# MASS SPECTROMETRY BASED ANALYSIS OF ELECTROPHILE INDUCED ADDUCTS OF BIOMOLECULES

A THESIS
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

**Amanda Eileen Degner**

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

**Dr. Natalia Y Tretyakova, Advisor**

**August 2019**

## Acknowledgments

The following dissertation could not have been accomplished without the help, support, and guidance of several individuals. First and foremost, I would like to thank my advisor, Dr. Natalia Tretyakova, as she has been a constant source of guidance and support for me. I would also like to thank Dr. Barry Finzel, Dr. Calin Campbell, and Dr. Stephen Hecht for agreeing to serve on my thesis committee and for their valuable feedback and suggestions. I am exceedingly grateful for our numerous collaborators who provided generous help and support for this research – Dr. Margareta Törnqvist (Stockholm University, Stockholm, SE), Dr. Vernon Walker (University of Vermont, Burlington, VT), Dr. Lisa Peterson (University of Minnesota, Minneapolis, MN), and Dr. Colin Campbell (University of Minnesota, Minneapolis, MN). I would also like to thank the NIH, the Sw. Research Council, Stockholm University, and the Department of Medicinal Chemistry at the University of Minnesota for their financial support.

I would like to thank Dr. Peter Villalta for his help with MS troubleshooting and for always being available to help when something went wrong with our instrument. I am extremely grateful for the help of Dr. Christopher Seiler, who trained me in HPLC and MS instrumentation and taught me how to be a good scientist. I would also like to thank all of the current and past members of the Tretyakova lab for their feedback on my research and their constant help and support. In particular, I would like to express my sincere gratitude to Dr. Dewakar Sangaraju, Dr. Luke Erber, Dr. Suresh Pujari, Emily Boldry, Caitlin Jokipii-Krueger, Emina Dzafic, and Dominic Najjar for making significant contributions

to my thesis. I would also like to thank Dr. Rashi Arora, Karin Vevang, Dr. Lisa Chesner, Dr. Henrik Carlsson, Dr. Isabella Karlsson, and Dr. Johan Eriksson, whose collaborations made this research possible. I am also appreciative of the help of Bob Carlson in preparing the graphics for this thesis and other publications.

Last but not least, I would like to thank my family and friends for all of their love and support. I would not have pursued or succeeded in my graduate studies without the support of my parents, Chris and Mary Degner, for which I am eternally thankful. I would also like to thank my sisters, Emily and Olivia, my brother Jacob, and my friend and roommate, Katie Schlasner.

## Dedication

This thesis is dedicated to my late friend and roommate through all four years of my undergraduate studies, Jenny Johnson.

## ABSTRACT

Humans are exposed to a wide range of electrophilic agents, which can form covalent adducts at nucleophilic sites in biomolecules including DNA and proteins. If not repaired, DNA adducts can inhibit biological processes such as DNA replication and transcription, which can lead to mutagenesis, carcinogenesis, and toxicity. DNA repair pathways exist to remove these lesions, but the role of different pathways in the repair of different DNA adducts is not fully understood. Additionally, electrophiles can be detoxified through further metabolism such as glutathione conjugation before they are able to form adducts. Interindividual differences in the balance between detoxification, adduct formation, and repair ultimately determines the risk a person has of developing cancer after carcinogen exposure.

In Chapter 2, we investigated the role of nucleotide excision repair (NER) and Fanconi Anemia (FA) repair pathways in repair of a 1,3-butadiene (BD) induced DNA-DNA interstrand crosslink (ICL), *bis*-N7G-BD. NanoLC-nanoESI$^+$-MS/MS methodology was employed to quantify the amount of crosslinks present after human cell lines deficient in components of NER or FA and their isogenic controls were treated with an electrophilic metabolite of BD, 1,2,3,4-diepoxybutane (DEB), and allowed to repair for up to 72 h. Despite observing increased sensitivity to DEB treatment in repair deficient cell lines in cytotoxicity assays, no change in repair kinetics was observed between repair deficient and control cells, suggesting that in humans, there is a redundancy in DNA repair pathways. The same nanoLC-nanoESI$^+$-MS/MS methodology was employed to measure *bis*-N7G-

BD formation in HL60 cells, where crosslink levels correlated with cytotoxicity and micronuclei formation.

In Chapter 3, we investigated the roll of a glutathione S-transferase theta 1 (GSTT1) in detoxification of another epoxide metabolite of BD, 3,4-epoxy-1-butene (EB). In a previous genome wide association study, differences in GSTT1 copy number explained a fraction of ethnic differences in BD metabolism, which could contribute to ethnic differences in smoking-induced lung cancer risk. A HPLC-$ESI^+$-MS/MS method was developed for the quantitation of EB-GSH conjugates in cells that did or did not express GSTT1 after treatment of EB. No difference in EB-GSH conjugates or EB-GII DNA adduct levels were observed between $GSTT1^{-/-}$ and $GSTT1^{+/+}$ cell lines. However, the expression of GSTT1 was had a protective effect against EB-induced apoptosis.

In Chapter 4, an unknown electrophile-induced adduct to hemoglobin (Hb) was identified. Previous adductomic screens of human blood samples using the FI*R*E procedure to screen for adducts to N-terminal valine in Hb found an unknown adduct corresponding to an added mass of 106.042 Da. This adduct was identified as a 4-OHBn adduct to N-terminal valine, which was confirmed by HPLC-$ESI^+$-HRMS analysis of authentic synthesized standard. Accurate mass, retention time, and MS/MS fragmentation patterns between the standard and unknown adduct matched. Levels of the 4OHBn-Val adduct were quantified to be 380 ± 160 pmol/g Hb in 12 human blood samples. Both 4-quinone methide and 4-hydroxybenzaldehyde were found to be capable of forming this adduct and thus were identified as possible sources.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF SCHEMES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| μmol | micro mole ($10^{-6}$) |
| 2-OHBA | 2-hydroxybenzaldehyde |
| 2-OHBn | 2-hydroxybenzyl |
| 2-QM | *ortho*-quinone methide |
| 3-OHBA | 3-hydroxybenzaldehyde |
| 4-OHBA | 4-hydroxybenzaldehyde |
| 4-OHBn | 4-hydroxybenzyl |
| 4-QM | *para*-quinone methide |
| A | adenine |
| ACN | acetonitrile |
| AIBN | azobisisobutyronitrile |
| AP | abasic site |
| BA | benzaldehyde |
| bcr | breakpoint cluster region |
| BD | 1,3-butadiene |
| BER | base excision repair |
| *Bis*-N7G-BD | 1,4-*bis*-(guan-7-yl)-2,3-butanediol |
| BNC | binucleated cells |
| C | cytosine |
| cDNA | complementary DNA |

| | |
|---|---|
| CID | collision-induced dissociation |
| CML | chronic myeloid leukemia |
| CYP | cytochrome P450 monooxygenase |
| Cys | cysteine |
| DEB | 1,2,3,4-diepoxybutane |
| DEB-GSH | S-(2-hydroxy-3,4-epoxybutyl)glutathione |
| DIA | data independent acquisition |
| DMF | dimethylformamide |
| DMSO | dimethyl sulfoxide |
| DNA | deoxyribonucleic acid |
| DNA-PK | DNA-dependent protein kinase |
| DSB | double strand break |
| EB | 3,4-epoxy-1-butene |
| EBD | 3,4-epoxy-1,2-butanediol |
| EB-diol | 1,butene-3,4-diol |
| EB-GI | N7-(2-hydroxy-3-buten-1-yl)guanine |
| EB-GII | N-7-(1-hydroxy-3-buten-2-yl)guanine |
| EB-GSH | (1-hydroxy-3-buten-2-yl)glutathione |
| EPHX1 | epoxide hydrolase 1 |
| ESI | electrospray ionization |
| EtOH | ethanol |

| | |
|---|---|
| FA | Fanconi Anemia |
| FANC | Fanconi Anemia complementation group |
| Fapy | formamidopyridine |
| FITC | fluorescein isothiocyanate |
| fmol | femto mole ($10^{-15}$) |
| FTH | fluorescein thiohydantoin |
| G | guanine |
| GG-NER | global genomic nucleotide excision repair |
| GLU | glutamate |
| GSH | glutathione |
| GST | glutathione S-transferase |
| GSTM1 | glutathione S-transferase mu 1 |
| GSTT1 | glutathione S-transferase theta 1 |
| GSTT2 | glutathione S-transferase theta 2 |
| Gua | guanine |
| Hb | hemoglobin |
| HL60 | human promyelocytic leukemia cell line |
| HPLC | high performance liquid chromatography |
| HPLC-ESI$^+$-MS/MS | high-pressure liquid chromatography tandem mass spectrometry |
| HRMS | high resolution mass spectrometry |

