Velos instrument will increase the chances of discovery of the novel glycosylases using MS-based proteomics (**Scheme 8-1**). Elimination of protein hits from salt eluates and unmodified DNA controls can aid in revealing the BD-dA adduct-specific glycosylases. We have already developed methods to synthesize these adducts on solid support (**Scheme 8-2** and **Figure 8-1**). Alternatively, an affinity capture-based approach can also be employed using 5'-biotinylated adducted DNA.[85,294,425,426] If the glycosylase involved is a bifunctional enzyme having both glycosylase and AP lyase activity,[6,427,428] reduction of the Schiff base formed between the enzyme and the AP site upon incubation of adducted DNA with nuclear extracts can be reduced to form a stable amine conjugate. Covalently trapped DNA-proteins cross-links on solid support/streptavidin beads can be used for nanoLC-MS/MS analysis.

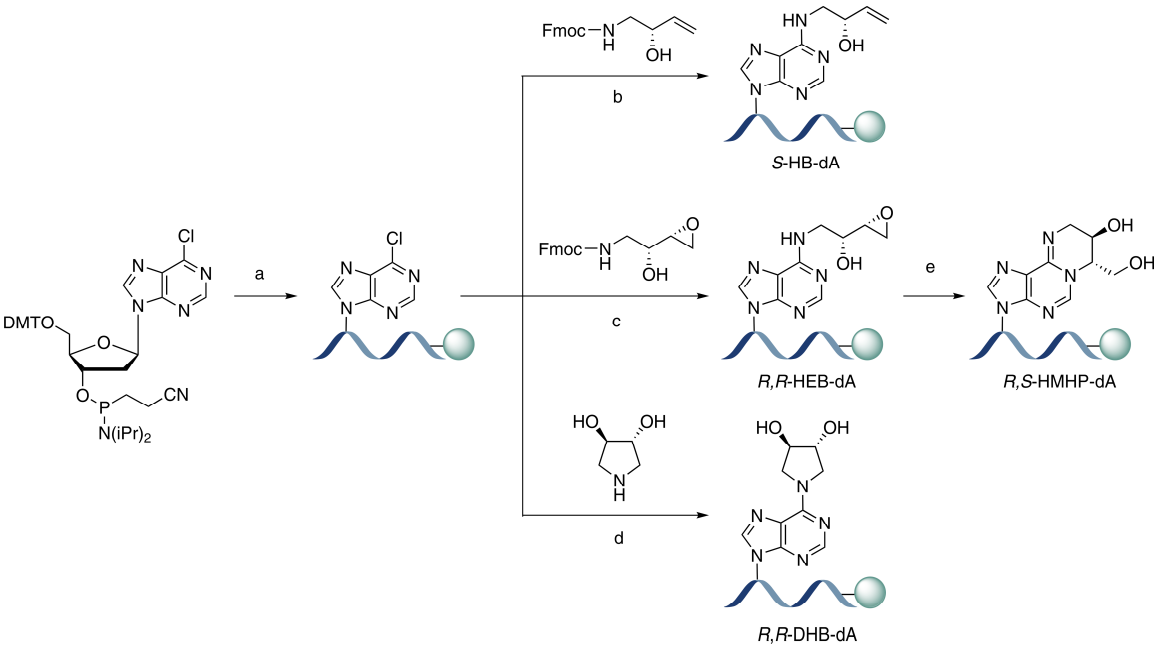
Isotope labeling-based proteomics approach such as isobaric tag for relative and absolute quantitation (iTRAQ) can also be employed.[429-432] BD-dA adducted dsDNA (on solid support or biotinylated) will be incubated with nuclear extracts. Unmodified

dsDNA will be used as controls. Following protein elution, fractionation and in-gel digestion, samples will be labeled with iTRAQ reagents (iTRAQ reagents–4plex, Ab Sciex, Foster City, CA). Peptide backbone fragment ions from nanoLC-MS/MS analysis will be used for protein identification, while relative quantities of iTRAQ reporter ions will be used to identify the proteins specifically bound to BD-dA adducts.[430,431] A higher abundance of reporter ions corresponding to the samples from BD-dA adducted DNA as compared to the negative control will indicate any glycosylases responsible for the repair of the adducts.

Scheme 8-1 Mass spectrometry-based proteomics approach to identifying the base excision repair glycosylases responsible for the repair of 1,3-butadiene-induced 2'-deoxyadenosine adducts.

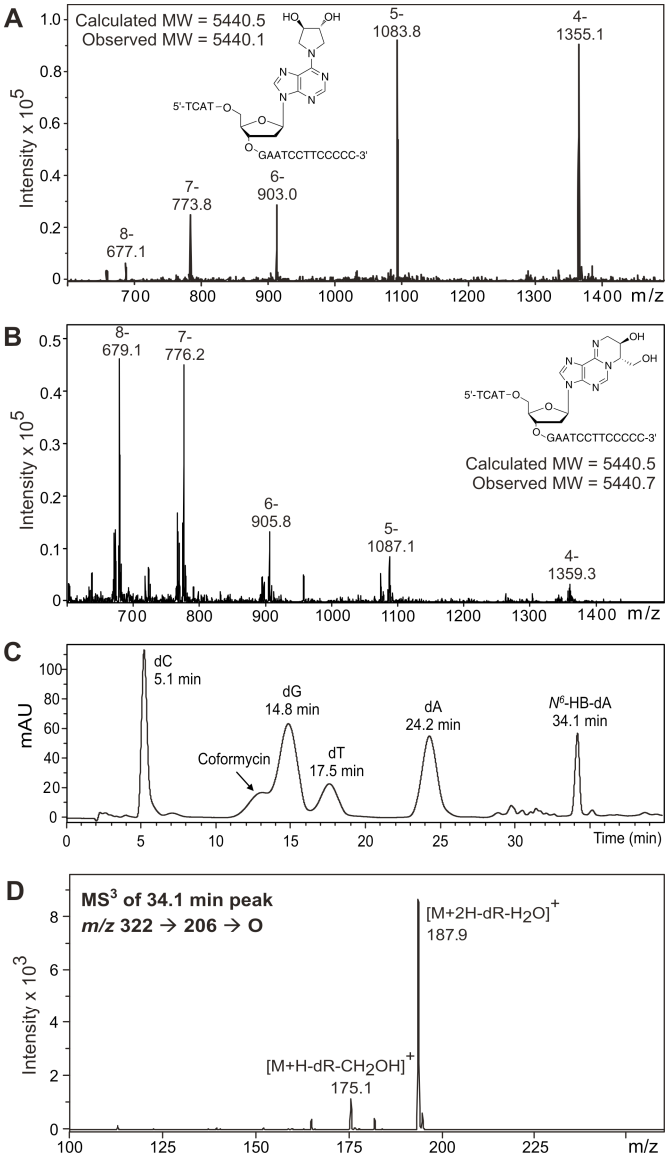


Scheme 8-2 Synthesis of 1,3-butadiene-induced 2'-deoxyadenosine adducts on solid support.[§§]



[§§] (a) solid phase synthesis, (b) DIPEA, DMSO, 60 °C, 24 h, (c) DIPEA, DMSO, 37 °C, 24 h, (d) DIPEA, DMSO, 37 °C, 72 h, (e) H_2O, rt, 3h.

Figure 8-1 Characterization of 18-mer oligodeoxynucleotides containing site- and stereo-specific 1,3-butadiene lesions after cleavage from the solid support. Capillary HPLC-ESI$^-$-MS spectra of 18-mer containing *R*,*R*-DHB-dA (A) and *R*,*S*-HMHP-dA (B), HPLC-UV analysis of enzymatic digests of the 18-mer containing *S*-HB-dA (C) and ESI$^+$-MS^3 spectrum of *S*-HB-dA (D).



8.2 *IN VITRO* AND *IN VIVO* INVESTIGATION OF REPAIR OF 1,3-BUTADIENE-INDUCED N^6-ADENINE ADDUCTS

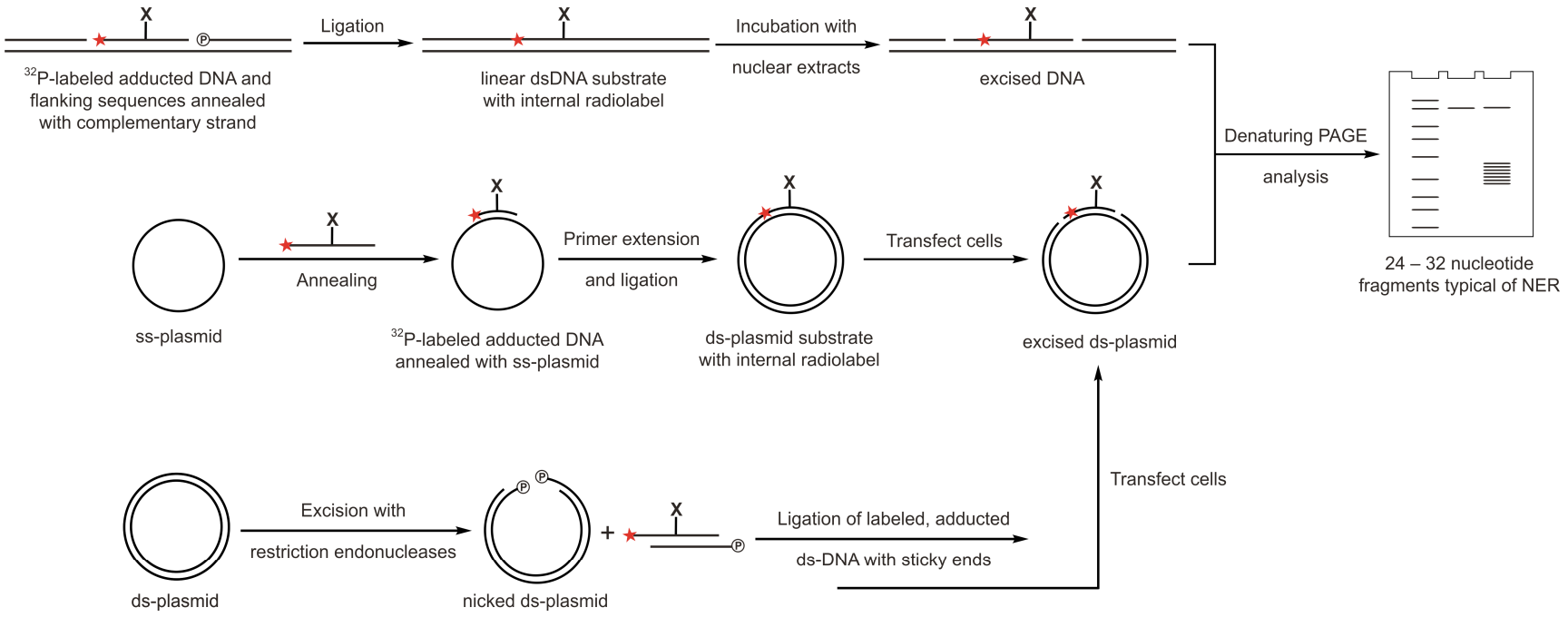
Our *in vitro* repair studies revealed efficient recognition of BD-dA adducts by BER in human cell free extracts (Chapter 2). Nonetheless, it will be interesting to examine whether these adducts can be recognized by other cellular repair mechanisms. For example, it is reported that guanidinohydantoin (Gh), spiroiminodihydantoin (Sp) and hydantoin amine adducts formed upon oxidation of guanine are good substrates for both BER and NER.[433]

To examine whether NER plays a role in removal of BD-dA lesions, *in vitro* repair experiments can be performed using HeLa nuclear extracts (**Scheme 8-3**). The site-specifically adducted 18-mer (5'-TCA T***X***G AAT CCT TCC CCC-3') oligodeoxynucleotides prepared in Chapter 2 will be 5'-endlabeled, ligated with 5' and 3' flanking sequences and annealed to the complementary strands to construct longer (50 nucleotides or more), linear ds-DNA substrates (**Scheme 8-3**).[128,433,434] Alternatively, ^{32}P-end-labeled, site-specific BD-dA adducted 18-mer DNA strands will either be annealed to single-stranded plasmids followed by primer extension[435] or be inserted into restriction endonuclease-nicked double-stranded plasmids,[436] and ligated to prepare double-stranded plasmids containing the adducts (**Scheme 8-3**). Following incubation with cell free protein extracts, samples will be analyzed by denaturing PAGE, and the incision products will be visualized by phosphorimaging.[128,433-435] Unmodified and aminofluorene adducted dG (*N*-(2'-deoxyguanosin-8-yl)-2-aminofluorene or *N*-(2'-deoxy-

guanosin-8-yl)-2-acetylaminofluorene) containing DNA substrates will be used as negative and positive controls, respectively.[435]

For time course studies of DNA repair, unlabeled dsDNA substrates will be incubated with nuclear extracts or human cells will be tranfected with adducted plasmids for predetermined time points (0–180 min) and purified using a commercial DNA isolation kit. Following enzymatic digestion to nucleosides, BD-dA adducts will be isolated by solid phase extraction.[111] Isotope dilution liquid chromatography-tandem mass spectrometry (HPLC-ESI-MS^n) methods similar to those developed by Tretyakova and coworkers[111] will be used for adduct quantitation. The mass spectrometry-based methods can be advantageous over biochemical methods described above due to the ability of absolute quantitation.

Scheme 8-3 *In vitro* repair experiments to examine the involvement of nucleotide excision repair in the removal of 2'-deoxyadenosine lesions of 1,3-butadiene.



8.3 EXAMINATION OF *IN VIVO* MUTAGENESIS OF 2'-DEOXY-ADENOSINE LESIONS OF 1,3-BUTADIENE

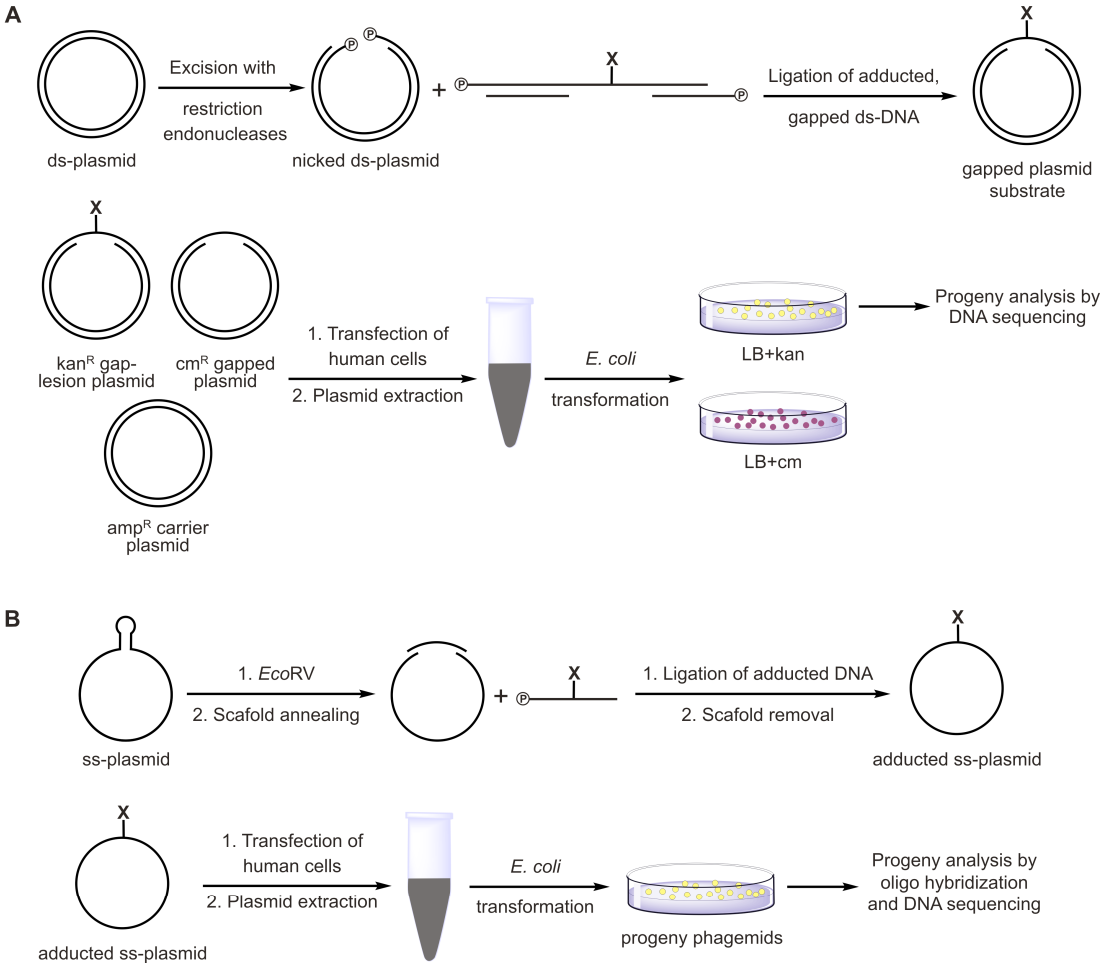
In vitro replication studies conducted in our laboratory using human TLS polymerases have provided evidence for the mutagenic potential of both exocyclic BD-dA adducts.[112,157] Yet, the applicability of these mutation spectra *in vivo* should be evaluated as cellular responses to DNA damage is complex: DNA repair and tolerance mechanisms that operate prior to or simultaneously with DNA replication can mitigate the effects of damaged DNA. Quantitative lesion bypass and mutagenesis experiments of BD-dA adducts are currently in progress (in collaboration with Dr. John Essigmann, Massachusetts Institute of Technology). Site- and stereo-specific 16-mers, 5'-GAA GAC CT***X*** GGC GTC C-3', containing BD-dA adducts prepared in our laboratory were ligated to ss-M13 plasmids linearized using *Eco*RV with the aid of a scaffold DNA, and consequently transfected into *E. coli*.[437-439] The effects of BD-dA adducts on DNA replication will be assessed by using the competitive replication of adduct bypass (CRAB) assay,[438,439] while site-specific mutation frequency will be calculated using the restriction endonuclease and post-labeling analysis of mutation frequency (REAP) assay.[437-439] Preliminary data indicates that *S*-N^6-HB-dA and *R,R*-N^6,N^6-DHB-dA do not block replication, while *R,S*-1,N^6-HMHP-dA strongly blocks replication (~10% bypass in non-induced cells). However, none of the lesions were strongly mutagenic in *E. coli*. *R,R*-N^6,N^6-DHB-dA induced ~1% A $\rightarrow$ G and A $\rightarrow$ T mutations. *R,S*-1,N^6-HMHP-dA caused ~2% A $\rightarrow$ T mutations, while *S*-N^6-HB-dA did not appear to be mutagenic. No deletion mutations were observed for any of the three adducts under the assay conditions.