| | |
|---|---|
| HRR | homologous recombination repair |
| HSA | human serum albumin |
| ICL | interstrand crosslink |
| IPA | isopropanol |
| LC | liquid chromatography |
| *m/z* | mass/charge ratio |
| MA | mercapturic acid |
| MAX | mixed-mode, strong anion-exchange |
| MCX | mixed-mode, strong cation-exchange |
| MeOH | methanol |
| MHBMA | 1-(*N*-acetyl-L-cystein-S-yl)-1-hydroxybut-3-ene and 2-(*N*-acetyl-L-cystein-S-yl)-1-hydroxybut-3-ene |
| mmol | milli mole ($10^{-3}$) |
| MN | micronuclei |
| MS | mass spectrometry |
| MS/MS | tandem mass spectrometry |
| N7-THBG | N7-(2,3,4-trihydroxybut-1-yl)guanine |
| nanoLC-nanoESI-MS/MS | nanoflow liquid chromatography-nanoelectrospray ionization tandem mass spectrometry |
| NER | nucleotide excision repair |
| $NH_4OAc$ | ammonium acetate |

| | |
|---|---|
| $NH_4OH$ | ammonium hydroxide |
| NHEJ | nonhomologous end-joining |
| nmol | nano mole ($10^{-9}$) |
| PBS | phosphate buffered saline |
| PD20 | human dermal fibroblast deficient in *FANCD2* |
| PD20corr | PD20 isogenic cell line with functional *FANCD2* |
| pmol | pico mole ($10^{-12}$) |
| ppm | parts per million |
| PRM | parallel reaction monitoring |
| QM | quinone methide |
| QMP | quinone methide precursor |
| RBC | red blood cell |
| RF | radio frequency voltage |
| RNA | ribonucleic acid |
| RP | reverse phase |
| RT PCR | reverse transcription polymerase chain reaction |
| RT | retention time |
| SEM | standard error of the mean |
| siRNA | short interfering ribonucleic acid |
| SPE | solid-phase extraction |
| SRM | selected reaction monitoring, aka MRM |

| | |
|---|---|
| ssDNA | single-stranded DNA |
| T | thymine |
| TBS-Cl | *tert*-butyldimethylsilyl chloride |
| TC-NER | transcription-coupled nucleotide excision repair |
| TFIIH | transcription factor II H |
| TLC | thin-layer chromatography |
| Val | valine |
| ValpNA | l-valine *p*-nitroanilide |
| XP | Xeroderma Pigmentosum |
| XPA | human dermal fibroblast deficient in *XPA* |
| XPAcorr | XPA isogenic cell line with functional *XPA* |

# I. LITERATURE REVIEW

## 1.1 Electrophile-DNA adduct formation and their biological outcomes

### 1.1.1 The role of DNA adducts in chemical carcinogenesis

Humans are exposed to a wide range of electrophilic agents from exogenous sources such as air or food and endogenous sources such as metabolism byproducts.[1-4] These electrophiles can form covalent bonds with nucleophilic sites in biomacromolecules, including proteins, DNA, and RNA. Especially in the case of DNA, the formation of covalent adducts is associated with an increased risk of adverse health effects such as cancer, aging, and chronic diseases.[5-7]

Among the nucleophilic sites in DNA, the N7-position of guanine is the most reactive towards electrophiles.[6, 8] Formation of N7-guanine adducts by alkylating agents forms positively charged quaternary nitrogens which destabilize the glycosidic bond and leads to spontaneous depurinaion.[8] Other nucleophilic sites, such as N7-adenine, N3-adenine, N1-adenine, $N^6$-adenine, and $N^2$-guanine, also form nucleobase adducts. However, the peripheral location of the N7-position of purines and the fact that it is not involved in Watson-Crick base-pairing makes it the most easily accessible site for adduct formation.

Some alkylating agents such as nitrogen mustards and aldehydes can directly react with DNA.[9] These compounds already contain electrophilic functional groups that can form covalent adducts. Many other molecules such as 1,3-butadiene (BD) must first undergo some type of metabolic activation in order to generate electrophilic compounds

that can form DNA adducts.[10] Cytochrome P450 monooxygenases are the most important enzymes involved in bioactivation of DNA-reactive molecules, as they are capable of metabolically oxidizing and reducing a wide range of substrates.

Although the majority of DNA adducts are recognized and removed by DNA repair pathways (discussed further in 1.1.2), those left unrepaired can interfere with DNA replication.[11] Some adducts, such as crosslinks, lead to toxic double strand breaks if unrepaired, while others can be misread by polymerases – leading to point mutations. Not all mutations are carcinogenic – some may cause no biological consequences at all while others may be non-viable, leading to apoptosis. However, if a mutation is both viable and occurs within tumor suppressor genes or protooncogenes, it can lead to the initiation of cancer through uncontrolled cell growth.[12]

Carcinogen-DNA adducts represent excellent biomarkers of carcinogen exposure as they are directly implicated in the development of cancer (Scheme 1.1).[13] DNA adducts are commonly used as biomarkers of carcinogen exposure. For example, aflatoxin B1, a food contaminant and liver carcinogen, can induce the formation of 2,3-dihydro-2-(N7-guanyl)-3-hydrocyaflatoxin B1 lesions (AFB1-N7-dG).[14] Measurement of urinary AFB1-N7-dG has been used to as a biomarker of human exposure to aflatoxin.[15] Tobacco smoke contains a number of carcinogens that can form DNA adducts after metabolic activation, such as 1,3-butadiene (BD), acrolein, benzo[a]pyrene, and tobacco specific nitrosamine 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK). These can form a range of DNA adducts including N7-(2,3,4-trihydroxybut-1-yl)guanine (N7-THBG),[16] γ-

hydroxypropano-2-deoxyguanosine (γ-OH-Acr-dG),[17] (+)trans-$N^2$-benzo[a]pyrene-7,8-diol-9,10-epoxide-2-deoxyguanosine ($N^2$-BPDE-dG),[18-19] and 7-[4-(3-pyridyl)-4-oxobut-1-yl]-2-deoxyguanosine (N7-POB-dG).[20] Acetaldehyde, a metabolite of ethanol, can form $N^2$-ethyl-dG adducts (Figure 1.1).[21]

In the case of cancer treatments, toxic DNA adducts are induced intentionally in an attempt to kill tumor cells. The anticancer drugs cyclophosphamide and cisplatin are *bis*-electrophiles that form interstrand and intrastrand DNA-DNA crosslinks.[22-24] Cyclophosphamide, a phosphoramide mustard, forms interstrand N,N-*bis*-[2-(N7-guaninyl) ethyl] amine crosslinks (G-NOR-G),[25] while cisplatin forms the intrastand crosslink cis-$Pt(NH_3)_2$d(pGpG) (Pt-(GG))[26] (Figure 1.1). Quantitative measurements of these adducts in patients can be useful in biomonitoring the pharmacological activity of these drugs.[25, 27]

**Scheme 1.1** Biological outcomes of carcinogen exposure.

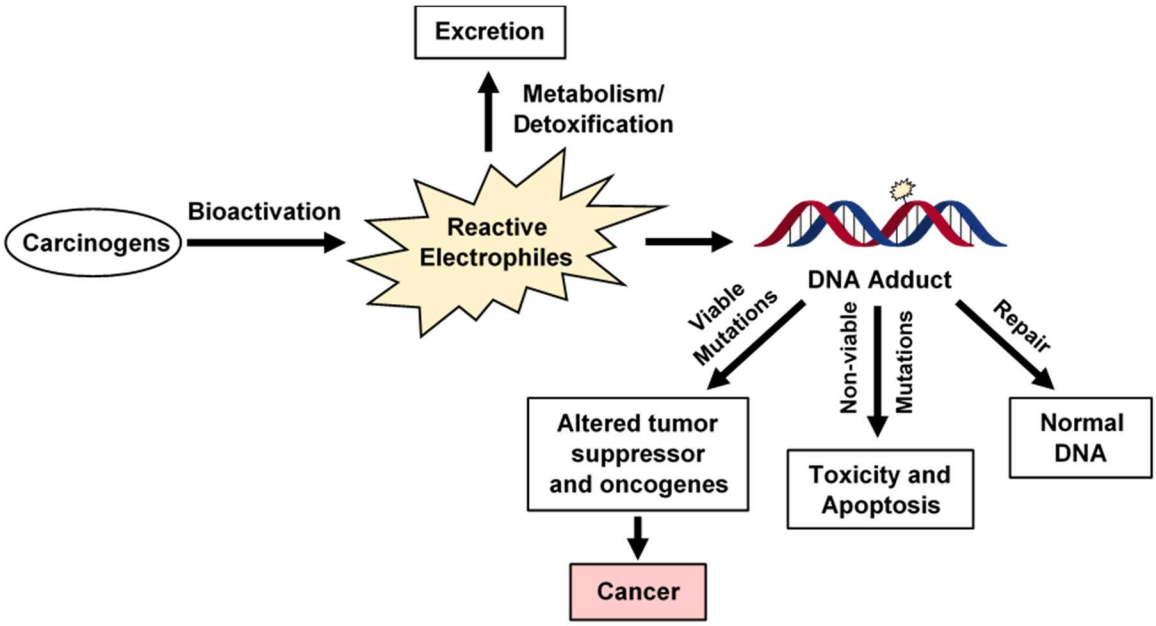
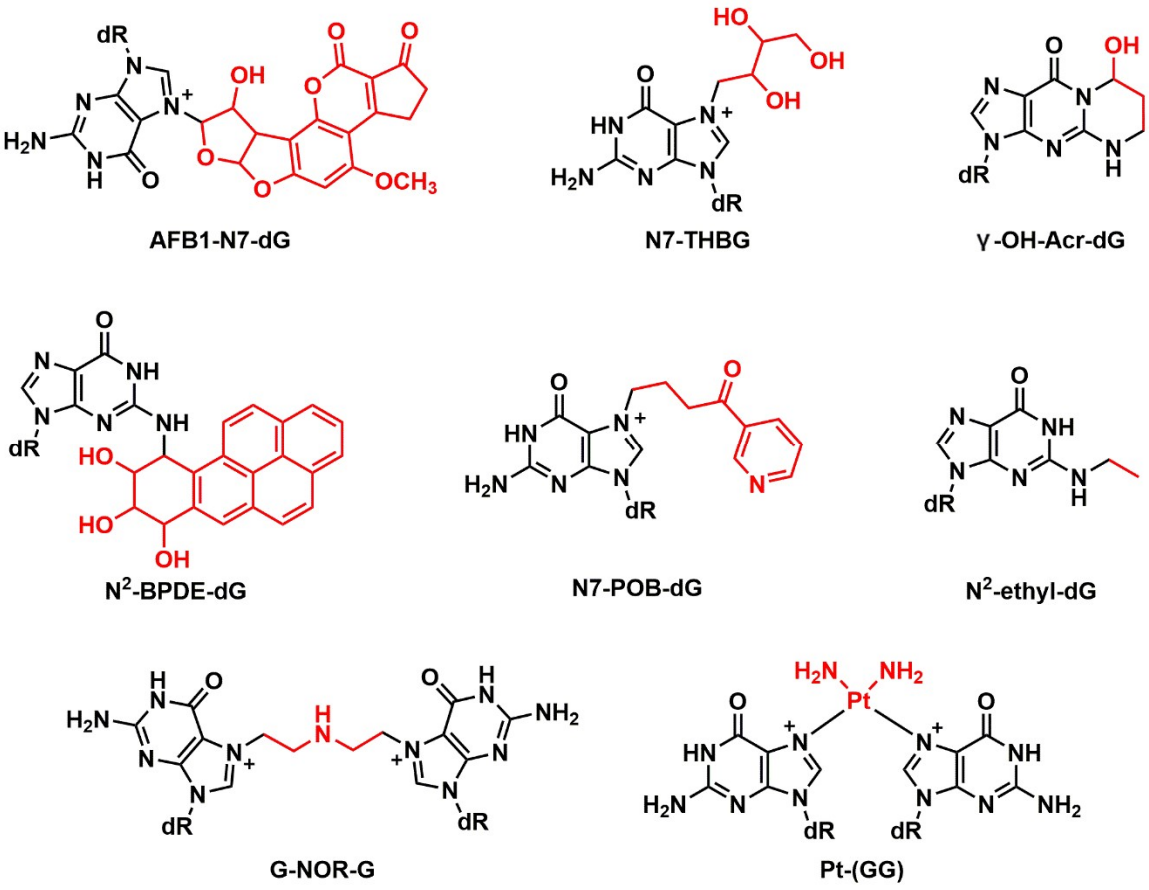
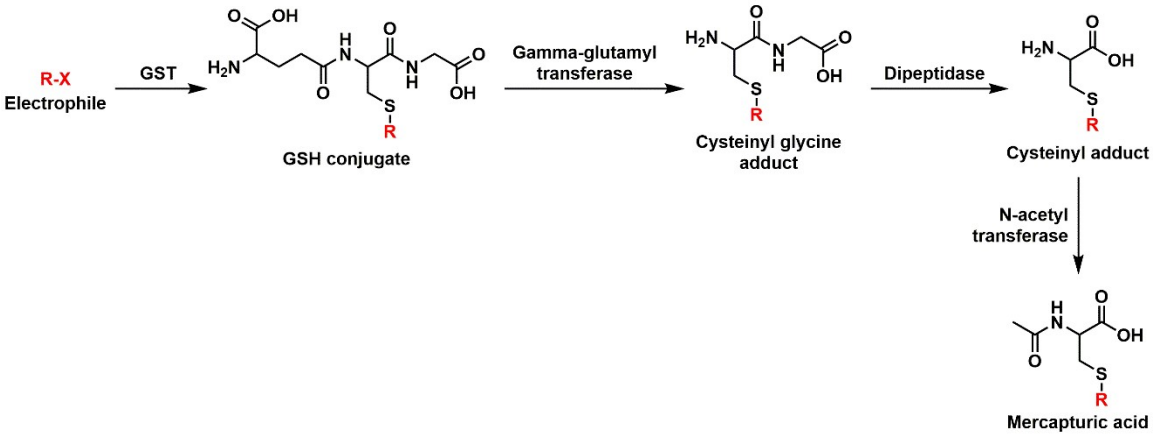


**Figure 1.1** Chemical structures of representative DNA adducts formed by chemical carcinogens and anticancer drugs. DNA adducts induced by aflatoxin B1 (AFB1-N7-dG), 1,3-butadiene (N7-THBG), acrolein (γ-OH-Acr-dG), benzo[a]pyrene (N2-BPDE-dG), 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (N7-POB-dG), acetaldehyde (N2-ethyl-dG), cyclophosphamide (G-NOR-G), and cisplatin (Pt-(GG)).



Many directly acting carcinogens and electrophilic metabolites can undergo metabolic detoxification via Phase II metabolic pathways to form harmless metabolites that can be excreted from the body.[28] These pathways involve the conjugation of molecules such as glutathione or glucuronic acid to neutralize reactive groups and/or increase excretion of carcinogens through bile or urine. Glutathione (GSH) conjugation is of particular importance for the detoxification of electrophiles. Glutathione, or γ-glutamyl cysteine glycine, has many roles in cells including as an antioxidant, an enzyme cofactor, a nontoxic form to store cysteine, and it facilitates xenobiotic metabolism and amino acid transport. It is highly prevalent in most cells, at concentrations between 0.5 and 12 mM. In xenobiotic metabolism, the thiol of GSH can react with electrophiles leading to non-reactive metabolites. This process can occur spontaneously or catalyzed by glutathione s-transferase enzymes. There are four different families of cytosolic GSTs: alpha, mu, pi, and theta, which include at least 11 isoforms with differing substrate specificities. GSH conjugates are further processed through the mercapturic acid pathway before excretion by the kidneys into urine (Scheme 1.2).

**Scheme 1.2** Glutathione conjugation and the mercapturic acid pathway.

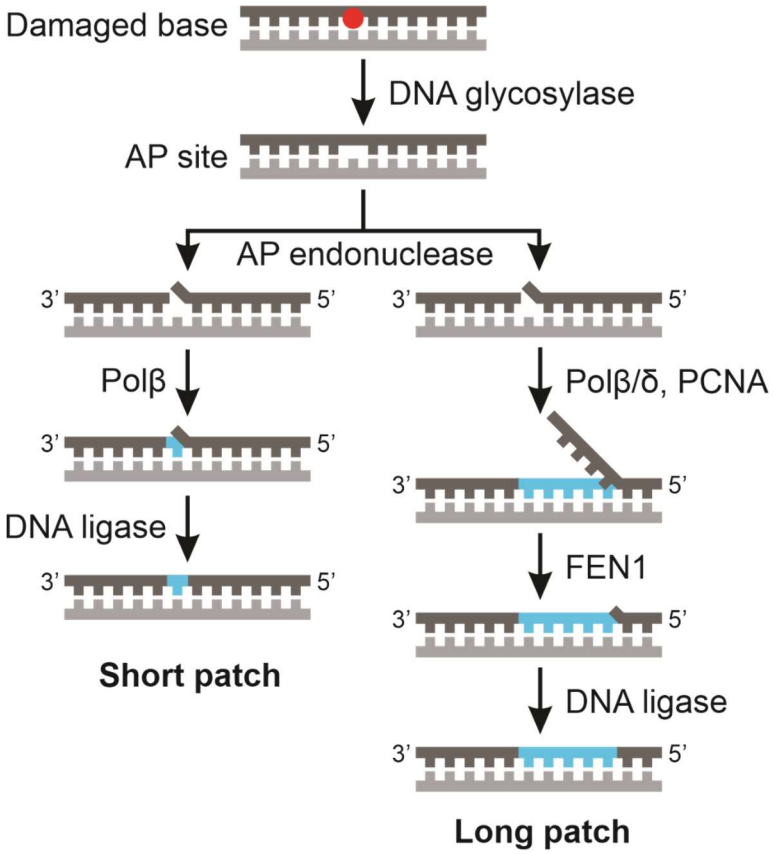


### 1.1.2 Enzymatic repair of carcinogen-DNA adducts

DNA repair pathways can prevent toxicity and mutations caused by DNA damage by restoring DNA to its native state. The main DNA repair pathways include base excision repair (BER), nucleotide excision repair (NER), and homologous recombination repair (HRR) pathways.[29] A specific DNA adduct may be recognized and repaired by just one, multiple, or none of these pathways.