These *in vivo* mutagenicity data contradict what has been observed *in vitro* with human TLS polymerases,[112,157] suggesting that these adducts are efficiently repaired in cells. However, the kinetics of repair and replication of these adducts can be different in bacterial and human cells.

To study mutagenicity of these adducts in human cell cultures, quantitative gap-filling assay[440,441] and single-stranded pMS2-based oligonucleotide hybridization assay[442] can be employed (**Scheme 8-4**). Gapped plasmids (chloramphenicol resistant, cm^R GP20 and kanamycin resistant, kan^R GP21) will be constructed by ligating synthetic gapped duplex oligonucleotides (unmodified or BD-dA adducted) to restriction endonuclease cleaved plasmids (**Scheme 8-4**A).[441] Human cells (for example, embryonic kidney cells, HEK293T) will be co-transfected with gapped plasmids (unmodified and BD-dA lesions) and a carrier plasmid.[440,441] After allowing for gap filling and replication, filled-in progeny plasmids will be isolated using alkali and transformed into *E. coli recA* strain. Progeny cells will be plated in parallel on LB+kan and LB+cm to select for GP21 plasmids that underwent TLS and GP20 control plasmids (**Scheme 8-4**A).[440,441] Lesion bypass is the ratio of kan^R/cm^R transformants, and mutation analysis will be done by sequencing the kan^R transformants.[440,441] Alternatively, 18-mer DNA strands (5'-TCA T***X***G AAT CCT TCC CCC-3', where ***X*** = dA or BD-dA adducts) will be ligated to a single-stranded pMS2 vector linearized using *Eco*RV with the aid of a scaffold DNA (**Scheme 8-4**B).[442-445] Human embryonic kidney cells (HEK293T) will be transfected with the adducted and control plasmids.[442] Following replication and transformation in *E. coli* HD10B strain, progeny plasmids will be analyzed by oligonucleotide

hybridization and DNA sequencing (**Scheme 8-4**B).[442-445] *In vivo* mutagenesis of BD-dA adducts, if conducted using biologically more relevant cell types such as human bronchial epithelial cells (HBEC), will provide more useful information compared to bacterial systems to understand the adverse effects of BD-mediated DNA damage in humans.

Scheme 8-4 *In vivo* mutagenesis studies of 1,3-butadiene-induce 2'-deoxyadenosine adducts. Quantitative gap filling assay (A)[440,441] and single-stranded plasmid-based oligo hybridization assay (B).[442]

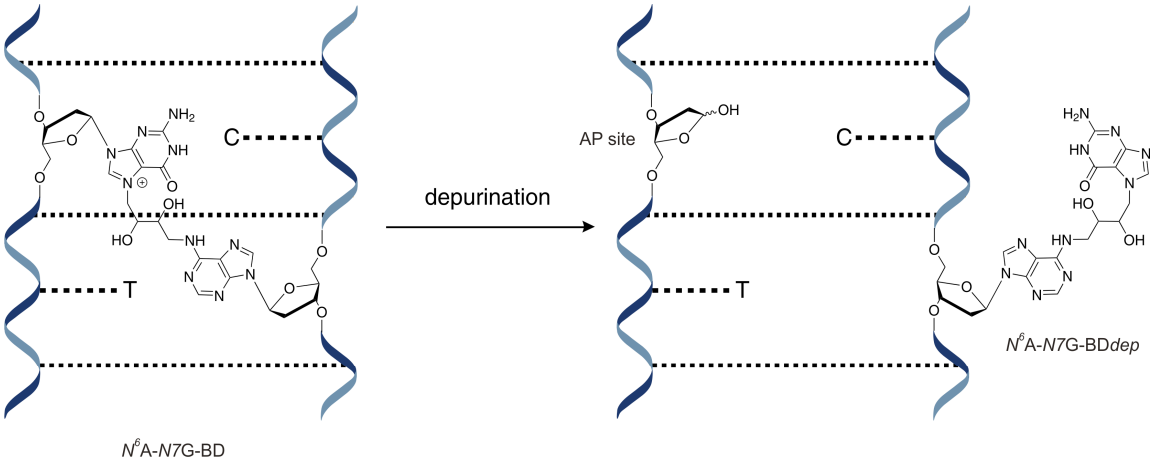


8.4 SYNTHESIS OF SITE- AND STEREO-SPECIFIC OLIGO-DEOXYNUCLEOTIDES CONTAINING 1-(ADEN-N^6-YL)-4-(GUAN-7-YL)-BUTAN-2,3-DIOL FOR STRUCTURAL AND BIOLOGICAL STUDIES

BD and its epoxide metabolites induce a large number of A → T mutations (section 1.2.4).[90-93] The exact structures and/or mechanisms responsible for these mutagenic effects are not well understood. 1-(Aden-N^6-yl)-4-(guan-7-yl)-butan- 2,3-diol (N^6A-*N7*G-BD) is a partially hydrolytically stable bulky DNA adducts induced by DEB.[446] The sequential alkylation of *N7*G and N^6A by DEB forms N^6A-*N7*G-BD adducts (**Scheme 8-5**). The resulting positive charge on alkylated guanine promotes the hydrolysis of the *N*-glycosidic bond to generate a partially depurinated bulky N^6A-*N7*G-BD*dep* adducts protruding out of the DNA major groove (**Scheme 8-5**). Hydrolytically stable N^6A-*N7*G-BD*dep* adducts can have significant effects on DNA structure, stability, repair and replication as observed with similar bulky adenine lesions. For example, site-specifically adducted DNA containing N^6-(14*R*-(+)-*trans,anti*-((11*S*,12*R*,13*R*)-trihydroxy-11,12,13,14-tetrahydrodibenzo[*a,l*]pyrenyl)-dA and N^6-(14*S*-(-)-*trans,anti*-((11*R*,12*S*,13*S*)-trihydroxy-11,12,13,14-tetrahydrodibenzo[*a,l*]pyrenyl)-dA were found to intercalate differentially (5'- and 3'-intercalation from the major groove of DNA, respectively), leading to significant differences in thermal stabilities.[447] However, both stereoisomers were resistant to nucleotide excision repair, when subjected to NER dual incision assay using HeLa cell free extracts.[447] Further, N^6-(10-(7,8,9-trihydroxy-7,8,9,10-tetrahydrobenzo[*a*]pyrenyl))-dA adducts represented a significant block to

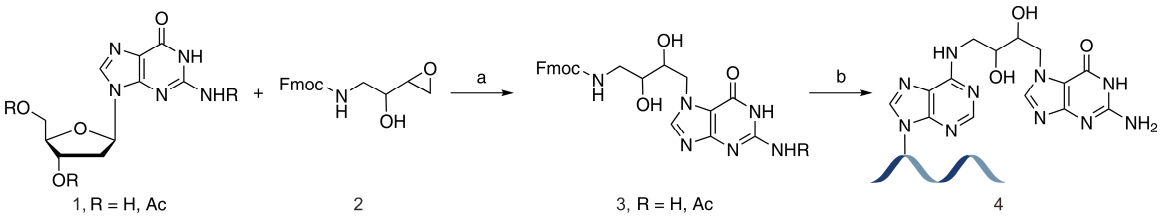
human bypass polymerases *in vitro*,[448] and produced mutagenic products with recombinant polymerases and human cell extracts.[448,449] Similarly, N^6-[(chrysene-5-yl)methyl]-adenine has revealed sequence-specific replication block with *E. coli* DNA polymerase I.[450] These results suggest that N^6A-*N7*G-BD*dep* adducts can alter the DNA structure contributing to adverse effects on DNA transactions.

Scheme 8-5 Formation of hydrolytically stable 1-(adenosin-N^6-yl)-4-(guan-7-yl)-butan-2,3-diol adduct.



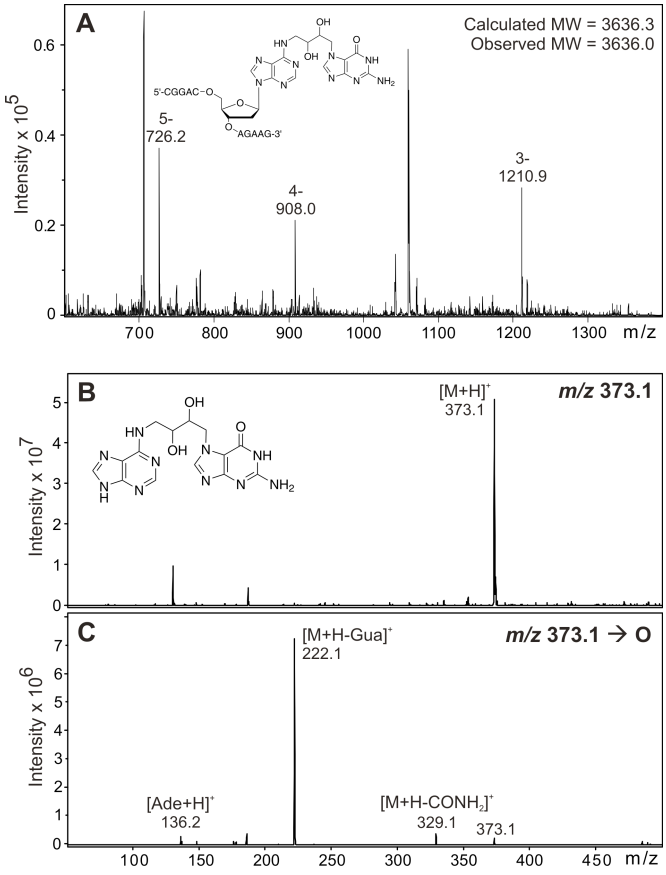
We have developed the synthetic methodologies to prepare site-specifically modified synthetic DNA containing N^6A-*N7*G-BD*dep* (**Scheme 8-6** and **Figure 8-2**). *N*-Fmoc-1-amino-3,4-epoxybutan-2-ol (2) was coupled to either 2'-deoxy-N^2,3',5'-*O*-tri-acetylguanosine (1, R=Ac) or 2'-deoxyguanosine (1, R=H) in an acidic medium to generate *N*-Fmoc-4-(N^2-acetylguan-*N7*-yl)-1-aminobutane-2,3-diol (3, R=Ac) and *N*-Fmoc-4-(guan-*N7*-yl)-1-amino-butane-2,3-diol (3, R=H), respectively (**Scheme 8-6**). The resulting *N7*G adduct was coupled with 6-chloropurine containing DNA in the presence of Hunig's base in DMSO to generate the N^6A-*N7*G-BD adducted oligodeoxynucleotides (4, **Scheme 8-6**). This synthetic methodology will be used to prepare site-specific DNA strands containing bulky N^6A-*N7*G-BD*dep* adducts for structural studies including solution state NMR, CD spectroscopic and UV melting studies to evaluate the effect of this bulky lesion on DNA duplex structure and stability. Further, replication bypass and repair experiments will be performed to elucidate its consequences on genomic stability and cell viability.

Scheme 8-6 Synthesis of oligodeoxynucleotides containing 1-(adenosin-N^6-yl)-4-(guan-7-yl)-butan-2,3-diol.[***]



[***] (a) AcOH, 80 °C, 5 h, (b) 6-Cl-Pu oligodeoxynucleotide, DIPEA, DMSO, rt, 3 d, (c) 0.1 M NaOH, rt, 3 d.

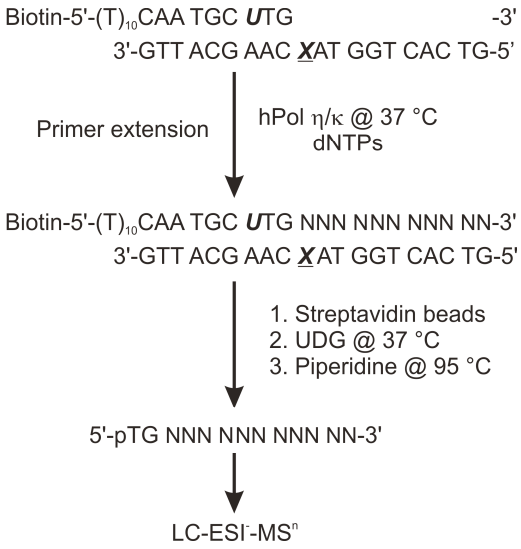
Figure 8-2 Capillary HPLC-ESI$^-$-MS spectrum of a 11-mer, 5'-CGG AC***X*** AGA AG-3', containing 1-(adenosin-N^6-yl)-4-(guan-7-yl)-butan-2,3-diol (A) and capillary HPLC-ESI$^+$-MS (B) and MS/MS spectra (C) of 1-(adenin-N^6-yl)-4-(guan-7-yl)-butan-2,3-diol.



8.5 FURTHER INVESTIGATION OF *IN VITRO* REPLICATION OF MODEL DNA-PROTEIN CROSS-LINKS

As discussed in Chapter 4, the decapeptide cross-linked to *C7* of 7-deaza-guanine in DNA was bypassed with high fidelity by human TLS polymerases η and κ. In contrast, the post-lesion bypass of the same peptide conjugated to *C5* of thymidine was highly mutagenic (Chapter 6). Therefore, it is critical to test the post-lesion bypass of *C7*-G conjugated peptide, since further extension past the damaged site may be mutagenic. To this end, quantification and sequencing of the replication products by a similar strategy used in Chapter 6 is currently underway to determine the fidelity of replication past *N7*-G adducted DNA (**Scheme 8-7**). Biotinylated 19-mer primer, Biotin-5'-$(T)_{10}$ CAA TGC **U**TG-3', containing a uridine at the 3rd position from the 3'-end will be annealed to the 20-mer, GTT ACG AAC ***X***AT GGT CAC TG-5', where ***X*** = dG or deaza-DHP-dG-decapeptide cross-link. The resulting primer-template complexes will be subjected to primer extension using hPol η and κ, subsequently captured on streptavidin beads and washed to remove enzymes and buffer solution. Incubation with UDG followed by heating in the presence of piperidine will release the extended primer (5'-pTG NNN NNN NNN NN-3' and insertion/deletion products, if any), which will be dried *in vacuo*. The samples will be reconstituted in water and a 14-mer will be added as an internal standard, immediately prior to analysis. Quantitation and sequencing of primer extension products will be performed by separate capillary HPLC-ESI^--FTMS and MS/MS analyses.

Scheme 8-7 Streptavidin capture in combination with capillary HPLC-ESI$^-$-MS/MS methodology for sequencing and quantitation of primer extension products.

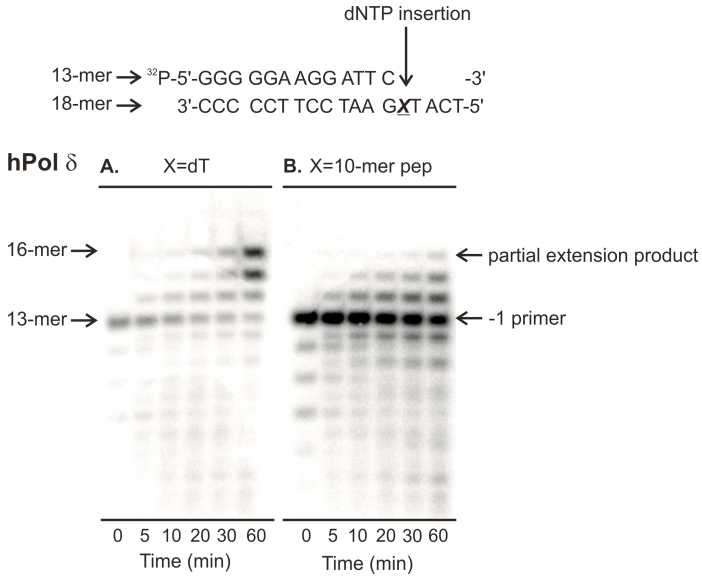


Although it is rare, replicative polymerases can also bypass DNA adducts. For example, Lloyd and coworkers reported that *E. coli* replicative polymerase, Pol III catalyzed efficient bypass of N^6-(1-hydroxy-3-buten-2-yl)-dA, a regiosiomer of the HB-dA adduct studied in this work, and N^6-(2,3,4-trihydroxybutan-1-yl)-dA adducts.[70] It will be interesting to examine whether human replicative polymerases can replicate past DNA-protein cross-links. Our pilot experiments using the decapeptide cross-linked to *C5*-T revealed that human replicative polymerase δ can partially extend a primer-template duplex past the peptide cross-link with extremely low efficiency (**Figure 8-3**). We also observed some degradation of the primer, which is typical for hPol δ having 3' → 5' exonuclease activity.[451] These observations warrant further experiments using hPol δ. Ideally, the enzyme lacking the exonuclease activity should be expressed to conduct steady-state kinetic studies.

It will be interesting to investigate the cooperativity and synergy of replicative polymerases and translesion synthesis polymerases. The cooperativity of polymerases will be examined by conducting the primer extension assays with an enzyme that catalyze partial extension of adducted DNA, followed by the addition of another enzyme after a predetermined time.[112,157] Full extension products will be observed if the enzymes act cooperatively. Alternatively, primer extension assays will be performed in the presence of both polymerases to test whether these enzymes can act synergistically. Higher amounts of full extension products will be observed with the dual enzyme reaction compared to the reactions with individual enzymes or heat inactivation of one enzyme

after a preselected time followed by the addition of the other, confirming the synergy of the polymerases in replication past the adduct.[112,157]

Figure 8-3 Standing start assay for replication bypass of DNA-peptide cross-links by human replicative polymerase δ. 50 nM primer-template duplexes were incubated at 37 °C with 5 nM polymerase in the presence 150 μM individual dNTPs. The reactions were quenched at predetermined time points (0–60 min) and analyzed by 20% denaturing PAGE.



8.6 IDENTIFICATION OF REPAIR MECHANISMS RESPONSIBLE FOR REMOVAL OF DNA-PROTEIN CROSS-LINKS

As mentioned in section 1.6, mechanisms responsible for the repair of DPCs in cells are not fully understood. Previous studies have yielded contradicting data dependent on structure of model DPC substrates. This requires a systematic investigation of *in vivo* repair using biologically relevant, site-specific DPCs. To this end, model DPC substrates prepared in the present study (Chapters 3 and 5) can be employed.

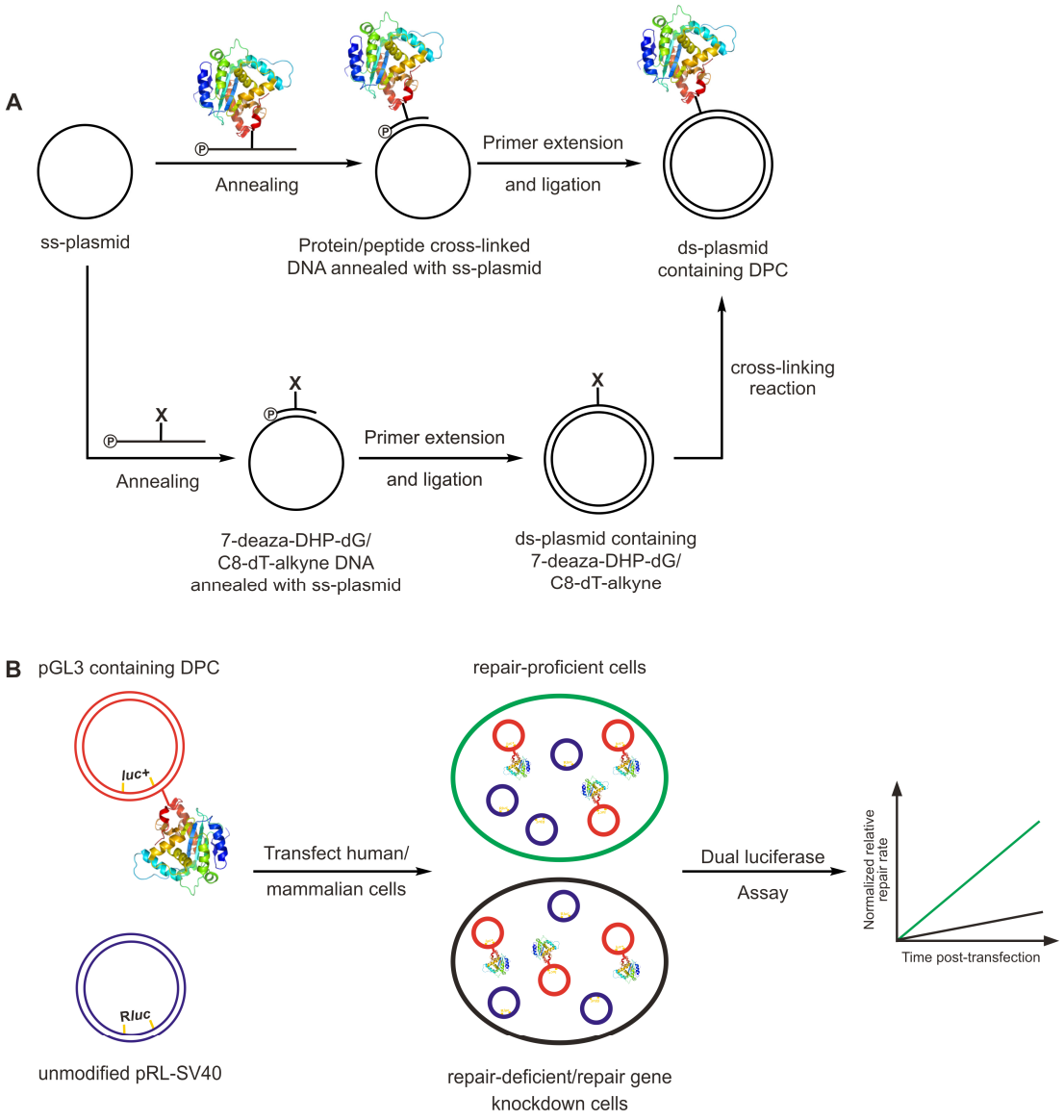
Repair experiments with phosphoramide mustard-induced DPCs in human cells has revealed that nucleotide excision repair pathway may be involved in removal of these bulky DNA adducts (Arnie Groehler and Natalia Tretyakova, unpublished data). Furthermore, the repair rates were slower upon inhibition of the proteasome suggesting the involvement of proteolytic processing of DPCs prior to repair (Arnie Groehler and Natalia Tretyakova, unpublished data).

To test whether NER and proteolytic degradation play a role in removing DPC lesions, *in vitro* repair experiments can be carried out using linear dsDNA containing cross-linked proteins and peptides of varying sizes (as described in section 8.2 for exocyclic N^6-dA adducts).

Further *in vivo* experiments can be performed using repair proficient and deficient human cell culture, and gene knockdown methods[452-455] to assess the effects of size and identity of cross-linked protein, site of cross-linking within DNA and proteins, and DNA sequence context on repair of these super bulky lesions. Host cell reactivation assay can

be employed for *in vivo* repair experiments (**Scheme 8-8**).[456] Proteins or peptides cross-linked to a 23-mer, 5'-CCA TGG TG***G***$^{\#}$ C***T****T TAC CAA CAG TA-3', containing either ***G***$^{\#}$ = 7-deaza-DHP-dG or ***T**** = C8-dT-alkyne, will be annealed to single-stranded pGL3 plasmids, followed by primer extension and ligation to generate double-stranded plasmids containing site-specific DPCs (**Scheme 8-8**A). Alternatively, proteins or peptides can be cross-linked to double-stranded plasmids containing 7-deaza-DHP-dG or C8-dT-alkyne (**Scheme 8-8**A). The 23-mer sequence, 5'-CCA TGG TG***G***$^{\#}$ C***T****T TAC CAA CAG TA-3', containing the site-specific DPC lesion is complementary to a sequence between the SV40 promoter and the *Photinus pyralis* (firefly) luciferase gene (*luc*+) on single-stranded pGL3 plasmid. Unmodified pRL-CMV plasmids expressing *Renilla reniformis* (sea pansy) luciferase protein will be used as an internal control. Human cells (wild-type and deficient in specific repair pathways) will be co-transfected with adducted and unadducted double-stranded plasmids. A dual luciferase assay will be used to quantify DPC repair (**Scheme 8-8**B).[457,458] Cell lines deficient in specific repair pathways (BER, NER, MMR, HR and NHEJ) can be obtained from Coriell Institute for Medical Research (Camden, NJ) or Trevigen (Gaithersburg, MD). These cell lines can be used to identify the repair mechanisms involved in the removal of these bulky DNA lesions, since relative repair rates will be significantly lower compared to those of wild-type cells.[459,460] Using short interference RNA (siRNA) to knockdown specific genes will be beneficial to confirm the repair pathways involved, since repair rates will decrease considerably in the presence of appropriate siRNAs.[452-455]

Scheme 8-8 Preparation of double-stranded pGL3 plasmids containing site-specific DNA-protein cross-links (A) and host cell reactivation assay to study repair (B).



8.7 INVESTIGATION OF MUTAGENICITY OF DNA-PROTEIN CROSS-LINKS IN CELLS

High fidelity bypass of an *C7*-G cross-linked peptide (Chapter 4) and highly mutagenic replication of the same peptide adducted to *C5*-T on DNA (Chapter 6) by TLS polymerases merit further investigation of the mutagenic potential of these lesions in cells. *In vivo* mutagenesis studies are currently in progress (in collaboration with Dr. Ashis Basu, University of Connecticut). A 6×His-eGFP cross-linked 23-mer, 5'-AGG GTT TTC CCA G**X**C ACG ACG TT-3', was ligated to a single-stranded pMS2 vector, and human embryonic kidney cells (HEK293T) were transfected with the adducted plasmids (**Scheme 8-4**B).[442,443] Following replication and transformation, progeny were analyzed by oligonucleotide hybridization and DNA sequencing.[442,443] In preliminary experiments, we found ~7% of recovered progeny to contain both targeted and semi-targeted mutations (**Table 8-1**). A systematic analysis of *in vivo* mutagenesis using polypeptides of varying sizes cross-linked to *N7*-G and *C5*-T prepared in this thesis work (Chapters 3 and 5) will enhance the current understanding of the effects of DPCs on DNA replication in cells.

Table 8-1 Mutation analysis of ss-pMS2 plasmids containing 6×His-eGFP cross-link replicated in HEK293T cells.[†††]

Replication of a ss-pMS2 plasmid containing a site-specific DPC in HEK 293T cells

Targeted mutations (3.2%)	Semi-targeted mutations (3.9%)
CCA G***X***C ACG → CCA G**C**C ACG (4/154)	CCA G***X***C A**C**G → CCA GTC A**T**G (2/154)
CCA G***X***C ACG → CCA G_C ACG (4/154)	**C**CA G***X***C ACG → **T**CA GTC ACG (2/154)
	CCA **G*X***C ACG → CCA **T**TC ACG (1/154)
	CCA **G*X***C ACG → CCA **A**TC ACG (1/154)

[†††] ***X*** = 6×His-eGFP cross-linked to *C5*-T.

8.8 STRUCTURE ELUCIDATION OF DNA-PROTEIN CROSS-LINKS

Replication bypass studies described in this thesis (Chapters 4–6) discovered that the size of the cross-linked polypeptide, the cross-linking site within DNA, and the identity of the polymerase play an important role in determining the efficiency and fidelity of replication past DNA-protein cross-links. It is intriguing that the same peptide cross-linked to major groove of DNA *via* a purine vs. a pyrimidine nucleobase has different effect on replication fidelity. It is also unclear how these bulky adducts are accommodated in polymerase active site. Therefore, it will be beneficial to obtain structures of human TLS polymerases in the presence of DPCs and dNTPs to better understand these observations. NMR[461] or X-ray crystallographic techniques[148] can be employed in this regard. Sequences like 5'-pGGC GAA GCC GGG TGC GAA GCA CC-3'dd for NMR[461] and 5'-TTC ATT AGT CCT TCC CCC-3' for X-ray crystallography[148] that have been successfully used to obtain the structures of ternary complexes (the enzyme, dducted primer template duplex and the incoming dNTP) can be modified to contain either *C7*-G or *C5*-T cross-linked decapeptide for these studies. The resulting structural data will aid in fully comprehending the biological consequences of DPCs.

9 BIBLIOGRAPHY

1. Geacintov, N. E. and Broyde, S. (2010) *The chemical biology of DNA damage* Wiley, Hoboken, NJ.

2. Borges, H. L., Linden, R., and Wang, J. Y. (2008) DNA damage-induced cell death: lessons from the central nervous system, *Cell Res. 18*, 17-26.

3. Clancy, S. (2008) DNA damage & repair: mechanisms for maintaining DNA integrity, *Nat. Educ. 1*, 103.

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

5. Ide, H., Shoulkamy, M. I., Nakano, T., Miyamoto-Matsubara, M., and Salem, A. M. H. (2011) Repair and biochemical effects of DNA-protein crosslinks, *Mutat. Res. 711*, 113-122.

6. Hegde, M. L., Hazra, T. K., and Mitra, S. (2008) Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells, *Cell Res. 18*, 27-47.

7. Li, G. M. (2008) Mechanisms and functions of DNA mismatch repair, *Cell Res. 18*, 85-98.

8. Hoeijmakers, J. H. J. (2001) Genome maintenance mechanisms for preventing cancer, *Nature 411*, 366-374.

9. Ghosal, G. and Chan, J. (2013) DNA damage tolerance: a double-edge sword guarding the genome, *Transl. Cancer Res. 2*, 107-129.

10. Shrivastav, N., Li, D., and Essigmann, J. M. (2010) Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation, *Carcinogenesis 31*, 59-70.

11. Houtgraaf, J. H., Versmissen, J., and van der Giessen, W. J. (2006) A concise review of DNA damage checkpoints and repair in mammalian cells, *Cardiovasc. Revasc. Med. 7*, 165-172.

12. Iyama, T. and Wilson, D. M., III (2013) DNA repair mechanisms in dividing and non-dividing cells, *DNA Repair (Amst) 12*, 620-636.

13. Sancar, A., Lindsey-Boltz, L. A., Unsal-Kacmaz, K., and Linn, S. (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints, *Annu. Rev. Biochem. 73*, 39-85.

14. Hoeijmakers, J. H. (2009) DNA damage, aging, and cancer, *N. Engl. J. Med. 361*, 1475-1485.

15. De Bont, R. and van Larebeke, N. (2004) Endogenous DNA damage in humans: a review of quantitative data, *Mutagenesis 19*, 169-185.

16. Sugiyama, H., Fujiwara, T., Ura, A., Tashiro, T., Yamamoto, K., Kawanishi, S., and Saito, I. (1994) Chemistry of thermal degradation of abasic sites in DNA. Mechanistic investigation on thermal DNA strand cleavage of alkylated DNA, *Chem. Res. Toxicol. 7*, 673-683.

17. Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *Nature 362*, 709-715.

18. Marnett, L. J. (2000) Oxyradicals and DNA damage, *Carcinogenesis 21*, 361-370.

19. Cadet, J., Berger, M., Douki, T., and Ravanat, J. L. (1997) Oxidative damage to DNA: formation, measurement, and biological significance, *Rev. Physiol Biochem. Pharmacol. 131*, 1-87.

20. Apel, K. and Hirt, H. (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction, *Annu. Rev. Plant Biol 55*, 373-399.

21. Burney, S., Caulfield, J. L., Niles, J. C., Wishnok, J. S., and Tannenbaum, S. R. (1999) The chemistry of DNA damage from nitric oxide and peroxynitrite, *Mutat. Res. 424*, 37-49.

22. Ravanat, J. L., Douki, T., and Cadet, J. (2001) Direct and indirect effects of UV radiation on DNA and its components, *J. Photochem. Photobiol. B 63*, 88-102.

23. Irigaray, P. and Belpomme, D. (2010) Basic properties and molecular mechanisms of exogenous chemical carcinogens, *Carcinogenesis 31*, 135-148.

24. Wogan, G. N., Hecht, S. S., Felton, J. S., Conney, A. H., and Loeb, L. A. (2004) Environmental and chemical carcinogenesis, *Semin. Cancer Biol 14*, 473-486.

25. Hecht, S. S. (1999) Tobacco smoke carcinogens and lung cancer, *J. Natl. Cancer Inst. 91*, 1194-1210.

26. Hecht, S. S. (2003) Tobacco carcinogens, their biomarkers and tobacco-induced cancer, *Nat. Rev. Cancer 3*, 733-744.

27. Reedijk, J. (1987) The mechanism of action of platinum anti-tumor drugs, *Pure & Appl. Chem. 59*, 181-192.

28. Tomasz, M., Chawla, A. K., and Lipman, R. (1988) Mechanism of monofunctional and bifunctional alkylation of DNA by mitomycin C, *Biochemistry 27*, 3182-3187.

29. Kohn, K. W. (1977) Interstrand cross-linking of DNA by 1,3-bis(2-chloroethyl)-1-nitrosourea and other 1-(2-haloethyl)-1-nitrosoureas, *Cancer Res. 37*, 1450-1454.

30. Wang, K., Zhou, Y., Wu, K., Ding, S., and Yang X (2009) Studies on formation of liquid and gaseous formaldehyde-induced DNA-protein crosslinks in rat marrow cells, *Asian J. Ecotoxicol. 4*, 780-785.

31. Waris, G. and Ahsan, H. (2006) Reactive oxygen species: role in the development of cancer and various chronic conditions, *J. Carcinog. 5*, 14.

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

33. Lu, K., Collins, L. B., Ru, H., Bermudez, E., and Swenberg, J. (2010) Distribution of DNA adducts caused by inhaled formaldehyde is consistent with induction of nasal carcinoma but not leukemia, *Toxicol. Sci. 116*, 441-451.

34. De Flora, S., Izzotti, A., Randerath, K., Randerath, E., Bartsch, H., Nair, J., Balansky, R., van Schooten, F., Degan, P., Fronza, G., Walsh, D., and Lewtas, J. (1996) DNA adducts and chronic degenerative disease. Pathogenetic relevance and implications in preventive medicine, *Mutat. Res. 366*, 197-238.

35. Ames, B. N., Shigenaga, M. K., and Hagen, T. M. (1993) Oxidants, antioxidants, and the degenerative diseases of aging, *Proc. Natl. Acad. Sci. U. S. A. 90*, 7915-7922.

36. Zahn, R. K., Zahn-Daimler, G., Ax, S., Hosokawa, M., and Takeda, T. (1999) Assessment of DNA-protein crosslinks in the course of aging in two mouse strains by use of a modified alkaline filter elution applied to whole tissue samples, *Mech. Ageing Dev. 108*, 99-112.

37. Martin, L. J. (2008) DNA damage and repair: relevance to mechanisms of neurodegeneration, *J. Neuropathol. Exp. Neurol. 67*, 377-387.

38. Reddy, V. P., Zhu, X., Perry, G., and Smith, M. A. (2009) Oxidative stress in diabetes and Alzheimer's disease, *J. Alzheimers. Dis. 16*, 763-774.

39. Sang, N., Hou, L., Yun, Y., and Li, G. (2009) SO_2 inhalation induces protein oxidation, DNA-protein crosslinks and apoptosis in rat hippocampus, *Ecotox. Environ. Safe. 72*, 879-884.

40. Coussens, L. M. and Werb, Z. (2002) Inflammation and cancer, *Nature 420*, 860-867.

41. Dedon, P. C. and Tannenbaum, S. R. (2004) Reactive nitrogen species in the chemical biology of inflammation, *Arch. Biochem. Biophys. 423*, 12-22.

42. Brooks, P. J. (2002) DNA repair in neural cells: basic science and clinical implications, *Mutat. Res. 509*, 93-108.

43. Malik, Q. and Herbert, K. E. (2012) Oxidative and non-oxidative DNA damage and cardiovascular disease, *Free Radic. Res. 46*, 554-564.