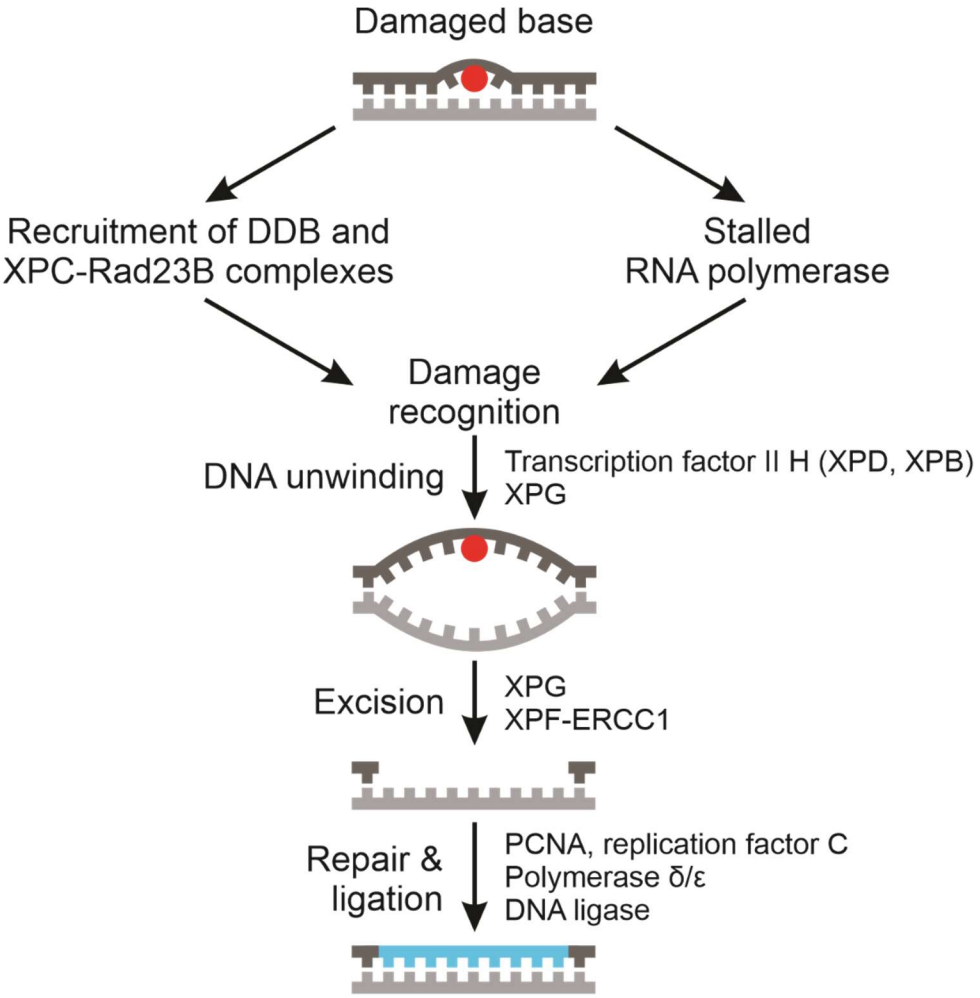
Base excision repair (BER) pathway is primarily involved in the repair of relatively small nucleobase adducts that do not significantly distort the DNA helix. These include imidazole ring opened purines (formamidopyridine or Fapy) adducts,[30] uracil resulting from cytosine deamination,[31] and 7,8-dihydro-8-oxoguanine (8-oxoG).[32] In this pathway, after the damaged nucleobase is recognized, DNA glycosylases cleave the N-glycosidic bond between the base and the deoxyribose sugar, resulting in an abasic (AP) site.[33] The DNA backbone is then cleaved by either AP endonuclease-1 (APEX1) or an AP lyase, creating a nick in the DNA strand. DNA polymerase β (Polβ) then inserts a correct nucleotide into position, and finally a DNA ligase seals the nick, resulting in a repaired DNA strand. This is known as the short patch BER pathway.[34] Alternatively, in long patch BER, APEX1 forms a 5' nick in the DNA strand and several enzymes are recruited. Proliferating cell nuclear antigen (PCNA) together with Polβ or DNA polymerase δ (Polδ) displaces the nicked DNA strand and polymerizes a new DNA strand of 2-12 nucleotides long. Flap structure-specific endonuclease 1 (FEN1) then removes the displaced section of DNA, allowing the repaired DNA to be ligated to the original strand (Scheme 1.3).[33]

**Scheme 1.3** Overview of base excision repair (BER) pathway.



Nucleotide excision repair (NER) pathway, in contrast to BER, is typically involved in the repair of bulky, helix-distorting lesions. In this pathway, the 3' and 5' phosphodiester bonds on either side of the DNA adduct are cleaved, and the strand containing the damaged nucleobase, ~25-30 nucleotides long, is excised. DNA polymerase δ or ε then fills in the gap, and the repaired strand is ligated with a DNA ligase.[35-36] NER involves a much larger complex of enzymes to recognize, remove, and repair DNA adducts than BER.[37] NER is further divided into two subtypes, global genomic NER (GG-NER) and transcription coupled NER (TC-NER). GG-NER is not dependent on transcription, and employs damage detecting proteins such as DNA-damage binding (DDB) and XPC-Rad23B complexes to find conformationally distorted DNA lesions. TC-NER, in contrast, does not require DDB or XPC complexes to detect DNA damage, and instead employs stalled RNA polymerases to detect DNA lesions (Scheme 1.4).[38]

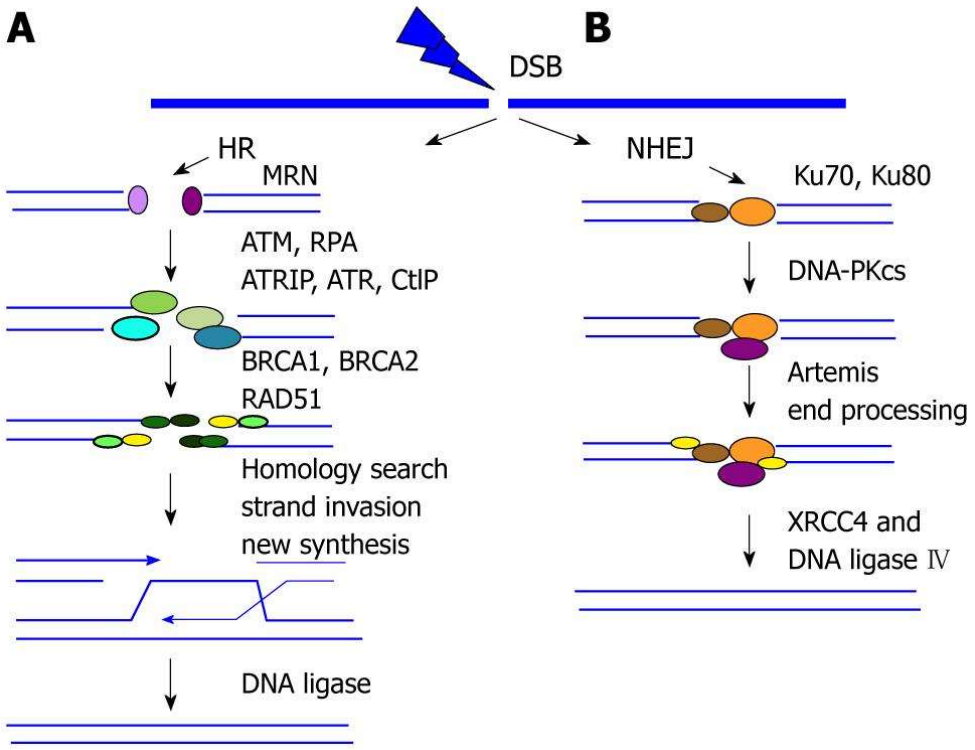
**Scheme 1.4** Overview of nucleotide excision repair (NER) pathway.



As stated previously, bifunctional alkylating agents can sequentially react with two sites of DNA to form intrastrand and interstrand DNA-DNA crosslinks. Interstrand crosslinks are of particular concern as they completely block DNA replication and can lead to toxic double strand breaks (DSBs). DSBs can be recognized and repaired by separate pathways such as homologous recombination repair (HRR) and Fanconi Anemia repair (FA) pathways.[39]

HRR requires a long homologous sequence, such as from the sister chromatid, to use as a repair template, and only occurs immediately before a cell enters mitosis. Sequence information is exchanged between the damaged and homologous DNA sequences, allowing the damaged strand to be correctly repaired (Scheme 1.5A).[40] This repair pathway also facilitates chromosomal crossover during meiosis, allowing homologous chromosomes to exchange regions leading to greater genetic diversity in gamete cells.[40] If the cell is not about to enter mitosis, or if a homologous repair template is not available, the nonhomologous end-joining (NHEJ) pathway can be used to repair DSBs. This pathway involves ligating the ends directly back together, but can be more inaccurate than HRR (Scheme 1.5B).[41] This is because it does not use a template for repair – any damaged or missing bases at the ends can lead to miscoding or deletions.