44. Shukla, P. C., Singh, K. K., Yanagawa, B., Teoh, H., and Verma, S. (2010) DNA damage repair and cardiovascular diseases, *Can. J. Cardiol. 26 Suppl A*, 13A-16A.

45. White, W. C. (2007) Butadiene production process overview, *Chem. Biol. Interact. 166*, 10-14.

46. National Toxicology Program (2011) 1,3-Butadiene, *Rep. Carcinogen. 12*, 75-77.

47. Himmelstein, M. W., Acquavella, J. F., Recio, L., Medinsky, M. A., and Bond, J. A. (1997) Toxicology and epidemiology of 1,3-butadiene, *Crit. Rev. Toxicol. 27*, 1-108.

48. Pelz, N., Dempster, N. M., and Shore, P. R. (1990) Analysis of low molecular weight hydrocarbons including 1,3-butadiene in engine exhaust gases using an aluminum oxide porous-layer open-tubular fused-silica column, *J. Chromatogr. Sci. 28*, 230-235.

49. Hoffmann, D., Hoffmann, I., and el Bayoumy, K. (2001) The less harmful cigarette: a controversial issue. a tribute to Ernst L. Wynder, *Chem. Res. Toxicol. 14*, 767-790.

50. Brunnemann, K. D., Kagan, M. R., Cox, J. E., and Hoffmann, D. (1990) Analysis of 1,3-butadiene and other selected gas-phase components in cigarette mainstream and sidestream smoke by gas chromatography-mass selective detection, *Carcinogenesis 11*, 1863-1868.

51. Delzell, E., Sathiakumar, N., Hovinga, M., Macaluso, M., Julian, J., Larson, R., Cole, P., and Muir, D. C. (1996) A follow-up study of synthetic rubber workers, *Toxicology 113*, 182-189.

52. Clapp, R. W., Jacobs, M. M., and Loechler, E. L. (2008) Environmental and occupational causes of cancer: new evidence 2005-2007, *Rev. Environ. Health 23*, 1-37.

53. Cheng, H., Sathiakumar, N., Graff, J., Matthews, R., and Delzell, E. (2007) 1,3-Butadiene and leukemia among synthetic rubber industry workers: exposure-response relationships, *Chem. Biol Interact. 166*, 15-24.

54. Sielken, R. L., Jr., Valdez-Flores, C., Gargas, M. L., Kirman, C. R., Teta, M. J., and Delzell, E. (2007) Cancer risk assessment for 1,3-butadiene: dose-response modeling from an epidemiological perspective, *Chem. Biol. Interact. 166*, 140-149.

55. Delzell, E., Sathiakumar, N., Graff, J., and Matthews, R. (2005) Styrene and ischemic heart disease mortality among synthetic rubber industry workers, *J. Occup. Environ. Med. 47*, 1235-1243.

56. Graff, J. J., Sathiakumar, N., Macaluso, M., Maldonado, G., Matthews, R., and Delzell, E. (2005) Chemical exposures in the synthetic rubber industry and lymphohematopoietic cancer mortality, *J. Occup. Environ. Med. 47*, 916-932.

57. Macaluso, M., Larson, R., Delzell, E., Sathiakumar, N., Hovinga, M., Julian, J., Muir, D., and Cole, P. (1996) Leukemia and cumulative exposure to butadiene, styrene and benzene among workers in the synthetic rubber industry, *Toxicology 113*, 190-202.

58. Sathiakumar, N., Graff, J., Macaluso, M., Maldonado, G., Matthews, R., and Delzell, E. (2005) An updated study of mortality among North American synthetic rubber industry workers, *Occup. Environ. Med. 62*, 822-829.

59. Divine, B. J. and Hartman, C. M. (2001) A cohort mortality study among workers at a 1,3 butadiene facility, *Chem. Biol Interact. 135-136*, 535-553.

60. Agency for Toxic Substances and Disease Registry (1992) Toxicological profile for 1,3-butadiene, *Public Health Service,U. S. Department of Health and Human Services, Atlanta, GA*.

61. Melnick, R. L., Huff, J., Chou, B. J., and Miller, R. A. (1990) Carcinogenicity of 1,3-butadiene in C57BL/6 x C3H F1 mice at low exposure concentrations, *Cancer Res. 50*, 6592-6599.

62. Owen, P. E., Glaister, J. R., Gaunt, I. F., and Pullinger, D. H. (1987) Inhalation toxicity studies with 1,3-butadiene. 3. Two year toxicity/carcinogenicity study in rats, *Am. Ind. Hyg. Assoc. J 48*, 407-413.

63. Fowles, J. and Dybing, E. (2003) Application of toxicological risk assessment principles to the chemical constituents of cigarette smoke, *Tob. Control 12*, 424-430.

64. United States Environmental protection Agency (2002) Health assessment of 1,3-butadiene, *National Center for Environmental Assessment, Washignton, DC*.

65. International Agency for Research on Cancer (2008) 1,3-Butadiene, ethylene oxide and vinyl halides (vinyl fluoride, vinyl chloride and vinyl bromide), *IARC Monogr. Eval. Carcinog. Risks Hum. 97*, 3-471.

66. Duescher, R. J. and Elfarra, A. A. (1994) Human liver microsomes are efficient catalysts of 1,3-butadiene oxidation: evidence for major roles by cytochromes P450 2A6 and 2E1, *Arch. Biochem. Biophys. 311*, 342-349.

67. Csanady, G. A., Guengerich, F. P., and Bond, J. A. (1992) Comparison of the biotransformation of 1,3-butadiene and its metabolite, butadiene monoepoxide, by hepatic and pulmonary tissues from humans, rats and mice, *Carcinogenesis 13*, 1143-1153.

68. Krause, R. J., Sharer, J. E., and Elfarra, A. A. (1997) Epoxide hydrolase-dependent metabolism of butadiene monoxide to 3-butene-1,2-diol in mouse, rat, and human liver, *Drug Metab Dispos. 25*, 1013-1015.

69. Krause, R. J. and Elfarra, A. A. (1997) Oxidation of butadiene monoxide to meso- and (+/-)-diepoxybutane by cDNA-expressed human cytochrome P450s and by mouse, rat, and human liver microsomes: evidence for preferential hydration of meso-diepoxybutane in rat and human liver microsomes, *Arch. Biochem. Biophys. 337*, 176-184.

70. Carmical, J. R., Nechev, L. V., Harris, C. M., Harris, T. M., and Lloyd, R. S. (2000) Mutagenic potential of adenine N^6 adducts of monoepoxide and diolepoxide derivatives of butadiene, *Environ. Mol. Mutagen. 35*, 48-56.

71. van Sittert, N. J., Megens, H. J., Watson, W. P., and Boogaard, P. J. (2000) Biomarkers of exposure to 1,3-butadiene as a basis for cancer risk assessment, *Toxicol. Sci. 56*, 189-202.

72. Carmella, S. G., Chen, M., Han, S., Briggs, A., Jensen, J., Hatsukami, D. K., and Hecht, S. S. (2009) Effects of smoking cessation on eight urinary tobacco carcinogen and toxicant biomarkers, *Chem. Res. Toxicol. 22*, 734-741.

73. Tretyakova, N., Sangaiah, R., Yen, T. Y., Gold, A., and Swenberg, J. A. (1997) Adenine adducts with diepoxybutane: isolation and analysis in exposed calf thymus DNA, *Chem. Res. Toxicol. 10*, 1171-1179.

74. Tretyakova, N. Y., Sangaiah, R., Yen, T. Y., and Swenberg, J. A. (1997) Synthesis, characterization, and *in vitro* quantitation of N7-guanine adducts of diepoxybutane, *Chem. Res. Toxicol. 10*, 779-785.

75. Selzer, R. R. and Elfarra, A. A. (1999) In vitro reactions of butadiene monoxide with single- and double- stranded DNA: characterization and quantitation of several purine and pyrimidine adducts, *Carcinogenesis 20*, 285-292.

76. Citti, L., Gervasi, P. G., Turchi, G., Bellucci, G., and Bianchini, R. (1984) The reaction of 3,4-epoxy-1-butene with deoxyguanosine and DNA *in vitro*: synthesis and characterization of the main adducts, *Carcinogenesis 5*, 47-52.

77. Tretyakova, N. Y., Lin, Y. P., Upton, P. B., Sangaiah, R., and Swenberg, J. A. (1996) Macromolecular adducts of butadiene, *Toxicology 113*, 70-76.

78. 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*, 118-133.

79. Zhang, X. Y. and Elfarra, A. A. (2003) Identification and characterization of a series of nucleoside adducts formed by the reaction of 2'-deoxyguanosine and 1,2,3,4-diepoxybutane under physiological conditions, *Chem. Res. Toxicol. 16*, 1606-1615.

80. 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*, 839-847.

81. Park, S. and Tretyakova, N. (2004) Structural characterization of the major DNA-DNA cross-link of 1,2,3,4-diepoxybutane, *Chem. Res. Toxicol. 17*, 129-136.

82. Tretyakova, N., Livshits, A., Park, S., Bisht, B., and Goggin, M. (2007) Structural elucidation of a novel DNA-DNA cross-link of 1,2,3,4-diepoxybutane, *Chem. Res. Toxicol. 20*, 284-289.

83. Park, S., Anderson, C., Loeber, R., Seetharaman, M., Jones, R., and Tretyakova, N. (2005) Interstrand and intrastrand DNA-DNA cross-linking by 1,2,3,4-diepoxybutane: role of stereochemistry, *J. Am. Chem. Soc. 127*, 14355-14365.

84. Park, S., Hodge, J., Anderson, C., and Tretyakova NY (2004) Guanine-adenine DNA cross-linking by 1,2,3,4-diepoxybutane: potential basis for biological activity., *Chem. Res. Toxicol. 17*, 1638-1651.

85. 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*, 4356-4367.

86. Gherezghiher, T. B., Ming, X., Villalta, P. W., Campbell, C., and Tretyakova, N. Y. (2013) 1,2,3,4-Diepoxybutane-induced DNA–protein cross-linking in human fibrosarcoma (HT1080) cells, *J. Proteome Res. 12*, 2151-2164.

87. United States Environmental protection Agency (2009) Integrated Risk Information System (IRIS) on 1,3-Butadiene, *National Center for Environmental Assessment, Office of Research and Development, Washington, DC*.

88. Rice, J. M. and Boffetta, P. (2001) 1,3-Butadiene, isoprene and chloroprene: reviews by the IARC monographs programme, outstanding issues, and research priorities in epidemiology, *Chem. Biol. Interact. 135-136*, 11-26.

89. California Environmental Protection Agency (1997) Technical support document for the determination of noncancer chronic reference exposure levels. Draft for public comment, *Office of Environmental Health Hazard Assessment, Berkeley, CA*.

90. Recio, L., Steen, A. M., Pluta, L. J., Meyer, K. G., and Saranko, C. J. (2001) Mutational spectrum of 1,3-butadiene and metabolites 1,2-epoxybutene and 1,2,3,4-diepoxybutane to assess mutagenic mechanisms, *Chem. Biol. Interact. 135-136*, 325-341.

91. Carmical, J. R., Zhang, M., Nechev, L., Harris, C. M., Harris, T. M., and Lloyd, R. S. (2000) Mutagenic potential of guanine N^2 adducts of butadiene mono- and diolepoxide, *Chem. Res. Toxicol. 13*, 18-25.

92. Ma, H., Wood, T. G., Ammenheuser, M. M., Rosenblatt, J. I., and Ward, J. B., Jr. (2000) Molecular analysis of hprt mutant lymphocytes from 1, 3-butadiene-exposed workers, *Environ. Mol. Mutagen. 36*, 59-71.

93. Liu, S., Ao, L., Du, B., Zhou, Y., Yuan, J., Bai, Y., Zhou, Z., and Cao, J. (2008) HPRT mutations in lymphocytes from 1,3-butadiene-exposed workers in China, *Environ. Health Perspect. 116*, 203-208.

94. Kligerman, A. D., DeMarini, D. M., Doerr, C. L., Hanley, N. M., Milholland, V. S., and Tennant, A. H. (1999) Comparison of cytogenetic effects of 3,4-epoxy-1-butene and 1,2,3,4-diepoxybutane in mouse, rat and human lymphocytes following *in vitro* G_0 exposures, *Mutat. Res. 439*, 13-23.

95. Cochrane, J. E. and Skopek, T. R. (1994) Mutagenicity of butadiene and its epoxide metabolites: I. Mutagenic potential of 1,2-epoxybutene, 1,2,3,4-diepoxybutane and 3,4-epoxy-1,2- butanediol in cultured human lymphoblasts, *Carcinogenesis 15*, 713-717.

96. Sasiadek, M. and Chichlowska-Sliwinska, M. (1991) Genotoxic properties of 1,3-butadiene and its derivatives, *Med. Pr. 42*, 193-198.

97. Sasiadek, M., Norppa, H., and Sorsa, M. (1991) 1,3-Butadiene and its epoxides induce sister-chromatid exchanges in human lymphocytes in vitro, *Mutat. Res. 261*, 117-121.

98. Millard, J. T. and Wilkes, E. E. (2001) Diepoxybutane and diepoxyoctane interstrand cross-linking of the 5S DNA nucleosomal core particle, *Biochemistry 40*, 10677-10685.

99. Verly, W. G., Brakier, L., and Feit, P. W. (1971) Inactivation of the T7 coliphage by the diepoxybutane stereoisomers, *Biochim. Biophys. Acta 228*, 400-406.

100. Matagne, R. (1969) Induction of chromosomal aberrations and mutations with isomeric forms of L-threitol-1,4-bismethanesulfonate in plant materials, *Mutat. Res. 7*, 241-247.

101. Kim, M. Y., Tretyakova, N., and Wogan, G. N. (2007) Mutagenesis of the supF gene by stereoisomers of 1,2,3,4-diepoxybutane, *Chem. Res. Toxicol. 20*, 790-797.

102. Millard, J. T., Hanly, T. C., Murphy, K., and Tretyakova, N. (2006) The 5'-GNC site for DNA interstrand cross-linking is conserved for diepoxybutane stereoisomers, *Chem. Res. Toxicol. 19*, 16-19.

103. Fred, C., Kautiainen, A., Athanassiadis, I., and Tornqvist, M. (2004) Hemoglobin adduct levels in rat and mouse treated with 1,2,3,4-diepoxybutane, *Chem. Res. Toxicol. 17*, 785-794.

104. Osterman-Golkar, S., Kautiainen, A., Bergmark, E., Hakansson, K., and Maki-Paakkanen, J. (1991) Hemoglobin adducts and urinary mercapturic acids in rats as biological indicators of butadiene exposure, *Chem. Biol Interact. 80*, 291-302.

105. Perez, H. L., Lahdetie, J., Landin, H., Kilpelainen, I., Koivisto, P., Peltonen, K., and Osterman-Golkar, S. (1997) Haemoglobin adducts of epoxybutanediol from exposure to 1,3-butadiene or butadiene epoxides, *Chem. Biol. Interact. 105*, 181-198.

106. Boysen, G., Georgieva, N. I., Upton, P. B., Jayaraj, K., Li, Y., Walker, V. E., and Swenberg, J. A. (2004) Analysis of diepoxide-specific cyclic N-terminal globin adducts in mice and rats after inhalation exposure to 1,3-butadiene, *Cancer Res. 64*, 8517-8520.

107. Swenberg, J. A., Christova-Gueorguieva, N. I., Upton, P. B., Ranasinghe, A., Scheller, N., Wu, K. Y., Yen, T. Y., and Hayes, R. (2000) 1,3-butadiene: cancer, mutations, and adducts. Part V: Hemoglobin adducts as biomarkers of 1,3-butadiene exposure and metabolism, *Res. Rep. Health Eff. Inst.* 191-210.

108. Kotapati, S., Sangaraju, D., Esades, A., Hallberg, L., Walker, V. E., Swenberg, J. A., and Tretyakova, N. Y. (2014) Bis-butanediol-mercapturic acid (bis-BDMA) as a urinary biomarker of metabolic activation of butadiene to its ultimate carcinogenic species, *Carcinogenesis 35*, 1371-1378.

109. Lawley, P. D. and Brookes, P. (1967) Interstrand cross-linking of DNA by difunctional alkylating agents, *J. Mol. Biol. 25*, 143-160.

110. Swenberg, J. A., Koc, H., Upton, P. B., Georguieva, N., Ranasinghe, A., Walker, V. E., and Henderson, R. (2001) Using DNA and hemoglobin adducts to improve the risk assessment of butadiene, *Chem. Biol. Interact. 135-136*, 387-403.

111. Goggin, M., Sangaraju, D., Walker, V. E., Wickliffe, J., Swenberg, J. A., and Tretyakova, N. (2011) Persistence and repair of bifunctional DNA adducts in tissues of laboratory animals exposed to 1,3-butadiene by inhalation, *Chem. Res. Toxicol. 24*, 809-817.

112. Kotapati, S., Maddukuri, L., Wickramaratne, S., Seneviratne, U., Goggin, M., Pence, M. G., Villalta, P., Guengerich, F. P., Marnett, L., and Tretyakova, N. (2012) Translesion synthesis across 1,N^6-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2'-deoxyadenosine (1,N^6-γ-HMHP-dA) adducts by human and archebacterial DNA polymerases, *J. Biol Chem. 287*, 38800-38811.

113. Fu, D., Calvo, J. A., and Samson, L. D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents, *Nat. Rev. Cancer 12*, 104-120.

114. Christmann, M., Tomicic, M. T., Roos, W. P., and Kaina, B. (2003) Mechanisms of human DNA repair: an update, *Toxicology 193*, 3-34.

115. Mathews, L. A., Cabarcas, S. M., and Hurt, E. M. (2013) *DNA repair of cancer stem cells* Springer Netherlands.

116. Fousteri, M. and Mullenders, L. H. (2008) Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects, *Cell Res. 18*, 73-84.

117. McKinnon, P. J. (2009) DNA repair deficiency and neurological disease, *Nat. Rev. Neurosci. 10*, 100-112.

118. Shuck, S. C., Short, E. A., and Turchi, J. J. (2008) Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology, *Cell Res. 18*, 64-72.

119. Li, X. and Heyer, W. D. (2008) Homologous recombination in DNA repair and DNA damage tolerance, *Cell Res. 18*, 99-113.

120. Weterings, E. and Chen, D. J. (2008) The endless tale of non-homologous end-joining, *Cell Res. 18*, 114-124.

121. Yang, W. (2008) Structure and mechanism for DNA lesion recognition, *Cell Res. 18*, 184-197.

122. Wickliffe, J. K., Galbert, L. A., Ammenheuser, M. M., Herring, S. M., Xie, J., Masters, O. E., III, Friedberg, E. C., Lloyd, R. S., and Ward, J. B., Jr. (2006) 3,4-Epoxy-1-butene, a reactive metabolite of 1,3-butadiene, induces somatic mutations in Xpc-null mice, *Environ. Mol. Mutagen. 47*, 67-70.

123. Wickliffe, J. K., Herring, S. M., Hallberg, L. M., Galbert, L. A., Masters, O. E., III, Ammenheuser, M. M., Xie, J., Friedberg, E. C., Lloyd, R. S., Abdel-Rahman, S. Z., and Ward, J. B., Jr. (2007) Detoxification of olefinic epoxides and nucleotide excision repair of epoxide-mediated DNA damage: Insights from animal models examining human sensitivity to 1,3-butadiene, *Chem. Biol Interact. 166*, 226-231.

124. Kennedy, C. H., Catallo, W. J., Wilson, V. L., and Mitchell, J. B. (2009) Combustion products of 1,3-butadiene inhibit catalase activity and induce expression of oxidative DNA damage repair enzymes in human bronchial epithelial cells, *Cell Biol Toxicol. 25*, 457-470.

125. Hang, B. (2004) Repair of exocyclic DNA adducts: rings of complexity, *Bioessays 26*, 1195-1208.

126. Lee, C. Y., Delaney, J. C., Kartalou, M., Lingaraju, G. M., Maor-Shoshani, A., Essigmann, J. M., and Samson, L. D. (2009) Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG), *Biochemistry 48*, 1850-1861.

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

128. Carmical, J. R., Kowalczyk, A., Zou, Y., Van Houten, B., Nechev, L. V., Harris, C. M., Harris, T. M., and Lloyd, R. S. (2000) Butadiene-induced intrastrand DNA cross-links: a possible role in deletion mutagenesis, *J. Biol. Chem. 275*, 19482-19489.

129. Prakash, S., Johnson, R. E., and Prakash, L. (2005) Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function, *Annu. Rev. Biochem. 74*, 317-353.

130. Lehmann, A. R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J. F., Kannouche, P. L., and Green, C. M. (2007) Translesion synthesis: Y-family polymerases and the polymerase switch, *DNA Repair 6*, 891-899.

131. Sale, J. E., Lehmann, A. R., and Woodgate, R. (2012) Y-family DNA polymerases and their role in tolerance of cellular DNA damage, *Nat. Rev. Mol. Cell Biol. 13*, 141-152.

132. Kuo, H. K., Griffith, J. D., and Kreuzer, K. N. (2007) 5-Azacytidine induced methyltransferase-DNA adducts block DNA replication *in vivo*, *Cancer Res. 67*, 8248-8254.

133. Woodgate, R. (1999) A plethora of lesion-replicating DNA polymerases, *Genes Dev. 13*, 2191-2195.

134. Friedberg, E. C., Lehmann, A. R., and Fuchs, R. P. P. (2005) Trading places: how do DNA polymerases switch during translesion DNA synthesis?, *Mol. Cell 18*, 499-505.

135. Huen, M. S. and Chen, J. (2008) The DNA damage response pathways: at the crossroad of protein modifications, *Cell Res. 18*, 8-16.

136. McCulloch, S. D. and Kunkel, T. A. (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases, *Cell Res. 18*, 148-161.

137. Burgers, P. M., Koonin, E. V., Bruford, E., Blanco, L., Burtis, K. C., Christman, M. F., Copeland, W. C., Friedberg, E. C., Hanaoka, F., Hinkle, D. C., Lawrence, C. W., Nakanishi, M., Ohmori, H., Prakash, L., Prakash, S., Reynaud, C. A., Sugino, A., Todo, T., Wang, Z., Weill, J. C., and Woodgate, R. (2001) Eukaryotic DNA polymerases: proposal for a revised nomenclature, *J. Biol Chem. 276*, 43487-43490.

138. Ohmori, H., Friedberg, E. C., Fuchs, R. P. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z., and Woodgate, R. (2001) The Y-family of DNA polymerases, *Mol. Cell 8*, 7-8.

139. Waters, L. S., Minesinger, B. K., Wiltrout, M. E., D'Souza, S., Woodruff, R. V., and Walker, G. C. (2009) Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance, *Microbiol. Mol. Biol Rev. 73*, 134-154.

140. Marini, F., Kim, N., Schuffert, A., and Wood, R. D. (2003) POLN, a nuclear PolA family DNA polymerase homologous to the DNA cross-link sensitivity protein Mus308, *J. Biol. Chem. 278*, 32014-32019.

141. Cordeiro-Stone, M., Zaritskaya, L. S., Price, L. K., and Kaufmann, W. K. (1997) Replication fork bypass of a pyrimidine dimer blocking leading strand DNA synthesis, *J. Biol. Chem. 272*, 13945-13954.

142. Friedberg, E. C. (2005) Suffering in silence: the tolerance of DNA damage, *Nat. Rev. Mol. Cell Biol 6*, 943-953.

143. Goodman, M. F. (2002) Error-prone repair DNA polymerases in prokaryotes and eukaryotes, *Annu. Rev. Biochem. 71*, 17-50.

144. Yang, W. and Woodgate, R. (2007) What a difference a decade makes: insights into translesion DNA synthesis, *Proc. Natl. Acad. Sci. U. S. A. 104*, 15591-15598.

145. Kunkel, T. A. (2004) DNA replication fidelity, *J. Biol Chem. 279*, 16895-16898.

146. Yang, W. (2005) Portraits of a Y-family DNA polymerase, *FEBS Lett. 579*, 868-872.

147. Trincao, J., Johnson, R. E., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2001) Structure of the catalytic core of *S. cerevisiae* DNA polymerase η: implications for translesion DNA synthesis, *Mol. Cell 8*, 417-426.

148. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication, *Cell 107*, 91-102.

149. Mailand, N., Gibbs-Seymour, I., and Bekker-Jensen, S. (2013) Regulation of PCNA-protein interactions for genome stability, *Nat. Rev. Mol. Cell Biol 14*, 269-282.

150. Matsuda, T., Bebenek, K., Masutani, C., Hanaoka, F., and Kunkel, T. A. (2000) Low fidelity DNA synthesis by human DNA polymerase η, *Nature 404*, 1011-1013.

151. 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*, 1163-1170.

152. Goggin, M., Seneviratne, U., Swenberg, J. A., Walker, V. E., and Tretyakova, N. (2010) Column switching HPLC-ESI$^+$-MS/MS methods for quantitative analysis of exocyclic dA adducts in the DNA of laboratory animals exposed to 1,3-butadiene, *Chem. Res. Toxicol. 23*, 808-812.

153. Fernandes, P. H., Hackfeld, L. C., Kozekov, I. D., Hodge, R. P., and Lloyd, R. S. (2006) Synthesis and mutagenesis of the butadiene-derived *N3* 2'-deoxyuridine adducts, *Chem. Res. Toxicol. 19*, 968-976.

154. Kanuri, M., Nechev, L. V., Tamura, P. J., Harris, C. M., Harris, T. M., and Lloyd, R. S. (2002) Mutagenic spectrum of butadiene-derived *N1*-deoxyinosine adducts and N^6,N^6-deoxyadenosine intrastrand cross-links in mammalian cells, *Chem. Res. Toxicol. 15*, 1572-1580.

155. Scholdberg, T. A., Nechev, L. V., Merritt, W. K., Harris, T. M., Harris, C. M., Lloyd, R. S., and Stone, M. P. (2005) Mispairing of a site specific major groove (2S,3S)-N^6-(2,3,4-trihydroxybutyl)-2'-deoxyadenosyl DNA Adduct of butadiene diol epoxide with deoxyguanosine: formation of a dA(anti).dG(anti) pairing interaction, *Chem. Res. Toxicol. 18*, 145-153.

156. Fernandes, P. H. and Lloyd, R. S. (2007) Mutagenic bypass of the butadiene-derived 2'-deoxyuridine adducts by polymerases η and ζ, *Mutat. Res. 625*, 40-49.

157. Kotapati, S. (2013) Ethnic/racial differences in metabolism of 1,3-butadiene (BD) and influence of BD-DNA adducts on DNA replication, *PhD thesis, University of Minnesota*.

158. 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*, 1151-1162.

159. 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. 10*, 2785-2796.

160. 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*, 33826-33838.

161. Loeber, R., Michaelson, E., Fang, Q., Campbell, C., Pegg, A. E., and Tretyakova, N. (2008) Cross-linking of the DNA repair protein O^6-alkylguanine DNA alkyltransferase to DNA in the presence of antitumor nitrogen mustards, *Chem. Res. Toxicol. 21*, 787-795.

162. 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*, 1297-1303.

163. 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*, 1281-1296.

164. Nakano, T., Terato, H., Asagoshi, K., Masaoka, A., Mukuta, M., Ohyama, Y., Suzuki, T., Makino, K., and Ide, H. (2003) DNA-protein cross-link formation mediated by oxanine. A novel genotoxic mechanism of nitric oxide-induced DNA damage, *J. Biol. Chem. 278*, 25264-25272.

165. 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*, 1074-1081.

166. 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*, 10080-10081.

167. 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*, 703-709.

168. 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*, 681-686.

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

170. Dexheimer, T. S., Antony, S., Marchand, C., and Pommier, Y. (2008) Tyrosyl-DNA phosphodiesterase as a target for anticancer therapy, *Anticancer Agents Med. Chem. 8*, 381-389.

171. 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*, 976-984.

172. Pommier, Y., Kohlhagen, G., Pourquier, P., Sayer, J. M., Kroth, H., and Jerina, D. M. (2000) Benzo[*a*]pyrene diol epoxide adducts in DNA are potent suppressors of a normal topoisomerase I cleavage site and powerful inducers of other topoisomerase I cleavages, *Proc. Natl. Acad. Sci. U. S. A. 97*, 2040-2045.

173. Li, T. K. and Liu, L. F. (2001) Tumor cell death induced by topoisomerase-targeting drugs, *Annu. Rev. Pharmacol. Toxicol. 41*, 53-77.

174. Chen, S. H., Chan, N. L., and Hsieh, T. S. (2013) New mechanistic and functional insights into DNA topoisomerases, *Annu. Rev. Biochem. 82*, 139-170.

175. Cao, L., Alani, E., and Kleckner, N. (1990) A pathway for generation and processing of double-strand breaks during meiotic recombination in *S. cerevisiae*, *Cell 61*, 1089-1101.

176. Povirk, L. F. (2012) Processing of damaged DNA ends for double-strand break repair in mammalian cells, *ISRN. Mol. Biol 2012*, 1-16.

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

178. 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*, 3161-3162.

179. DeMott, M. S., Beyret, E., Wong, D., Bales, B. C., Hwang, J. T., Greenberg, M. M., and Demple, B. (2002) Covalent trapping of human DNA polymerase β by the oxidative DNA lesion 2-deoxyribonolactone, *J. Biol. Chem. 277*, 7637-7640.

180. Sung, J. S., DeMott, M. S., and Demple, B. (2005) Long-patch base excision DNA repair of 2-deoxyribonolactone prevents the formation of DNA-protein cross-links with DNA polymerase β, *J. Biol. Chem. 280*, 39095-39103.

181. 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*, 5660-5671.

182. Dizdaroglu, M. and Gajewski, E. (1989) Structure and mechanism of hydroxyl radical-induced formation of a DNA-protein cross-link involving thymine and lysine in nucleohistone, *Cancer Res. 49*, 3463-3467.

183. Dizdaroglu, M., Gajewski, E., Reddy, P., and Margolis, S. A. (1989) Structure of a hydroxyl radical induced DNA-protein cross-link involving thymine and tyrosine in nucleohistone, *Biochemistry 28*, 3625-3628.

184. Gajewski, E. and Dizdaroglu, M. (1990) Hydroxyl radical induced cross-linking of cytosine and tyrosine in nucleohistone, *Biochemistry 29*, 977-980.

185. Gajewski, E., Fuciarelli, A. F., and Dizdaroglu, M. (1988) Structure of hydroxyl radical-induced DNA-protein crosslinks in calf thymus nucleohistone *in vitro*, *Int. J. Radiat. Biol 54*, 445-459.

186. Ban, F., Lundqvist, M. J., Boyd, R. J., and Eriksson, L. A. (2002) Theoretical studies of the cross-linking mechanisms between cytosine and tyrosine, *J. Am. Chem. Soc. 124*, 2753-2761.

187. Suzuki, T., Nakamura, T., Yamada, M., Ide, H., Kanaori, K., Tajima, K., Morii, T., and Makino, K. (1999) Isolation and characterization of diazoate intermediate upon nitrous acid and nitric oxide treatment of 2'-deoxycytidine, *Biochemistry 38*, 7151-7158.

188. Suzuki, T., Yamada, M., Nakamura, T., Ide, H., Kanaori, K., Tajima, K., Morii, T., and Makino, K. (2000) Products of the reaction between a diazoate derivative of 2'-deoxycytidine and L-lysine and its implication for DNA-nucleoprotein cross-linking by NO or HNO_2, *Chem. Res. Toxicol. 13*, 1223-1227.

189. Takahashi, K. (1977) Further studies on the reactions of phenylglyoxal and related reagents with proteins, *J. Biochem. 81*, 403-414.

190. Shapiro, R., Cohen, B. I., Shiuey, S. J., and Maurer, H. (1969) On the reaction of guanine with glyoxal, pyruvaldehyde, and kethoxal, and the structure of the acylguanines. A new synthesis of N^2-alkylguanines, *Biochemistry 8*, 238-245.

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

192. 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*, 121-125.

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

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

195. Kehrer, J. P. and Biswal, S. S. (2000) The molecular effects of acrolein, *Toxicol. Sci. 57*, 6-15.

196. Ishii, T., Yamada, T., Mori, T., Kumazawa, S., Uchida, K., and Nakayama, T. (2007) Characterization of acrolein-induced protein cross-links, *Free Radic. Res. 41*, 1253-1260.

197. Kurtz, A. J. and Lloyd, R. S. (2003) 1,N^2-deoxyguanosine adducts of acrolein, crotonaldehyde, and trans-4-hydroxynonenal cross-link to peptides via Schiff base linkage, *J. Biol. Chem. 278*, 5970-5976.

198. Minko, I. G., Kozekov, I. D., Kozekova, A., Harris, T. M., Rizzo, C. J., and Lloyd, R. S. (2008) Mutagenic potential of DNA-peptide crosslinks mediated by acrolein-derived DNA adducts, *Mutat. Res. 637*, 161-172.

199. 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*, 433-449.

200. 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.

201. Loeber, R., Rajesh, M., Fang, Q., Pegg, A. E., and Tretyakova, N. (2006) Cross-linking of the human DNA repair protein O^6-alkylguanine DNA alkyltransferase to DNA in the presence of 1,2,3,4-diepoxybutane, *Chem. Res. Toxicol. 19*, 645-654.

202. Edara, S., Kanugula, S., and Pegg, A. E. (1999) Expression of the inactive C145A mutant human O^6-alkylguanine-DNA alkyltransferase in *E.coli* increases cell killing and mutations by *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine, *Carcinogenesis 20*, 103-108.

203. Abril, N. and Margison, G. P. (1999) Mammalian cells expressing *Escherichia coli* O^6-alkylguanine-DNA alkyltransferases are hypersensitive to dibromoalkanes, *Chem. Res. Toxicol. 12*, 544-551.

204. Abril, N., Luque-Romero, F. L., Christians, F. C., Encell, L. P., Loeb, L. A., and Pueyo, C. (1999) Human O^6-alkylguanine-DNA alkyltransferase: protection against alkylating agents and sensitization to dibromoalkanes, *Carcinogenesis 20*, 2089-2094.

205. Abril, N., Luque-Romero, F. L., Prieto-Alamo, M. J., Rafferty, J. A., Margison, G. P., and Pueyo, C. (1997) Bacterial and mammalian DNA alkyltransferases sensitize *Escherichia coli* to the lethal and mutagenic effects of dibromoalkanes, *Carcinogenesis 18*, 1883-1888.

206. Abril, N., Luque-Romero, F. L., Prieto-Alamo, M. J., Margison, G. P., and Pueyo, C. (1995) Ogt alkyltransferase enhances dibromoalkane mutagenicity in excision repair-deficient *Escherichia coli* K-12, *Mol. Carcinog. 12*, 110-117.

207. Liu, H., Xu-Welliver, M., and Pegg, A. E. (2000) The role of human O^6-alkylguanine-DNA alkyltransferase in promoting 1,2-dibromoethane-induced genotoxicity in *Escherichia coli*, *Mutat. Res. 452*, 1-10.

208. Liu, L., 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, *J. Biol. Chem. 279*, 4250-4259.

209. Liu, L., Williams, K. M., Guengerich, F. P., and Pegg, A. E. (2004) O^6-alkylguanine-DNA alkyltransferase has opposing effects in modulating the genotoxicity of dibromomethane and bromomethyl acetate, *Chem. Res. Toxicol. 17*, 742-752.

210. Valadez, J. G., Liu, L., 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, *Chem. Res. Toxicol. 17*, 972-982.

211. Bau, D. T., Wang, T. S., Chung, C. H., Wang, A. S. S., and Jan, K. Y. (2002) Oxidative DNA adducts and DNA-protein cross-links are the major DNA lesions induced by arsenite, *Environ. Health Perspect. 110*, 753-756.

212. Ramirez, P., Del Razo, L. M., Gutierrez-Ruiz, M. C., and Gonsebatt, M. E. (2000) Arsenite induces DNA-protein crosslinks and cytokeratin expression in the WRL-68 human hepatic cell line, *Carcinogenesis 21*, 701-706.

213. 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.