**Scheme 1.5** (A) Homologous recombination repair (HRR) and (B) nonhomologous end-joining (NHEJ) pathways. A: MRE11-RAD50-NBS1 (MRN) complex recognizes and senses DSBs and activates ATM kinase, initiating DNA damage response. CtIP-mediated nuclease activity forms single-strand DNA (ssDNA). Exposed ssDNA is coated with DNA replication protein A (RPA) and activates the Ataxia Telangiectasia and Rad3-related protein (ATR) response. RAD51 nucleoprotein filament replaces RPA-coated ssDNA, performs homology sequence searching, and mediates strand invasion. DSBs are restored by branch migration of this joint DNA molecule, DNA synthesis, ligation, and resolution of Holliday junctions; B: The two broken ends are processed and ligated directly by the end-binding KU70/80 complex and DNA-PKcs followed by XRCC4-ligase4. Figure adapted from Peng et al.[42]



The Fanconi Anemia (FA) repair pathway does not work on its own, but instead involves complexes of proteins from HRR, NER, and translesion synthesis.[43] It is specifically associated with the repair of interstrand crosslinks caused by ultraviolet radiation, ionizing radiation, and bifunctional electrophiles, such as nitrogen mustards and diepoxybutane.[44-45] It is named after a rare genetic disease, caused by mutations in proteins involved in FA complexes, which causes hypersensitivity to these crosslinking agents, leading to an markedly increased risk of cancer.[46]

### 1.1.3 Hemoglobin adducts as biomarkers of exposure to electrophiles

In addition to DNA adducts, electrophile adducts formed on human serum albumin (HSA) and hemoglobin (Hb) can be used as biomarkers of human exposures. Like DNA, HSA and Hb contain nucleophilic sites where electrophiles can form covalent adducts. DNA adducts are the most relevant exposure biomarkers as they directly reflect DNA damage that leads to carcinogenesis and mutagenesis. However, DNA adducts are often present at very low levels due to their rapid repair, which often makes them more difficult to detect. In contrast, HSA and Hb are the most abundant blood proteins and are much more abundant than DNA. In 1 mL of human blood, there is only approximately 0.005 – 0.008 mg of DNA, whereas there is ~30 mg of HSA and ~150 mg of Hb.[4] Additionally, protein adducts typically have much longer *in vivo* half-lives than DNA adducts which are constantly undergoing repair and thus can accumulate to high levels. Overall, this makes HSA and Hb adducts popular choices as biomarkers of exposure dose and screening.[4, 47-48]

Of these two blood proteins, Hb is an especially popular choice for use as an exposure biomarker. The major form of Hb in humans, Hb A, is a tetramer. All 4 subunits (2 α chains and 2 β chains) contain a nucleophilic N-terminal valine residue, which is very reactive towards electrophiles.[47] Its reactivity and prevalence (~150 mg/mL of human blood) make Hb a good choice for monitoring electrophile-induced adducts, and its long lifetime of 126 days in erythrocytes gives ample time for adducts to form and accumulate.[49-50] For example, quantitation of N-terminal valine Hb adducts in laboratory rats and mice treated with at low levels of 1,3-butadiene (BD) for up to 4 weeks showed that these adducts were stable *in vivo* and could be used to evaluate blood concentrations of the 1,3-butadiene metabolite 3,4-epoxy-1-butene (Scheme 1.7) during exposure.[51] As discussed in more detail in below (section 1.3.2.), N-terminal valine adducts have also been used for adductomic screens for unknown electrophile exposures.

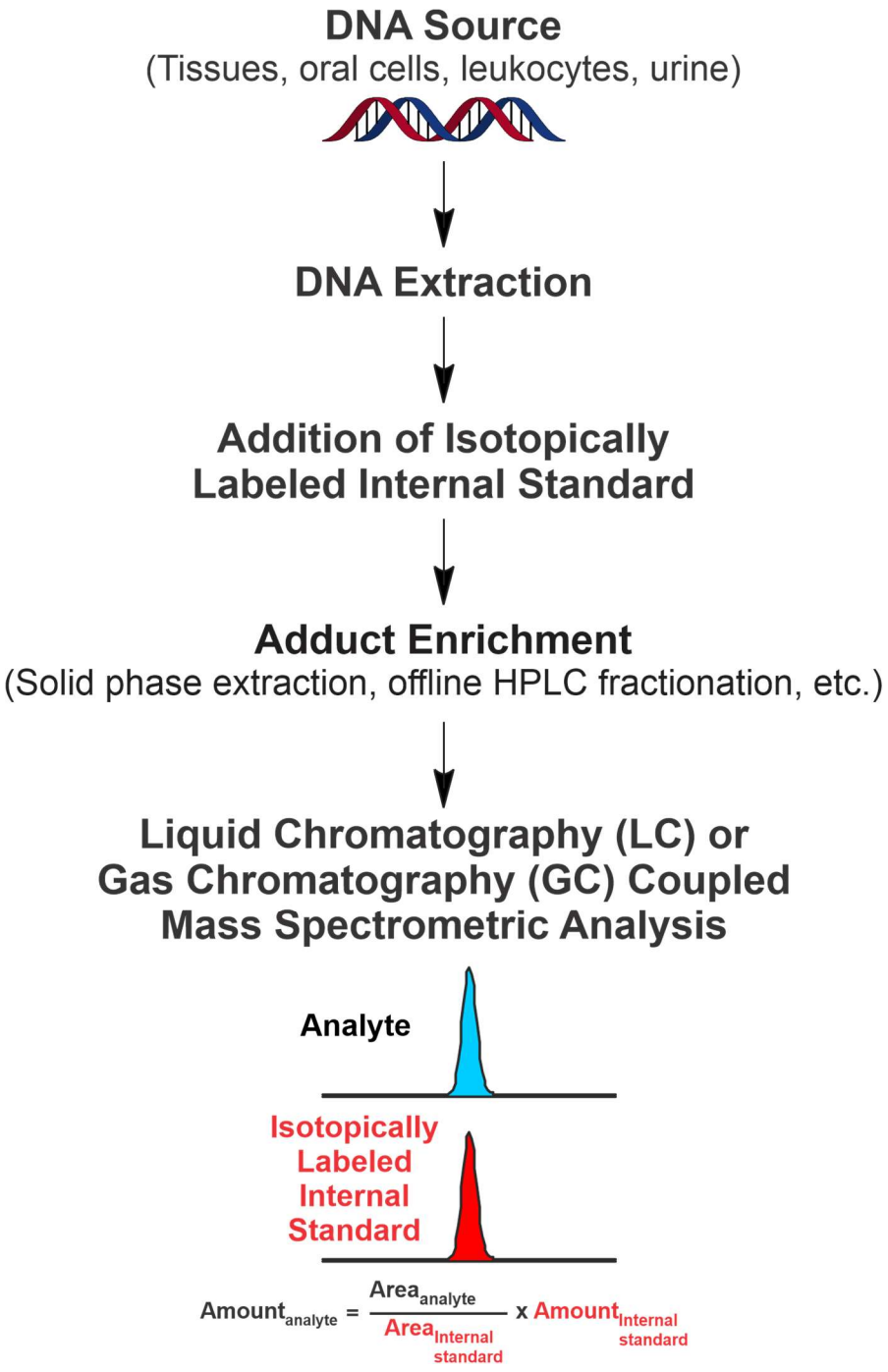
## 1.2 Mass spectrometry-based quantitation of biomacromolecule adducts

Recent advances in DNA and protein adduct analyses have been directly tied to the developments in mass spectrometry methodologies. Compared to older analytical methodologies such as $^{32}$P-postlabeling assays and high performance liquid chromatography-ultraviolet visible light absorbance detection (HPLC-UV/Vis), mass spectrometry (MS) based methods have higher sensitivity, selectivity, and can detect a wider range of compounds.[52-53]

### 1.2.1 Preparing DNA adducts for mass spectrometry analysis

Mass spectrometry, particularly isotope dilution MS, is currently the gold standard for absolute quantitation of DNA adducts.[54] Isotopically labeled internal standards are synthesized, where naturally occurring isotopes of hydrogen, carbon, or nitrogen at certain positions within the adduct are replaced by the stable isotopes $^{2}$H, $^{13}$C, and $^{15}$N. Thus, the synthesized isotopically labeled adduct has the same structure and behaves the same as the naturally occurring adduct, but has a distinct molecular mass due to the presence of an isotope label. This isotopically labeled compound is then used as an internal standard, with a known amount of the labeled standard added before sample preparation. Absolute quantitation is achieved by using the peak area ratios of the LC-MS/MS peaks for the unlabeled analyte and the labeled standard (Scheme 1.6). This accounts for any loss of analyte during sample preparation and adduct enrichment and also accounts for any matrix effects/signal suppression.[55]