214. Chakrabarti, S. K., Bai, C., and Subramanian, K. S. (1999) DNA-protein crosslinks induced by nickel compounds in isolated rat renal cortical cells and its antagonism by specific amino acids and magnesium ion, *Toxicol. Appl. Pharmacol. 154*, 245-255.

215. 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*, 153-165.

216. 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*, 3197-3203.

217. Ross, W. E., Ewig, R. A., and Kohn, K. W. (1978) Differences between melphalan and nitrogen mustard in the formation and removal of DNA cross-links, *Cancer Res. 38*, 1502-1506.

218. 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*, 3954-3958.

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

220. 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*, 5172-5177.

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

222. Hawkins, R. B. (1976) The measurement of lonizing radiation-induced cross linkage of DNA and protein in bacteriophage, *Radiat. Res. 68*, 300-307.

223. Cress, A. E., Kurath, K. M., Stea, B., and Bowden, G. T. (1990) The crosslinking of nuclear protein to DNA using ionizing radiation, *J. Cancer Res. Clin. Oncol. 116*, 324-330.

224. Kirkinezos, I. G. and Moraes, C. T. (2001) Reactive oxygen species and mitochondrial diseases, *Semin. Cell Dev. Biol 12*, 449-457.

225. Niki, E., Yoshida, Y., Saito, Y., and Noguchi, N. (2005) Lipid peroxidation: mechanisms, inhibition, and biological effects, *Biochem. Biophys. Res. Commun. 338*, 668-676.

226. Martinez, M. C. and Andriantsitohaina, R. (2009) Reactive nitrogen species: molecular mechanisms and potential significance in health and disease, *Antioxid. Redox. Signal. 11*, 669-702.

227. Jena, N. R. (2012) DNA damage by reactive species: Mechanisms, mutation and repair, *Journal of Biosciences 37*, 503-517.

228. Roberts, M. J., Wondrak, G. T., Laurean, D. C., Jacobson, M. K., and Jacobson, E. L. (2003) DNA damage by carbonyl stress in human skin cells, *Mutat. Res. 522*, 45-56.

229. Murata-Kamiya, N. and Kamiya, H. (2001) Methylglyoxal, an endogenous aldehyde, crosslinks DNA polymerase and the substrate DNA, *Nucleic Acids Res. 29*, 3433-3438.

230. Shangari, N. and O'Brien, P. J. (2004) The cytotoxic mechanism of glyoxal involves oxidative stress, *Biochem. Pharmacol. 68*, 1433-1442.

231. Kalapos, M. P. (1999) Methylglyoxal in living organisms: chemistry, biochemistry, toxicology and biological implications, *Toxicol. Lett. 110*, 145-175.

232. Petrova, K. V., Millsap, A. D., Stec, D. F., and Rizzo, C. J. (2014) Characterization of the deoxyguanosine-lysine cross-link of methylglyoxal, *Chem. Res. Toxicol. 27*, 1019-1029.

233. Flyvholm, M. A. and Andersen, P. (1993) Identification of formaldehyde releasers and occurrence of formaldehyde and formaldehyde releasers in registered chemical products, *Am. J. Ind. Med. 24*, 533-552.

234. Zhao, W., Pen, G., and Yang, X. (2009) DNA-protein crosslinks induced by formaldehyde and its repair process, *Int. J. Environ. Pollut. 37*, 299-308.

235. 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*, 47-55.

236. European Food Safety Authority (2014) Endogenous formaldehyde turnover in humans compared with exogenous contribution from food sources, *EFSA Journal 12*, 3550-3560.

237. Metz, B., Kersten, G. F., Hoogerhout, P., Brugghe, H. F., Timmermans, H. A., de Jong, A., Meiring, H., ten Hove, J., Hennink, W. E., Crommelin, D. J., and Jiskoot, W. (2004) Identification of formaldehyde-induced modifications in proteins: reactions with model peptides, *J. Biol Chem. 279*, 6235-6243.

238. Lu, K., Ye, W., Zhou, L., Collins, L. B., Chen, X., Gold, A., Ball, L. M., and Swenberg, J. (2010) Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers, *J. Am. Chem Soc. 132*, 3388-3399.

239. 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*, 525-536.

240. 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*, 409-428.

241. Wilkins, R. J. and Macleod, H. D. (1976) Formaldehyde induced DNA-protein crosslinks in *Escherichia coli*, *Mutat. Res. 36*, 11-16.

242. Craft, T. R., Bermudez, E., and Skopek, T. R. (1987) Formaldehyde mutagenesis and formation of DNA-protein crosslinks in human lymphoblasts *in vitro*, *Mutat. Res. 176*, 147-155.

243. Miller, C. A., III and Costa, M. (1989) Analysis of proteins cross-linked to DNA after treatment of cells with formaldehyde, chromate, and cis-diamminedichloroplatinum(II), *Mol. Toxicol. 2*, 11-26.

244. 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*, 6391-6397.

245. Integrated Risk Information System (2003) Toxicological review of acrolein, *United States Environmental Protection Agency, Washington, DC*.

246. VanderVeen, L., Harris, T. M., Jen-Jacobson, L., and Marnett, L. J. (2008) Formation of DNA-protein cross-links between γ-hydroxypropanodeoxy-guanosine and EcoRI, *Chem. Res. Toxicol. 21*, 1733-1738.

247. Uchida, K., Kanematsu, M., Morimitsu, Y., Osawa, T., Noguchi, N., and Niki, E. (1998) Acrolein is a product of lipid peroxidation reaction. Formation of free acrolein and its conjugate with lysine residues in oxidized low density lipoproteins, *J. Biol. Chem. 273*, 16058-16066.

248. Loecken, E. M., Dasari, S., Hill, S., Tabb, D. L., and Guengerich, F. P. (2009) The *bis*-electrophile diepoxybutane cross-links DNA to human histones but does not result in enhanced mutagenesis in recombinant systems, *Chem. Res. Toxicol. 22*, 1069-1076.

249. Loecken, E. M. and Guengerich, F. P. (2008) Reactions of glyceraldehyde 3-phosphate dehydrogenase sulfhydryl groups with *bis*-electrophiles produce DNA-protein cross-links but not mutations, *Chem. Res. Toxicol. 21*, 453-458.

250. Xie, J., Fan, R., and Meng, Z. (2007) Protein oxidation and DNA-protein crosslink induced by sulfur dioxide in lungs, livers, and hearts from mice, *Inhal. Toxicol. 19*, 759-765.

251. Lei, Y. X., Zhang, Q., and Zhuang, Z. X. (1995) Study on DNA-protein crosslinks induced by chromate and nickel compounds in vivo with ^{125}I-postlabelling assay, *Mutat. Res. 329*, 197-203.

252. Olin, K. L., Cherr, G. N., Rifkin, E., and Keen, C. L. (1996) The effects of some redox-active metals and reactive aldehydes on DNA-protein cross-links *in vitro*, *Toxicology 110*, 1-8.

253. Kuykendall, J. R., Miller, K. L., Mellinger, K. M., Cain, A. J., Perry, M. W., Bradley, M., Jarvi, E. J., and Paustenbach, D. J. (2009) DNA-protein cross-links in erythrocytes of freshwater fish exposed to hexavalent chromium or divalent nickel, *Arch. Environ. Contam Toxicol. 56*, 260-267.

254. 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*, 987-997.

255. 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*, 33550-33560.

256. Macfie, A., Hagan, E., and Zhitkovich, A. (2010) Mechanism of DNA-protein cross-linking by chromium, *Chem. Res. Toxicol. 23*, 341-347.

257. Kasprzak, K. S., Sunderman, F. W., Jr., and Salnikow, K. (2003) Nickel carcinogenesis, *Mutat. Res. 533*, 67-97.

258. Peak, J. G. and Peak, M. J. (1991) Comparison of initial yields of DNA-to-protein crosslinks and single-strand breaks induced in cultured human cells by far- and near-ultraviolet light, blue light and X-rays, *Mutat. Res. 246*, 187-191.

259. Gueranger, Q., Kia, A., Frith, D., and Karran, P. (2011) Crosslinking of DNA repair and replication proteins to DNA in cells treated with 6-thioguanine and UVA, *Nucleic Acids Res. 39*, 5057-5066.

260. 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 8*, 135-140.

261. Vermund, H. and Gollin, F. F. (1968) Mechanisms of action of radiotherapy and chemotherapeutic adjuvants. A review, *Cancer 21*, 58-76.

262. Rajski, S. R. and Williams, R. M. (1998) DNA cross-linking agents as antitumor drugs, *Chem Rev. 98*, 2723-2796.

263. Povirk, L. F. and Shuker, D. E. (1994) DNA damage and mutagenesis induced by nitrogen mustards, *Mutat. Res. 318*, 205-226.

264. Caldecott, K. W. (2008) Single-strand break repair and genetic disease, *Nat. Rev. Genet. 9*, 619-631.

265. Khanna, K. K. and Jackson, S. P. (2001) DNA double-strand breaks: signaling, repair and the cancer connection, *Nat. Genet. 27*, 247-254.

266. Hurley, L. H. (2002) DNA and its associated processes as targets for cancer therapy, *Nat. Rev. Cancer 2*, 188-200.

267. Lawley, P. D. and Brookes, P. (1965) Molecular mechanism of the cytotoxic action of difunctional alkylating agents and of resistance to this action, *Nature 206*, 480-483.

268. Pallis, A. G. and Karamouzis, M. V. (2010) DNA repair pathways and their implication in cancer treatment, *29*, 677-685.

269. Hall, A. G. and Tilby, M. J. (1992) Mechanisms of action of, and modes of resistance to, alkylating agents used in the treatment of haematological malignancies, *Blood Rev. 6*, 163-173.

270. Pacheco, D. Y., Stratton, N. K., and Gibson, N. W. (1989) Comparison of the mechanism of action of busulfan with hepsulfam, a new antileukemic agent, in the L1210 cell line, *Cancer Res. 49*, 5108-5110.

271. Iwamoto, T., Hiraku, Y., Oikawa, S., Mizutani, H., Kojima, M., and Kawanishi, S. (2004) DNA intrastrand cross-link at the 5'-GA-3' sequence formed by busulfan and its role in the cytotoxic effect, *Cancer Sci. 95*, 454-458.

272. Silverman, R. B. (2004) *The organic chemistry of drug design and drug action* Elsevier Academic Press, Burlington, MA.

273. Lemke, T. L., Roche, V. F., Williams, D. A., Zito, S. W., and Foye, W. O. (2012) *Foye's principles of medicinal chemistry* Lippincott Williams & Wilkins, Baltimore, MD.

274. Wang, C. C., Li, J., Teo, C. S., and Lee, T. (1999) The delivery of BCNU to brain tumors, *J. Control Release 61*, 21-41.

275. Tong, W. P., Kirk, M. C., and Ludlum, D. B. (1982) Formation of the cross-link 1-[*N3*-deoxycytidyl),2-[*N1*-deoxyguanosinyl]ethane in DNA treated with *N*,*N'*-bis(2-chloroethyl)-*N*-nitrosourea, *Cancer Res. 42*, 3102-3105.

276. Seidenfeld, J., Barnes, D., Block, A. L., and Erickson, L. C. (1987) Comparison of DNA interstrand cross-linking and strand breakage by 1,3-bis(2-chloroethyl)-1-nitrosourea in polyamine-depleted and control human adenocarcinoma cells, *Cancer Res. 47*, 4538-4543.

277. Dirven, H. A., van Ommen, B., and van Bladeren, P. J. (1996) Glutathione conjugation of alkylating cytostatic drugs with a nitrogen mustard group and the role of glutathione S-transferases, *Chem. Res. Toxicol. 9*, 351-360.

278. Gilman, A. and Philips, F. S. (1946) The biological actions and therapeutic applications of the β-chloroethyl amines and sulfides, *Science 103*, 409-436.

279. Brookes, P. and Lawley, P. D. (1961) The alkylation of guanosine and guanylic acid, *J. Chem. Soc.* 3923-3927.

280. Balcome, S., Park, S., Quirk Dorr, D. R., Hafner, L., Phillips, L., and Tretyakova, N. (2004) Adenine-containing DNA-DNA cross-links of antitumor nitrogen mustards, *Chem. Res. Toxicol. 17*, 950-962.

281. Kohn, K. W., Hartley, J. A., and Mattes, W. B. (1987) Mechanisms of DNA sequence selective alkylation of guanine-N7 positions by nitrogen mustards, *Nucleic Acids Res. 15*, 10531-10549.

282. Rink, S. M., Solomon, M. S., Taylor, M. J., Rajur, S. B., McLaughlin, L. W., and Hopkins, P. B. (1993) Covalent structure of a nitrogen mustard-induced DNA interstrand cross-link: An *N7*-to-*N7* linkage of deoxyguanosine residues at the duplex sequence 5'-d(GNC), *J. Am. Chem. Soc. 115*, 2551-2557.

283. Osborne, M. R., Wilman, D. E., and Lawley, P. D. (1995) Alkylation of DNA by the nitrogen mustard bis(2-chloroethyl)methylamine, *Chem. Res. Toxicol. 8*, 316-320.

284. Haapala, E., Hakala, K., Jokipelto, E., Vilpo, J., and Hovinen, J. (2001) Reactions of *N*,*N*-bis(2-chloroethyl)-p-aminophenylbutyric acid (chlorambucil) with 2'-deoxyguanosine, *Chem. Res. Toxicol. 14*, 988-995.

285. 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*, 2114-2122.

286. Kelland, L. (2007) The resurgence of platinum-based cancer chemotherapy, *Nat. Rev. Cancer 7*, 573-584.

287. Donzelli, E., Carfi, M., Miloso, M., Strada, A., Galbiati, S., Bayssas, M., Griffon-Etienne, G., Cavaletti, G., Petruccioli, M. G., and Tredici, G. (2004) Neurotoxicity of platinum compounds: comparison of the effects of cisplatin and oxaliplatin on the human neuroblastoma cell line SH-SY5Y, *J. Neurooncol. 67*, 65-73.

288. Wheate, N. J., Walker, S., Craig, G. E., and Oun, R. (2010) The status of platinum anticancer drugs in the clinic and in clinical trials, *Dalton Trans. 39*, 8113-8127.

289. de Gramont, A., Figer, A., Seymour, M., Homerin, M., Hmissi, A., Cassidy, J., Boni, C., Cortes-Funes, H., Cervantes, A., Freyer, G., Papamichael, D., Le Bail, N., Louvet, C., Hendler, D., de Braud, F., Wilson, C., Morvan, F., and Bonetti, A. (2000) Leucovorin and fluorouracil with or without oxaliplatin as first-line treatment in advanced colorectal cancer, *J. Clin. Oncol. 18*, 2938-2947.

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

291. 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*, 1921-1926.

292. 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*, 6091-6095.

293. 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*, 201-208.

294. Ming, X. (2011) DNA-protein cross-linking by *cis*-1,1,2,2-diamminedichloro-platinum(II) (cisplatin), *MS thesis, University of Minnesota*.

295. Bjorklund, C. C. and Davis, W. B. (2007) Stable DNA-protein cross-links are products of DNA charge transport in a nucleosome core particle, *Biochemistry 46*, 10745-10755.

296. Toyokuni, S., Mori, T., Hiai, H., and Dizdaroglu, M. (1995) Treatment of Wistar rats with a renal carcinogen, ferric nitrilotriacetate, causes DNA-protein cross-linking between thymine and tyrosine in their renal chromatin, *Int. J. Cancer 62*, 309-313.

297. 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*, 1573-1580.

298. Kuykendall, J. R. and Bogdanffy, M. S. (1992) Reaction kinetics of DNA-histone crosslinking by vinyl acetate and acetaldehyde, *Carcinogenesis 13*, 2095-2100.

299. Luo, L., Jiang, L., Geng, C., Cao, J., and Zhong, L. (2008) Hydroquinone-induced genotoxicity and oxidative DNA damage in HepG2 cells, *Chem. Biol Interact. 173*, 1-8.

300. Neuss, S., Holzmann, K., and Speit, G. (2010) Gene expression changes in primary human nasal epithelial cells exposed to formaldehyde *in vitro*, *Toxicol. Lett. 198*, 289-295.

301. Wong, V. C., Cash, H. L., Morse, J. L., Lu, S., and Zhitkovich, A. (2012) S-phase sensing of DNA-protein crosslinks triggers TopBP1-independent ATR activation and p53-mediated cell death by formaldehyde, *Cell Cycle 11*, 2526-2537.

302. Nieves-Neira, W., Rivera, M. I., Kohlhagen, G., Hursey, M. L., Pourquier, P., Sausville, E. A., and Pommier, Y. (1999) DNA protein cross-links produced by NSC 652287, a novel thiophene derivative active against human renal cancer cells, *Molecular Pharmacology 56*, 478-484.

303. Yarema, K. J., Lippard, S. J., and Essigmann, J. M. (1995) Mutagenic and genotoxic effects of DNA adducts formed by the anticancer drug cis-diamminedichloroplatinum(II), *Nucleic Acids Res. 23*, 4066-4072.

304. Tretyakova, N. Y., Michaelson-Richie, E. D., Gherezghiher, T. B., Kurtz, J., Ming, X., Wickramaratne, S., Campion, M., Kanugula, S., Pegg, A. E., and Campbell, C. (2013) DNA-reactive protein monoepoxides induce cell death and mutagenesis in mammalian cells, *Biochemistry 52*, 3171-3181.

305. Voitkun, V., Zhitkovich, A., and Costa, M. (1998) Cr(III)-mediated crosslinks of glutathione or amino acids to the DNA phosphate backbone are mutagenic in human cells, *Nucleic Acids Res. 26*, 2024-2030.

306. Langevin, F., Crossan, G. P., Rosado, I. V., Arends, M. J., and Patel, K. J. (2011) Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice, *Nature 475*, 53-58.

307. Garaycoechea, J. I., Crossan, G. P., Langevin, F., Daly, M., Arends, M. J., and Patel, K. J. (2012) Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function, *Nature 489*, 571-575.

308. Costa, M., Zhitkovich, A., and Toniolo, P. (1993) DNA-protein cross-links in welders: molecular implications, *Cancer Res. 53*, 460-463.

309. Waksvik, H., Boysen, M., and Hogetveit, A. C. (1984) Increased incidence of chromosomal aberrations in peripheral lymphocytes of retired nickel workers, *Carcinogenesis 5*, 1525-1527.

310. Waksvik, H. and Boysen, M. (1982) Cytogenetic analyses of lymphocytes from workers in a nickel refinery, *Mutat. Res. 103*, 185-190.

311. Izzotti, a., Cartiglia, C., Taningher, M., De Flora, S., and Balansky, R. (1999) Age-related increases of 8-hydroxy-2'-deoxyguanosine and DNA-protein crosslinks in mouse organs, *Mutat. Res. 446*, 215-223.

312. Dunnick, J. K., Elwell, M. R., Radovsky, A. E., Benson, J. M., Hahn, F. F., Nikula, K. J., Barr, E. B., and Hobbs, C. H. (1995) Comparative carcinogenic effects of nickel subsulfide, nickel oxide, or nickel sulfate hexahydrate chronic exposures in the lung, *Cancer Res. 55*, 5251-5256.

313. Chen, C. J., Chen, C. W., Wu, M. M., and Kuo, T. L. (1992) Cancer potential in liver, lung, bladder and kidney due to ingested inorganic arsenic in drinking water, *Br. J. Cancer 66*, 888-892.

314. Bates, M. N., Smith, A. H., and Hopenhayn-Rich, C. (1992) Arsenic ingestion and internal cancers: a review, *Am. J. Epidemiol. 135*, 462-476.

315. 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*, 1146-1151.

316. Tsapakos, M. J., Hampton, T. H., and Wetterhahn, K. E. (1983) Chromium(VI)-induced DNA lesions and chromium distribution in rat kidney, liver, and lung, *Cancer Res. 43*, 5662-5667.

317. Sugiyama, M., Patierno, S. R., Cantoni, O., and Costa, M. (1986) Characterization of DNA lesions induced by $CaCrO_4$ in synchronous and asynchronous cultured mammalian cells, *Mol. Pharmacol. 29*, 606-613.

318. Mattagajasingh, S. N. and Misra, H. P. (1999) Analysis of EDTA-chelatable proteins from DNA-protein crosslinks induced by a carcinogenic chromium(VI) in cultured intact human cells, *Mol. Cell Biochem. 199*, 149-162.

319. Reedijk, J. (1999) Why does Cisplatin reach Guanine-*N7* with competing s-donor ligands available in the cell?, *Chem. Rev. 99*, 2499-2510.

320. Barnham, K. J., Djuran, M. I., Socorro Murdoch, P., 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, *J. Chem. Soc. Chem. Commun. 6*, 721-722.

321. van Boom, 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), *J. Chem. Soc. Chem. Commun.* 1397-1398.

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

323. Sancar, A. (1996) DNA excision repair, *Annu. Rev. Biochem. 65*, 43-81.

324. 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*, 1905-1909.

325. 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*, 3000-3009.

326. 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, *Mol. Cell 28*, 147-158.

327. 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*, 27065-27076.

328. Meyn, R. E., vanAnkeren, S. C., and Jenkins, W. T. (1987) The induction of DNA-protein crosslinks in hypoxic cells and their possible contribution to cell lethality, *Radiat. Res. 109*, 419-429.

329. Murray, D. and Rosenberg, E. (1996) The importance of the ERCC1/ERCC4[XPF] complex for hypoxic-cell radioresistance does not appear to derive from its participation in the nucleotide excision repair pathway, *Mutat. Res. 364*, 217-226.

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

331. 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*, 22592-22604.

332. Takahashi, K., Morita, T., and Kawazoe, Y. (1985) Mutagenic characteristics of formaldehyde on bacterial systems, *Mutat. Res. 156*, 153-161.

333. Nishioka, H. (1973) Lethal and mutagenic action of formaldehyde in Hcr + and Hcr - strains of Escherichia coli, *Mutat. Res. 17*, 261-265.

334. Novakova, O., Kasparkova, J., Malina, J., Natile, G., and Brabec, V. (2003) DNA-protein cross-linking by *trans*-[$PtCl_2$(E-iminoether)$_2$]. A concept for activation of the trans geometry in platinum antitumor complex, *Nucleic Acids Res. 31*, 6450-6460.

335. Minko, I. G., Yamanaka, K., Kozekov, I. D., Kozekova, A., Indiani, C., O'Donnell, M. E., Jiang, Q., Goodman, M. F., Rizzo, C. J., and Lloyd, R. S. (2008) Replication bypass of the acrolein-mediated deoxyguanine DNA-peptide cross-links by DNA polymerases of the DinB family, *Chem. Res. Toxicol. 21*, 1983-1990.

336. Yamanaka, K., Minko, I. G., Takata, K. i., Kolbanovskiy, A., Kozekov, I. D., Wood, R. D., Rizzo, C. J., and Lloyd, R. S. (2010) Novel enzymatic function of DNA polymerase ν in translesion DNA synthesis past major groove DNA-peptide and DNA-DNA cross-links, *Chem. Res. Toxicol.* 689-695.

337. Yamanaka, K., Minko, I. G., Finkel, S. E., Goodman, M. F., and Lloyd, R. S. (2011) Role of high-fidelity *Escherichia coli* DNA polymerase I in replication bypass of a deoxyadenosine DNA-peptide cross-link, *J. Bacteriol. 193*, 3815-3821.

338. Cho, S. H. and Guengerich, F. P. (2013) Replication past the butadiene diepoxide-derived DNA adduct S-[4-(N^6-deoxyadenosinyl)-2,3-dihydroxybutyl]glutathione by DNA polymerases, *Chem. Res. Toxicol. 26*, 1005-1013.

339. Kumari, A., Minko, I. G., Smith, R. L., Lloyd, R. S., and McCullough, A. K. (2010) Modulation of UvrD helicase activity by covalent DNA-protein cross-links, *J. Biol. Chem. 285*, 21313-21322.

340. Zhitkovich, A., Voitkun, V., and Costa, M. (1995) Glutathione and free amino acids form stable complexes with DNA following exposure of intact mammalian cells to chromate, *Carcinogenesis 16*, 907-913.

341. Copeland, K. D., Lueras, A. M., Stemp, E. D., and Barton, J. K. (2002) DNA cross-linking with metallointercalator-peptide conjugates, *Biochemistry 41*, 12785-12797.

342. Kuykendall, J. R. and Bogdanffy, M. S. (1994) Formation and stability of acetaldehyde-induced crosslinks between poly-lysine and poly-deoxyguanosine, *Mutat. Res. 311*, 49-56.

343. Marchand, C., Krajewski, K., Lee, H. F., Antony, S., Johnson, A. A., Amin, R., Roller, P., Kvaratskhelia, M., and Pommier, Y. (2006) Covalent binding of the natural antimicrobial peptide indolicidin to DNA abasic sites, *Nucleic Acids Res. 34*, 5157-5165.

344. Verdine, G. L. and Norman, D. P. (2003) Covalent trapping of protein-DNA complexes, *Annu. Rev. Biochem. 72*, 337-366.

345. Zharkov, D. O., Golan, G., Gilboa, R., Fernandes, A. S., Gerchman, S. E., Kycia, J. H., Rieger, R. A., Grollman, A. P., and Shoham, G. (2002) Structural analysis of an *Escherichia coli* endonuclease VIII covalent reaction intermediate, *EMBO J. 21*, 789-800.

346. Gilboa, R., Zharkov, D. O., Golan, G., Fernandes, A. S., Gerchman, S. E., Matz, E., Kycia, J. H., Grollman, A. P., and Shoham, G. (2002) Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA, *J. Biol Chem. 277*, 19811-19816.

347. Banerjee, A., Yang, W., Karplus, M., and Verdine, G. L. (2005) Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA, *Nature 434*, 612-618.

348. Fromme, J. C. and Verdine, G. L. (2002) Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM, *Nat. Struct. Biol 9*, 544-552.

349. Sanchez, A. M., Minko, I. G., Kurtz, A. J., Kanuri, M., Moriya, M., and Lloyd, R. S. (2003) Comparative evaluation of the bioreactivity and mutagenic spectra of acrolein-derived α-HOPdG and γ-HOPdG regioisomeric deoxyguanosine adducts, *Chem. Res. Toxicol. 16*, 1019-1028.

350. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance, *Science 282*, 1669-1675.

351. Stanojevic, D. and Verdine, G. L. (1995) Deconstruction of GCN4/GCRE into a monomeric peptide-DNA complex, *Nat. Struct. Biol. 2*, 450-457.

352. Duckworth, B. P., Chen, Y., Wollack, J. W., Sham, Y., Mueller, J. D., Taton, T. A., and Distefano, M. D. (2007) A universal method for the preparation of covalent protein-DNA conjugates for use in creating protein nanostructures, *Angew. Chem. Int. Ed Engl. 46*, 8819-8822.

353. Khatwani, S. L., Kang, J. S., Mullen, D. G., Hast, M. A., Beese, L. S., Distefano, M. D., and Taton, T. A. (2012) Covalent protein-oligonucleotide conjugates by copper-free click reaction, *Bioorg. Med. Chem. 20*, 4532-4539.

354. Wickramaratne, S., Mukherjee, S., Villalta, P. W., Scharer, O. D., and Tretyakova, N. Y. (2013) Synthesis of sequence-specific DNA-protein conjugates via a reductive amination strategy, *Bioconjug. Chem. 24*, 1496-1506.

355. Ward, J. B., Jr., Ammenheuser, M. M., Bechtold, W. E., Whorton, E. B., Jr., and Legator, M. S. (1994) hprt mutant lymphocyte frequencies in workers at a 1,3-butadiene production plant, *Environ. Health Perspect. 102 Suppl 9*, 79-85.

356. Melnick, R. L. and Huff, J. E. (1993) 1,3-Butadiene induces cancer in experimental animals at all concentrations from 6.25 to 8000 parts per million, *IARC Sci. Publ.* 309-322.

357. Boysen, G., Pachkowski, B. F., Nakamura, J., and Swenberg, J. A. (2009) The formation and biological significance of N7-guanine adducts, *Mutat. Res. 678*, 76-94.

358. Cochrane, J. E. and Skopek, T. R. (1994) Mutagenicity of butadiene and its epoxide metabolites: II. Mutational spectra of butadiene, 1,2-epoxybutene and diepoxybutane at the *hprt* locus in splenic T cells from exposed B6C3F1 mice, *Carcinogenesis 15*, 719-723.

359. Steen, A. M., Meyer, K. G., and Recio, L. (1997) Characterization of hprt mutations following 1,2-epoxy-3-butene exposure of human TK6 cells, *Mutagenesis 12*, 359-364.

360. Steen, A. M., Meyer, K. G., and Recio, L. (1997) Analysis of hprt mutations occurring in human TK6 lymphoblastoid cells following exposure to 1,2,3,4-diepoxybutane, *Mutagenesis 12*, 61-67.

361. 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*, 445-457.

362. Seneviratne, U., Antsypovich, S., Quirk-Dorr, D., Dissanayake, T., Kotapati, S., and Tretyakova NY (2010) DNA oligomers containing site-specific and stereospecific exocyclic deoxyadenosine adducts of 1,2,3,4-diepoxybutane: synthesis, characterization, and effects on DNA structure, *Chem. Res. Toxicol. 23*, 1556-1567.

363. Zhao, X., Muller, J. G., Halasyam, M., David, S. S., and Burrows, C. J. (2007) *In vitro* ligation of oligodeoxynucleotides containing *C8*-oxidized purine lesions using bacteriophage T4 DNA ligase, *Biochemistry 46*, 3734-3744.

364. Xu, W., Merritt, W. K., Nechev, L. V., Harris, T. M., Harris, C. M., Lloyd, R. S., and Stone, M. P. (2007) Structure of the 1,4-bis(2'-deoxyadenosin-N^6-yl)-2S,3S-butanediol intrastrand DNA cross-link arising from butadiene diepoxide in the human *N-ras* codon 61 sequence, *Chem. Res. Toxicol. 20*, 187-198.

365. Harris, C. M., Zhou, L., Strand, E. A., and Harris, T. M. (1991) New strategy for the synthesis of oligodeoxynucleotides bearing adducts at exocyclic amino sites of purine nucleosides, *J. Am. Chem. Soc. 113*, 4328-4329.

366. Liu, P., Burdzy, A., and Sowers, L. C. (2002) Substrate recognition by a family of uracil-DNA glycosylases: UNG, MUG, and TDG, *Chem. Res. Toxicol. 15*, 1001-1009.

367. Zhao, X., Krishnamurthy, N., Burrows, C. J., and David, S. S. (2010) Mutation versus repair: NEIL1 removal of hydantoin lesions in single-stranded, bulge, bubble, and duplex DNA contexts, *Biochemistry 49*, 1658-1666.

368. Hazra, T. K., Hill, J. W., Izumi, T., and Mitra, S. (2001) Multiple DNA glycosylases for repair of 8-oxoguanine and their potential *in vivo* functions, *Prog. Nucleic Acid Res. Mol. Biol 68*, 193-205.

369. Fersht, A. (1999) *Structure and mechanism in protein science* W. H. Freeman and Company, New York, NY.

370. Wickramaratne, S., Kotapati, S., and Tretyakova N.Y. Synthesis and biological evaluation of site-specific DNA lesions of 1,3-butadiene. 242nd ACS National Meeting, Denver, CO. 2011.

371. Kowal, E. A., Seneviratne, U., Wickramaratne, S., Doherty, K. E., Cao, X., Tretyakova, N., and Stone, M. P. (2014) Structures of exocyclic *R,R*- and *S,S*-N^6,N^6-(2,3-dihydroxybutan-1,4-diyl)-2'-deoxyadenosine adducts induced by 1,2,3,4-diepoxybutane, *Chem. Res. Toxicol. 27*, 805-817.

372. Kotapati, S., Wickramaratne, S., Kowal, E. A., Stone, M. P., Guengerich, F. P., and Tretyakova N.Y. Translesion synthesis and structural studies of 1,3-butadiene-induced deoxyadenosine adducts. 246th ACS National Meeting, Indianapolis, IN. 2013.

373. Liuzzi, M. and Talpaert-Borle, M. (1985) A new approach to the study of the base-excision repair pathway using methoxyamine, *J. Biol Chem. 260*, 5252-5258.

374. Talpaert-Borle, M. and Liuzzi, M. (1983) Reaction of apurinic/apyrimidinic sites with [^{14}C]methoxyamine. A method for the quantitative assay of AP sites in DNA, *Biochim. Biophys. Acta 740*, 410-416.

375. Singer, B. and Grunberger, D. (1983) *Molecular biology of mutagens and carcinogens* Plenum Press, New York and London.

376. Fedtke, N., Boucheron, J. A., Walker V.E., and Swenberg, J. A. (1990) Vinyl chloride-induced DNA adducts. II: Formation and persistence of 7-(2'-oxoethyl)guanine and N^2,3-ethenoguanine in rat tissue DNA, *Carcinogenesis 11*, 1287-1292.

377. Angelov, T., Guainazzi, A., and Scharer, O. D. (2009) Generation of DNA interstrand cross-links by post-synthetic reductive amination, *Org. Lett. 11*, 661-664.

378. Gates, K. S., Nooner, T., and Dutta, S. (2004) Biologically relevant chemical reactions of *N7*-alkylguanine residues in DNA, *Chem. Res. Toxicol. 17*, 839-856.

379. Westbye, M. P., Feyzi, E., Aas, P. A., Vagbo, C. B., Talstad, V. A., Kavli, B., Hagen, L., Sundheim, O., Akbari, M., Liabakk, N., Slupphaug, G., Otterlei, M., and Krokan, H. E. (2008) Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA, *J. Biol. Chem. 283*, 25046-25056.

380. Yang, C. G., Yi, C., Duguid, E. M., Sullivan, C. T., Jian, X., Rice, P. A., and He, C. (2008) Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA, *452*, 961-965.

381. Fischle, W., Wang, Y., and Allis, C. D. (2003) Histone and chromatin cross-talk, *Curr. Opin. Cell Biol. 15*, 172-183.

382. Bandyopadhyay, U., Das, D., and Banerjee, R. K. (1999) Reactive oxygen species: oxidative damage and pathogenesis, *Curr. Sci. 77*, 658-666.

383. Lovell, M. A. and Markesbery, W. R. (2007) Oxidative DNA damage in mild cognitive impairment and late-stage Alzheimer's disease, *Nucleic Acids Res. 35*, 7497-7504.

384. Choi, J. Y. and Guengerich, F. P. (2005) Adduct size limits efficient and error-free bypass across bulky N^2-guanine DNA lesions by human DNA polymerase η, *J. Mol. Biol. 352*, 72-90.

385. Zhao, L., Christov, P. P., Kozekov, I. D., Pence, M. G., Pallan, P. S., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2012) Replication of N^2,3-ethenoguanine by DNA polymerases, *Angew. Chem. Int. Ed Engl. 51*, 5466-5469.

386. Pence, M. G., Choi, J. Y., Egli, M., and Guengerich, F. P. (2010) Structural basis for proficient incorporation of dTTP opposite O^6-methylguanine by human DNA polymerase ι, *J. Biol. Chem. 285*, 40666-40672.

387. Chung, J. A., Wollack, J. W., Hovlid, M. L., Okesli, A., Chen, Y., Mueller, J. D., Distefano, M. D., and Taton, T. A. (2009) Purification of prenylated proteins by affinity chromatography on cyclodextrin-modified agarose, *Anal. Biochem. 386*, 1-8.

388. Yeo, J. E., Wickramaratne, S., Khatwani, S., Wang, Y. C., Vervacke, J. S., Distefano, M. D., and Tretyakova, N. Y. (2014) Synthesis of site-specific DNA-protein conjugates and their effects on DNA replication, *ACS Chem. Biol. 9*, 1860-1868.

389. Ho, T. V., Guainazzi, A., Derkunt, S. B., Enoiu, M., and Scharer, O. D. (2011) Structure-dependent bypass of DNA interstrand crosslinks by translesion synthesis polymerases, *Nucleic Acids Res. 39*, 7455-7464.

390. Nilsson, J. A. and Cleveland, J. L. (2003) Myc pathways provoking cell suicide and cancer, *Oncogene 22*, 9007-9021.