**Scheme 1.6** General scheme for isotope dilution mass spectrometry-based analysis of DNA adducts.



#### *1.2.1.1 Methodology for sample preparation and DNA adduct enrichment*

In order to detect and measure DNA adducts, sufficient amounts of DNA must be procured. DNA can be extracted from blood leukocytes,[56] buccal cells,[57] exfoliated urothelial cells in urine,[58] and tissues such as the liver, lungs, and kidneys.[59] Next, DNA adducts need to be released from DNA. For hydrolytically labile DNA adducts, like N7-guanine and N3-adenine adducts where a destabilizing positive charge is present, neutral thermal hydrolysis can be used. In this method, the samples are heated at 70-90°C for 30-60 min, causing the destabilized N-glycosidic bond to cleave, releasing the free adducted base.[60] For adducts formed at other positions, nonselective nucleobase release methods can be used. Mild acid hydrolysis treatment can be used to release all purines from the DNA backbone.[61] Alternatively, complete enzymatic hydrolysis down to free nucleosides can be performed. This method involves digesting DNA using an enzyme mixture including deoxyribonucleases DNAse I and DNAse II and the phosphodiesterase nuclease P1 to form deoxyribonucleotides, and further digestion with alkaline phosphatase to free nucleosides.[62-63]

Because they are present in very low amounts as compared to native nucleobases or nucleosides (typically 1 adduct per $10^7$-$10^9$ nucleotides), DNA adducts must be enriched before MS analysis. If not removed, sample matrix may increase background noise or lead to ion suppression during MS analysis.[64] Ultrafiltration can be used to remove high molecular weight components, such as proteins or the DNA backbone, both of which can

clog the lines of the MS instrument.[54] Samples are loaded onto 10K or 3K Da molecular weight cutoff filters and then centrifuged to remove large molecules.

After filtration, further adduct enrichment needs to be performed. Solid phase extraction (SPE) is a commonly used methodology for this. SPE cartridges can be packed with a variety of different stationary phases, such as C18, phenyl, anion exchange, cation exchange, or mixed-mode. Depending on the properties of the DNA adduct of interest, different stationary phases are used to selectively retain adducts based on hydrophobicity or ionizability. Following binding of the adduct to the stationary phase, interfering compounds such as salts and other small molecules can be removed by washing the SPE cartridge with weakly organic, acidic, or basic solvents, after which the adduct of interest and its internal standard can be eluted off under higher organic, acidic, or basic conditions. One of the major benefits of SPE is that the cartridges are all single use, therefore eliminating possible carryover between samples. SPE cartridges are also available in 96-well plates, allowing these methodologies to be easily scaled to high throughput analysis. In our laboratory, SPE has been previously used for the enrichment of N-7-(1-hydroxy-3-buten-2-yl)guanine (EB-GII)[65] and N7-(2,3,4-trihydroxybut-1-yl)guanine (N7-THBG),[66] among others.

As an alternative to SPE, offline high pressure liquid chromatography (HPLC) can be used to enrich DNA adducts. This is more cost effective, and the ability to use a mobile phase gradient instead of separate washes leads to a more complete purification of DNA adducts. However, only one sample can be enriched at a time, leading to longer sample

processing times. Additionally, the fact that the same instrument and column is being used to process each sample can lead to analyte carryover, especially with "sticky" compounds such as guanine adducts. For low level adduct analysis, retention time markers need to be used to account for the influence of the sample matrix on retention times, as the chromatographic peak for the analyte will not be visible by UV. In our laboratory, offline HPLC enrichment has been used for many different adducts including N7-1,3-butadiene adducts EB-GII[67] and 1,4-*bis*-(guan-7-yl)-2,3-butanediol (*bis*-N7G-BD).[68]

### 1.2.2 Quantitative mass spectrometric methods for biomacromolecule adducts

Mass spectrometry is the preferred instrumentation for quantitation of biomacromolecule adducts. Published methods exist for a wide variety of adducts, from proteins and DNA to lipids. While gas chromatography coupled MS (GC-MS) has proven useful in certain applications, liquid chromatography coupled mass spectrometry (LC-MS) accounts for the majority of currently used methods.

#### *1.2.2.1 Liquid chromatography-mass spectrometry*

Before analyzing a sample, separation of the sample must be performed in order to allow for the mass spectrometer to have the highest sensitivity possible. This is frequently done by coupling a HPLC or UHPLC (ultra high performance liquid chromatography) instrument to the MS. These LC systems separate analytes of interest from interfering compounds and biological matrices as discussed in 1.2.1.1. When trying to measure very low adduct levels, preventing ion suppression from these interfering compounds is critical

to analyte detection. Reverse phase HPLC is most commonly used in LC-MS, which uses a hydrophobic stationary phase, such as C18, C8, or polyphenyl, to retain slightly nonpolar to very nonpolar compounds under highly aqueous conditions before eluting them off with increasing organic solvent concentration (usually acetonitrile or methanol). If an analyte is too polar to retain on a reverse phase column, a hydrophilic interaction liquid chromatography (HILIC) column and method may be used to simulate normal phase HPLC on a reverse phase instrument. Normal phase HPLC is the opposite of reverse phase, using a hydrophilic column to retain highly polar compounds under organic conditions, and then eluting them off with increasing aqueous solvent. Typically, flow rates used for LC-MS are much lower than for offline HPLC enrichment, including high flow (0.1-0.2 mL/min) down to nano flow (100-600 nL/min). High flow rates tend to lead to shorter methods and higher reproducibility, while lower flow rates lead to increased sensitivity due to more efficient ionization and ion transfer.[69]

The popularity of LC-MS likely owes a lot to compatibility of HPLC with electrospray ionization (ESI), or vice versa.[70] Analyte ionization is a critical step in MS analysis, as only charged ions can be seen by the MS. Positive or negative ion modes can be conducted by ESI, depending on what charge your analyte of interest can best hold. Many biomacromolecules, including DNA adducts, can be protonated under mildly acidic conditions, making positive ion mode the voltage of choice. Solvents used for LC-MS analysis are more limited than those used for offline HPLC, as only volatile acids and salts can be added to the aqueous and organic mobile phases. Commonly used additives include

0.01-0.1% formic acid or acetic acid, or 5-15 mM ammonium acetate. ESI is performed by applying a high voltage as the liquid comes off of the LC, forming a fine aerosol as molecules are ionized. As the aerosol approaches the inlet to the MS (the ion transfer tube), a combination of nitrogen gas (sheath and auxiliary gasses) and the high temperature of the ion transfer tube cause the solvent to further evaporate and the aerosol droplets to shrink, until they break into even smaller droplets. The final result is charged ions entering the MS. ESI is a soft ionization technique, as the ion is not fragmented before reaching the MS. Additionally, ESI can lead to multiply charged ions, allowing large molecules such as intact proteins to be seen, as with multiple charges, their charge ratio (*m/z*) can be reduced to the point that it is within the *m/z* range of the instrument.

#### *1.2.2.2 Triple quadrupole mass analyzers and selected reaction monitoring*

Triple quadrupole mass analyzers are a type of tandem mass spectrometer that are most frequently used for accurate quantitation. A triple quadrupole MS consists of three quadrupole mass analyzers. Each quadrupole is made up of 4 parallel rods with alternating radio frequency (RF) voltages, allowing them to filter specific ions based on their *m/z* ratio. Tandem mass spectrometry (MS/MS) in the selected reaction monitoring (SRM) mode, also known as multiple reaction monitoring (MRM) mode, is commonly used with triple quadrupole MS for quantitation. After the analyte is ionized by ESI and enters the MS, it enters the first quadrupole (Q1) where RF voltages are applied so that only ions with a specified *m/z* (the precursor ion) can make it through. Next, it enters the second quadrupole (Q2), which acts as a cell for collision-induced dissociation (CID), fragmenting the ion.

These fragments enter the third quadrupole (Q3), where once again only fragments of a specified *m/z* (the product ion) are allowed to pass through and hit the detector. Thus, the ions detected by the mass spectrometer must have both the specified *m/z* of the precursor ion and fragment to the specified *m/z* of the product ion. As ions fragment predictably, the peak detected for the analyte of interest can be quantified with high confidence due to the selectivity. When high resolution/accurate mass is not needed, triple quadrupole mass spectrometers using SRM are the instrument of choice for most trace level small molecule analysis, such as for DNA adduct analysis. The sensitivity of the SRM mode allows for typical limits of quantitation of quantitation for DNA adducts on these instruments to be very low, as low as 1-5 adducts per $10^9$ nucleotides.[71]

## 1.3 Adductomics

Adding the suffix -omics to the end of a field of biology refers to adopting a holistic view of components of an organism, rather than focusing on single molecules.[72] For example, genomics refers to looking at global changes in the genome rather than focusing on changes in an individual gene. With sequencing of the human genome and the advancements in technologies such as DNA sequencing and mass spectrometry, the focus on these global approaches has increased exponentially. This is known as the "omics revolution". By shifting focus to the whole rather than a specific piece of a specific condition, disease, system, etc., unknown components and interactions can be discovered. This approach has proved to be incredibly powerful, spurring the growth of fields including genomics, proteomics, metabolomics, and transcriptomics.