391. Hoffman, B. and Liebermann, D. A. (2008) Apoptotic signaling by c-MYC, *Oncogene 27*, 6462-6472.

392. Schellenberger, M. T., Grova, N., Farinelle, S., Willieme, S., Revets, D., and Muller, C. P. (2012) Immunogenicity of a promiscuous T cell epitope peptide based conjugate vaccine against benzo[*a*]pyrene: redirecting antibodies to the hapten, *PLoS. One. 7*, e38329.

393. Dronkert, M. L. and Kanaar, R. (2001) Repair of DNA interstrand cross-links, *Mutat. Res. 486*, 217-247.

394. Levine, R. L., Miller, H., Grollman, A., Ohashi, E., Ohmori, H., Masutani, C., Hanaoka, F., and Moriya, M. (2001) Translesion DNA synthesis catalyzed by human pol η and pol κ across 1,N^6-ethenodeoxyadenosine, *J. Biol. Chem. 276*, 18717-18721.

395. Zhao, L., Pence, M. G., Christov, P. P., Wawrzak, Z., Choi, J. Y., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2012) Basis of miscoding of the DNA adduct N^2,3-ethenoguanine by human Y-family DNA polymerases, *J. Biol Chem. 287*, 35516-35526.

396. Duckworth, B. P., Xu, J., Taton, T. A., Guo, A., and Distefano, M. D. (2006) Site-specific, covalent attachment of proteins to a solid surface, *Bioconjug. Chem. 17*, 967-974.

397. Rashidian, M., Dozier, J. K., and Distefano, M. D. (2013) Enzymatic labeling of proteins: techniques and approaches, *Bioconjug. Chem. 24*, 1277-1294.

398. Pence, M. G., Blans, P., Zink, C. N., Hollis, T., Fishbein, J. C., and Perrino, F. W. (2009) Lesion bypass of N^2-ethylguanine by human DNA polymerase ι, *J. Biol. Chem. 284*, 1732-1740.

399. Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S., and Prakash, L. (2000) Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions, *Nature 406*, 1015-1019.

400. Tissier, A., McDonald, J. P., Frank, E. G., and Woodgate, R. (2000) Pol ι, a remarkably error-prone human DNA polymerase, *Genes Dev. 14*, 1642-1650.

401. Zhang, Y., Yuan, F., Wu, X., and Wang, Z. (2000) Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase ι, *Mol. Cell Biol 20*, 7099-7108.

402. Ziv, O., Geacintov, N., Nakajima, S., Yasui, A., and Livneh, Z. (2009) DNA polymerase ζ cooperates with polymerases κ and ι in translesion DNA synthesis across pyrimidine photodimers in cells from XPV patients, *Proc. Natl. Acad. Sci. U. S. A. 106*, 11552-11557.

403. Efthymiou, T., Gong, W., and Desaulniers, J. P. (2012) Chemical architecture and applications of nucleic acid derivatives containing 1,2,3-triazole functionalities synthesized via click chemistry, *Molecules. 17*, 12665-12703.

404. El Sagheer, A. H. and Brown, T. (2010) Click chemistry with DNA, *Chem. Soc. Rev. 39*, 1388-1405.

405. Efrati, E., Tocco, G., Eritja, R., Wilson, S. H., and Goodman, M. F. (1997) Abasic translesion synthesis by DNA polymerase β violates the "A-rule". Novel types of nucleotide incorporation by human DNA polymerase β at an abasic lesion in different sequence contexts, *J. Biol Chem. 272*, 2559-2569.

406. Becherel, O. J. and Fuchs, R. P. (2001) Mechanism of DNA polymerase II-mediated frameshift mutagenesis, *Proc. Natl. Acad. Sci. U. S. A. 98*, 8566-8571.

407. Zang, H., Goodenough, A. K., Choi, J. Y., Irimia, A., Loukachevitch, L. V., Kozekov, I. D., Angel, K. C., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2005) DNA adduct bypass polymerization by *Sulfolobus solfataricus* DNA polymerase Dpo4: analysis and crystal structures of multiple base pair substitution and frameshift products with the adduct 1,N^2-ethenoguanine, *J. Biol. Chem. 280*, 29750-29764.

408. Smith, S. S., Kaplan, B. E., Sowers, L. C., and Newman, E. M. (1992) Mechanism of human methyl-directed DNA methyltransferase and the fidelity of cytosine methylation, *Proc. Natl. Acad. Sci. U. S. A. 89*, 4744-4748.

409. Goll, M. G. and Bestor, T. H. (2005) Eukaryotic cytosine methyltransferases, *Annu. Rev. Biochem. 74*, 481-514.

410. Dizdaroglu, M. and Jaruga, P. (2012) Mechanisms of free radical-induced damage to DNA, *Free Radic. Res. 46*, 382-419.

411. Morimoto, S., Hatta, H., Fujita, S., Matsuyama, T., Ueno, T., and Nishimoto, S. (1998) Hydroxyl radical-induced cross-lonking of thymine and lysine: identification of the primary structure and mechanism, *Bioorg. Med. Chem. Lett. 8*, 865-870.

412. Charlton, T. S., Ingelse, B. A., Black, D. S., Craig, D. C., Mason, K. E., and Duncan, M. W. (1999) A covalent thymine-tyrosine adduct involved in DNA-protein cross-links: synthesis, characterization and quantification, *Free Radic. Biol Med. 27*, 254-261.

413. Christov, P. P., Angel, K. C., Guengerich, F. P., and Rizzo, C. J. (2009) Replication past the N^5-methyl-formamidopyrimidine lesion of deoxyguanosine by DNA polymerases and an improved procedure for sequence analysis of *in vitro* bypass products by mass spectrometry, *Chem. Res. Toxicol. 22*, 1086-1095.

414. Stover, J. S., Chowdhury, G., Zang, H., Guengerich, F. P., and Rizzo, C. J. (2006) Translesion synthesis past the C8- and N^2-deoxyguanosine adducts of the dietary mutagen 2-Amino-3-methylimidazo[4,5-*f*]quinoline in the *Nar*I recognition sequence by prokaryotic DNA polymerases, *Chem. Res. Toxicol. 19*, 1506-1517.

415. Matsuda, T., Bebenek, K., Masutani, C., Rogozin, I. B., Hanaoka, F., and Kunkel, T. A. (2001) Error rate and specificity of human and murine DNA polymerase η, *J. Mol. Biol 312*, 335-346.

416. Zhang, Y., Yuan, F., Xin, H., Wu, X., Rajpal, D. K., Yang, D., and Wang, Z. (2000) Human DNA polymerase κ synthesizes DNA with extraordinarily low fidelity, *Nucleic Acids Res. 28*, 4147-4156.

417. Johnson, R. E., Washington, M. T., Prakash, S., and Prakash, L. (2000) Fidelity of human DNA polymerase η, *J. Biol Chem. 275*, 7447-7450.

418. Zhang, Y., Yuan, F., Wu, X., Rechkoblit, O., Taylor, J. S., Geacintov, N. E., and Wang, Z. (2000) Error-prone lesion bypass by human DNA polymerase η, *Nucleic Acids Res. 28*, 4717-4724.

419. Zhang, Y., Yuan, F., Wu, X., Wang, M., Rechkoblit, O., Taylor, J. S., Geacintov, N. E., and Wang, Z. (2000) Error-free and error-prone lesion bypass by human DNA polymerase κ *in vitro*, *Nucleic Acids Res. 28*, 4138-4146.

420. Ohashi, E., Ogi, T., Kusumoto, R., Iwai, S., Masutani, C., Hanaoka, F., and Ohmori, H. (2000) Error-prone bypass of certain DNA lesions by the human DNA polymerase κ, *Genes Dev. 14*, 1589-1594.

421. Mukherjee, P., Lahiri, I., and Pata, J. D. (2013) Human polymerase kappa uses a template-slippage deletion mechanism, but can realign the slipped strands to favour base substitution mutations over deletions, *Nucleic Acids Res. 41*, 5024-5035.

422. Surova, O. and Zhivotovsky, B. (2013) Various modes of cell death induced by DNA damage, *Oncogene 32*, 3789-3797.

423. Pravetoni, M., Vervacke, J. S., Distefano, M. D., Tucker, A. M., Laudenbach, M., and Pentel, P. R. (2014) Effect of currently approved carriers and adjuvants on the pre-clinical efficacy of a conjugate vaccine against oxycodone in mice and rats, *PLoS. One. 9*, e96547.

424. Baggio, R., Burgstaller, P., Hale, S. P., Putney, A. R., Lane, M., Lipovsek, D., Wright, M. C., Roberts, R. W., Liu, R., Szostak, J. W., and Wagner, R. W. (2002) Identification of epitope-like consensus motifs using mRNA display, *J. Mol. Recognit. 15*, 126-134.

425. Rieger, R. A., Zaika, E. I., Xie, W., Johnson, F., Grollman, A. P., Iden, C. R., and Zharkov, D. O. (2006) Proteomic approach to identification of proteins reactive for abasic sites in DNA, *Mol. Cell Proteomics. 5*, 858-867.

426. Kliszczak, A. E., Rainey, M. D., Harhen, B., Boisvert, F. M., and Santocanale, C. (2011) DNA mediated chromatin pull-down for the study of chromatin replication, *Sci. Rep. 1*, 95.

427. McCullough, A. K., Dodson, M. L., and Lloyd, R. S. (1999) Initiation of base excision repair: glycosylase mechanisms and structures, *Annu. Rev. Biochem. 68*, 255-285.

428. Brooks, S. C., Adhikary, S., Rubinson, E. H., and Eichman, B. F. (2013) Recent advances in the structural mechanisms of DNA glycosylases, *Biochim. Biophys. Acta 1834*, 247-271.

429. Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., Reader, V., Sweetman, G., Bauer, A., Bouwmeester, T., Hopf, C., Kruse, U., Neubauer, G., Ramsden, N., Rick, J., Kuster, B., and Drewes, G. (2007) Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors, *Nat. Biotechnol. 25*, 1035-1044.

430. Leung, H. C. E. (2012) *Integrative proteomics* InTech, Rijeka, Croatia.

431. White, F. M. (2007) On the iTRAQ of kinase inhibitors, *Nat. Biotechnol. 25*, 994-996.

432. Meany, D. L., Xie, H., Thompson, L. V., Arriaga, E. A., and Griffin, T. J. (2007) Identification of carbonylated proteins from enriched rat skeletal muscle mitochondria using affinity chromatography-stable isotope labeling and tandem mass spectrometry, *Proteomics. 7*, 1150-1163.

433. McKibbin, P. L., Fleming, A. M., Towheed, M. A., Van Houten, B., Burrows, C. J., and David, S. S. (2013) Repair of hydantoin lesions and their amine adducts in DNA by base and nucleotide excision repair, *J. Am. Chem. Soc. 135*, 13851-13861.

434. Mu, H., Kropachev, K., Wang, L., Zhang, L., Kolbanovskiy, A., Kolbanovskiy, M., Geacintov, N. E., and Broyde, S. (2012) Nucleotide excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies, *Nucleic Acids Res. 40*, 9675-9690.

435. Yeo, J. E., Khoo, A., Fagbemi, A. F., and Scharer, O. D. (2012) The efficiencies of damage recognition and excision correlate with duplex destabilization induced by acetylaminofluorene adducts in human nucleotide excision repair, *Chem. Res. Toxicol. 25*, 2462-2468.

436. Ang, W. H., Myint, M., and Lippard, S. J. (2010) Transcription inhibition by platinum-DNA cross-links in live mammalian cells, *J. Am. Chem. Soc. 132*, 7429-7435.

437. Delaney, J. C. and Essigmann, J. M. (1999) Context-dependent mutagenesis by DNA lesions, *Chem. Biol. 6*, 743-753.

438. Delaney, J. C. and Essigmann, J. M. (2004) Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli, *Proc. Natl. Acad. Sci. U. S. A. 101*, 14051-14056.

439. Delaney, J. C. and Essigmann, J. M. (2006) Assays for determining lesion bypass efficiency and mutagenicity of site-specific DNA lesions *in vivo*, *Methods Enzymol. 408*, 1-15.

440. Avkin, S., Adar, S., Blander, G., and Livneh, Z. (2002) Quantitative measurement of translesion replication in human cells: evidence for bypass of abasic sites by a replicative DNA polymerase, *Proc. Natl. Acad. Sci. U. S. A. 99*, 3764-3769.

441. Avkin, S., Goldsmith, M., Velasco-Miguel, S., Geacintov, N., Friedberg, E. C., and Livneh, Z. (2004) Quantitative analysis of translesion DNA synthesis across a benzo[*a*]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase κ, *J. Biol Chem. 279*, 53298-53305.

442. Colis, L. C., Raychaudhury, P., and Basu, A. K. (2008) Mutational specificity of γ-radiation-induced guanine-thymine and thymine-guanine intrastrand cross-links in mammalian cells and translesion synthesis past the guanine-thymine lesion by human DNA polymerase η, *Biochemistry 47*, 8070-8079.

443. Jasti, V. P., Das, R. S., Hilton, B. A., Weerasooriya, S., Zou, Y., and Basu, A. K. (2011) (5'S)-8,5'-cyclo-2'-deoxyguanosine is a strong block to replication, a potent pol V-dependent mutagenic lesion, and is inefficiently repaired in *Escherichia coli*, *Biochemistry 50*, 3862-3865.

444. Malia, S. A., Vyas, R. R., and Basu, A. K. (1996) Site-specific frame-shift mutagenesis by the 1-nitropyrene-DNA adduct *N*-(deoxyguanosin-8-y1)-1-aminopyrene located in the $(CG)_3$ sequence: effects of SOS, proofreading, and mismatch repair, *Biochemistry 35*, 4568-4577.

445. Kalam, M. A. and Basu, A. K. (2005) Mutagenesis of 8-oxoguanine adjacent to an abasic site in simian kidney cells: tandem mutations and enhancement of G→T transversions, *Chem. Res. Toxicol. 18*, 1187-1192.

446. Goggin, M., Swenberg, J. A., Walker, V. E., and Tretyakova, N. (2009) Molecular dosimetry of 1,2,3,4-diepoxybutane-induced DNA-DNA cross-links in B6C3F1 mice and F344 rats exposed to 1,3-butadiene by inhalation, *Cancer Res. 69*, 2479-2486.

447. Kropachev, K., Kolbanovskiy, M., Liu, Z., Cai, Y., Zhang, L., Schwaid, A. G., Kolbanovskiy, A., Ding, S., Amin, S., Broyde, S., and Geacintov, N. E. (2013) Adenine-DNA adducts derived from the highly tumorigenic dibenzo[*a,l*]pyrene are resistant to nucleotide excision repair while guanine adducts are not, *Chem. Res. Toxicol. 26*, 783-793.

448. Rechkoblit, O., Zhang, Y., Guo, D., Wang, Z., Amin, S., Krzeminsky, J., Louneva, N., and Geacintov, N. E. (2002) trans-Lesion synthesis past bulky benzo[*a*]pyrene diol epoxide N^2-dG and N^6-dA lesions catalyzed by DNA bypass polymerases, *J. Biol. Chem. 277*, 30488-30494.

449. Lavrukhin, O. V. and Lloyd, R. S. (1998) Mutagenic replication in a human cell extract of DNAs containing site-specific and stereospecific benzo(*a*)pyrene-7,8-diol-9,10-epoxide DNA adducts placed on the leading and lagging strands, *Cancer Res. 58*, 887-891.

450. Suzuki, M., Kojima, M., Okuda, H., Watabe, T., and Tada, M. (1991) DNA polymerase action blocked by adenine adducts induced by 5-hydroxy-methylchrysene sulfate, *Biochem. Int. 25*, 19-27.

451. Meng, X., Zhou, Y., Zhang, S., Lee, E. Y., Frick, D. N., and Lee, M. Y. (2009) DNA damage alters DNA polymerase δ to a form that exhibits increased discrimination against modified template bases and mismatched primers, *Nucleic Acids Res. 37*, 647-657.

452. Miller, H. and Grollman, A. P. (2003) DNA repair investigations using siRNA, *DNA Repair (Amst) 2*, 759-763.

453. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells, *Nature 411*, 494-498.

454. Semizarov, D., Kroeger, P., and Fesik, S. (2004) siRNA-mediated gene silencing: a global genome view, *Nucleic Acids Res. 32*, 3836-3845.

455. Espeseth, A. S., Fishel, R., Hazuda, D., Huang, Q., Xu, M., Yoder, K., and Zhou, H. (2011) siRNA screening of a targeted library of DNA repair factors in HIV infection reveals a role for base excision repair in HIV integration, *PLoS. One. 6*, e17612.

456. Johnson, J. M. and Latimer, J. J. (2005) Analysis of DNA repair using transfection-based host cell reactivation, *Methods Mol. Biol 291*, 321-335.

457. Sherf, B. A., Navarro, S. L., Hannah, R. R., and Wood, K. V. (1996) Dual-luciferase reporter assay: an advanced co-reporter technology integrating firefly and *Renilla* luciferase assays, *Promega Notes Magazine 57*, 2-8.

458. Paddison, P. J., Caudy, A. A., and Hannon, G. J. (2002) Stable suppression of gene expression by RNAi in mammalian cells, *Proc. Natl. Acad. Sci. U. S. A. 99*, 1443-1448.

459. Matijasevic, Z., Precopio, M. L., Snyder, J. E., and Ludlum, D. B. (2001) Repair of sulfur mustard-induced DNA damage in mammalian cells measured by a host cell reactivation assay, *Carcinogenesis 22*, 661-664.

460. Pitsikas, P., Lee, D., and Rainbow, A. J. (2007) Reduced host cell reactivation of oxidative DNA damage in human cells deficient in the mismatch repair gene hMSH2, *Mutagenesis 22*, 235-243.

461. Wu, W. J., Su, M. I., Wu, J. L., Kumar, S., Lim, L. H., Wang, C. W., Nelissen, F. H., Chen, M. C., Doreleijers, J. F., Wijmenga, S. S., and Tsai, M. D. (2014) How a low-fidelity DNA polymerase chooses non-Watson-Crick from Watson-Crick incorporation, *J. Am. Chem. Soc. 136*, 4927-4937.