### 1.3.1 The exposome and the adductome

First introduced in 2005, the exposome was suggested as a complement to the genome.[73] While the sequence of one's genome could be studied, the effects of environmental exposure on a person, aka everything non-genetic, was largely unknown.[74] The exposome was introduced in an attempt to fill this gap. Genetic factors do not account for the entirety of disease development, and especially in chronic diseases and cancers, environmental factors are thought to be major contributors.[75-77] Thus, detecting and studying these environmental exposures is critical to fully understanding risk and disease development. The exposome aims to encompass all environmental exposures, both

exogenous and endogenous, a person is exposed to throughout their entire life. Knowing what these exposures are and how they factor into disease development will allow for more accurate risk assessment and hopefully disease prevention.

An important class of environmental chemicals includes electrophiles. As discussed previously, the ability of these molecules to form covalent adducts with nucleophilic sites in DNA and proteins can lead to adverse biological effects, such as carcinogenesis. However, their inherent reactivity, leading to short half-lives *in vivo*, causes them to be very difficult to accurately measure. However, by measuring the amount of adducts formed with DNA or proteins, the initial exposure can be detected, and a more biologically relevant measurement of active electrophile exposure can be made. The adductome is defined as the totality of these adducts in humans, and untargeted adductomic approaches are being developed to help detect and identify the entirety of the adductome.

As discussed in 1.1.3, the blood proteins HSA and Hb have been popular biomarkers of electrophile exposure. These proteins are especially useful in untargeted adductomics screens, where the amount of DNA needed to screen for low level unknown exposures makes these experiments less feasible. Untargeted adductomics methods using either HSA or Hb have been previously published and used to detect numerous adducts.[78] In HSA, the most common site for adductomic screens is Cys34, and 75 adducts at this site have been detected.[79-82] In Hb, the N-terminal valine is the most common site for adductomics screens, and 25 adducts have been detected so far at this site, 13 of which have been identified.[83-86] So far there has been little overlap between the identified adducts

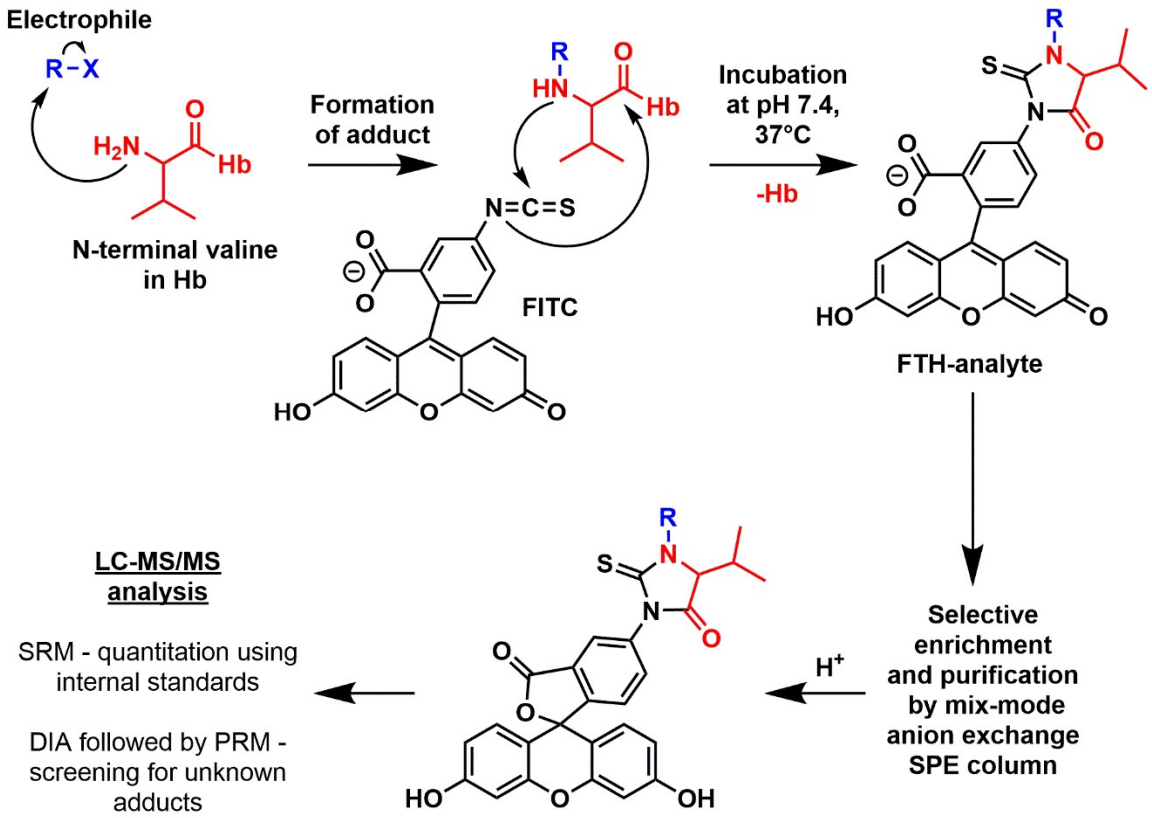
between HSA and Hb, emphasizing the need for studying adducts to both proteins, and the development of methodologies for detecting adducts at other nucleophilic sites.

### 1.3.2 Use of the FI*R*E procedure to identify novel hemoglobin adducts

The methodology that uses N-terminal valine (Val) of Hb for adductomics screens is known as the FI*R*E procedure. This procedure is based on a modified Edman degradation using fluorescein isothiocyanate (FITC) to detach and derivatize adducted N-terminal Val. The detachment of adducted Val is favored over non-adducted Val, allowing for the enrichment of derivatized adducts.[87-88] This derivatization procedure and the accompanying LC-MS methodology were developed in the Törnqvist lab, based off of a method developed by the Ehrenberg lab, both from Stockholm University.[87, 89]

In the FI*R*E procedure, 5 mg of fluorescein isothiocyanate (FITC) is added to 250 µL of whole blood, lysate, or red blood cells (RBCs) and allowed to react overnight while mixing at 37°C. This generates derivatives known as fluorescein thiohydantoins (FTHs) (Scheme 1.7). Proteins are then precipitated out with acetonitrile (ACN), and stable isotope labeled internal standards corresponding to FTH derivatives of known adducts are added. The derivatized sample is then purified using mixed-mode anion exchange SPE cartridges, as the carboxylic acid group on FTH holds a negative charge under basic conditions. The samples are then worked up in acid and subjected to LC-MS analysis (Scheme 1.7).

**Scheme 1.7** Derivatization and analysis of adducted N-terminal valine from hemoglobin using the FI*R*E procedure.



Derivatized samples can be analyzed using SRM mode on a triple quadrupole mass spectrometer looking for diagnostic fragments common to FTH fragmentation. However, different transitions must be used for each adduct due to their differing *m/z* and thus multiple injections must be done for each sample to cover a reasonable mass range, making this method incredibly slow adductomics screening.[83] To solve this problem, the use of high resolution mass spectrometers (HRMS), such as the Q Exactive hybrid quadrupole orbitrap MS (Thermo Scientific, MA, USA), have been especially advantageous.[90] Using these instruments, samples can be analyzed using the data independent acquisition (DIA) mode. In this mode, all ions with a selected *m/z* range are fragmented and analyzed. In order to maximize the number of adducts detected, the sample is injected multiple times (1 injection per 50 *m/z* range), but it is still faster than SRM using a triple quadrupole MS. Any adducts seen in DIA mode are then re-analyzed in parallel reaction monitoring (PRM) mode, a similar mode to SRM where specific precursor ions are selected for and fragmented.

A major disadvantage of Hb N-terminal Val adductomics screens compared to HSA Cys34 is that the selection of potential adduct must be done manually. The data resulting from DIA experiments requires extensive evaluation and unlike HSA Cys34, no in-house software exists for adduct detection. Adduct candidates are typically selected if at least 2 of the 4 diagnostic fragments common to FTH derivatives are observed. After confirming adduct detection, the lengthy process of adduct identification begins. From the adductomics screen, molecular weight (accurate mass if using HRMS), fragmentation pattern, and

retention time are known. If these do not match the values for a known adduct, the identity of the adduct must be hypothesized and reference compounds must be synthesized and compared to the unknown adduct. If accurate mass data is available, the elemental composition of the adduct can be found, greatly helping to propose the correct structure. Adduct retention time (RT) is also helpful in proposing adduct structure, as when using reverse phase chromatography, the retention time correlates with a molecule's lipophilicity, or LogP value.[85] As FTH derivatives have relatively low molecular weight compared to the much larger T3 peptide analyzed in HSA experiments, the R group has a large enough effect on RT for this information to be useful. Thus, despite initial adduct detection of Hb N-terminal Val adducts being more difficult than HSA Cys34 adducts, the identification of the structure of Hb adducts is actually easier than HSA adducts. However, adduct identification is still a long and sometimes expensive process.

## 1.4 Overview of 1,3-butadiene as a carcinogen

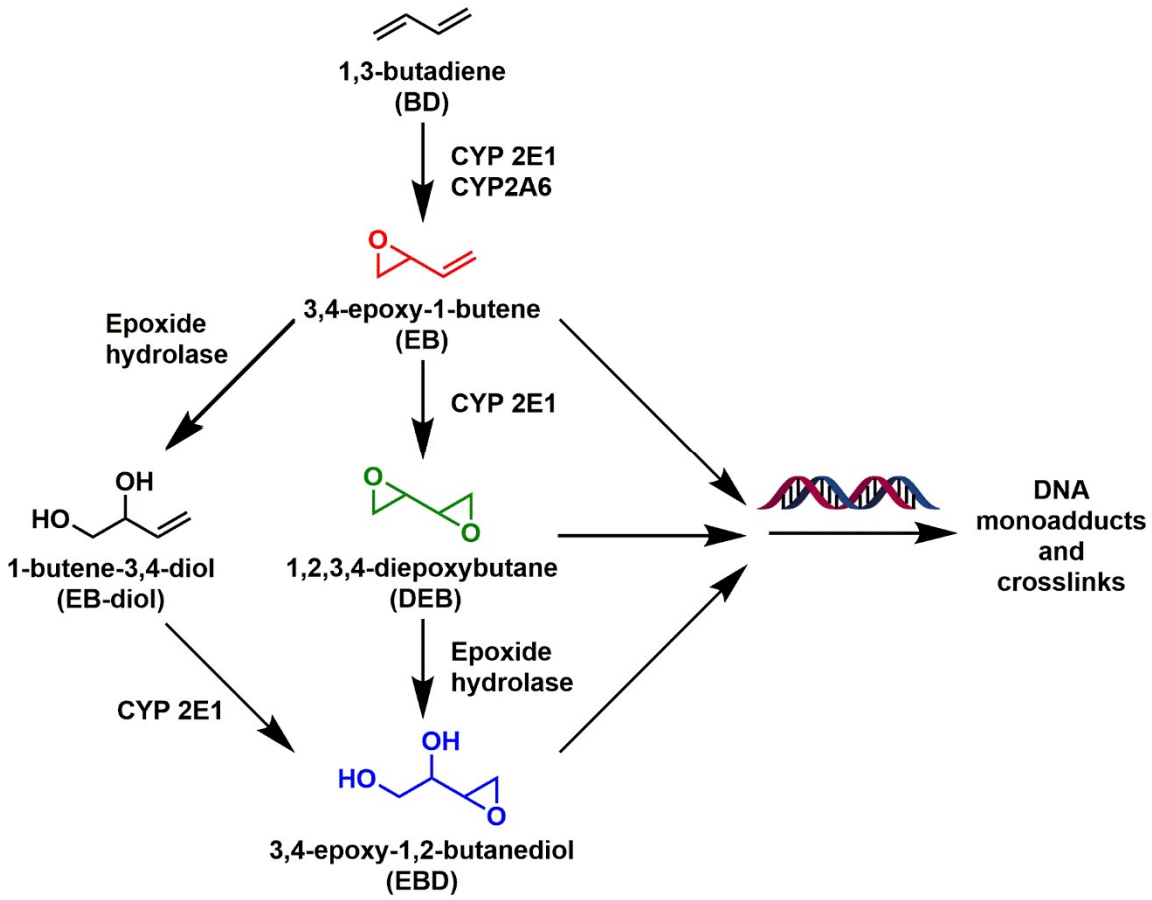
### 1.4.1 Sources of exposure and toxicity of 1,3-butadiene

1,3-Butadiene (BD) is designated as a group 1 carcinogen (carcinogenic to humans) by IARC.[91] It is a volatile, colorless gas or liquid that is used in the production of synthetic rubber and polymers.[92] Humans outside the polymer industry are also exposed to lower levels of BD from cigarette smoke, automobile exhaust, and wood burning fires.[93-95] In studies with both mice and rats, exposure to BD led to the development of tumors in multiple organs.[96-97] Additionally, occupationally exposed workers showed an increased risk of leukemia.[98-99]

### 1.4.2 Metabolism of 1,3-butadiene

BD on its own is not carcinogenic. However, following exposure, it is metabolically activated by cytochrome P450 monooxygenases CYP 2E1 and CYP 2A6 to the reactive epoxide 3,4-epoxy-1-butene (EB).[100] EB can then either undergo hydrolysis by epoxide hydrolase to form 1-butene-3,4-diol (EB-diol)[101] or oxidation by CYP 2E1 to form to diepoxide 1,2,3,4-diepoxybutane (DEB).[102] DEB can then be hydrolyzed by epoxide hydrolase to form 3,4-epoxy-1,2-butanediol (EBD), which can also be formed by CYP 2E1 oxidation of EB-diol.[103] These 3 reactive epoxides are responsible for the carcinogenicity of BD (Scheme 1.8).

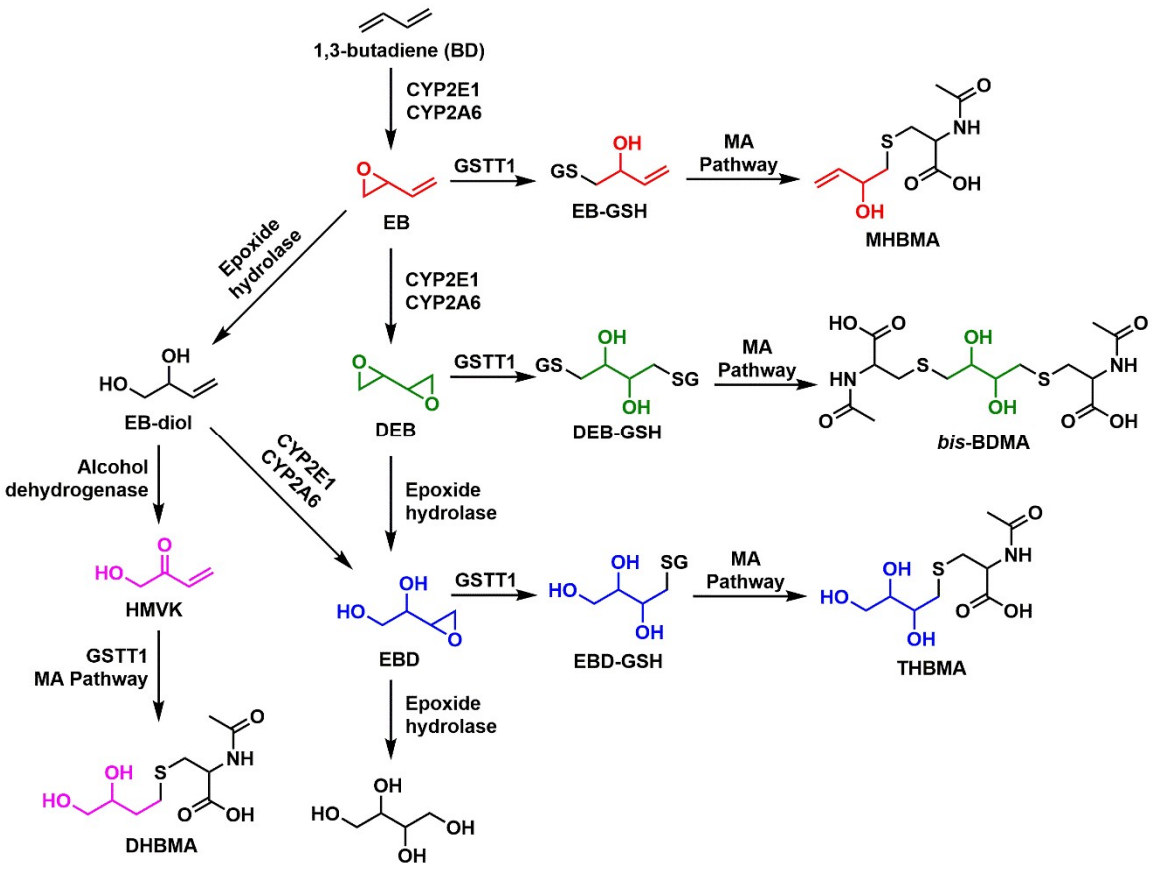
**Scheme 1.8** Metabolic activation of 1,3-butadiene (BD) to reactive epoxides.



EB, DEB, and EBD all contain electrophilic epoxides. These epoxides can form covalent adducts with nucleophilic sites in biomacromolecules including DNA, leading to toxicity and mutagenesis. EBD is the most abundant epoxide formed, followed by EB. While least abundant, DEB accounts for most of BD's genotoxic and carcinogenic effects, as it is a bis-electrophile and thus can form toxic DNA interstrand crosslinks[103-104] and DNA-protein crosslinks.[105-106]

Instead of forming DNA adducts, these reactive epoxides can undergo further metabolism and detoxification. Glutathione S-transferase (GST) mediated conjugation of these epoxides with glutathione (GSH) is the major route of detoxification (scheme 1.9). This conjugation can happen spontaneously, but the reaction rate is increased significantly through the use of GSTs. Among the GSTs, GST theta 1 (GSTT1) is thought to have the biggest role in 1,3-butadiene detoxification, though GST mu 1 (GSTM1) and GST theta 2 (GSTT2) may also be involved. These glutathione conjugates are further metabolized to mercapturic acids and excreted in the urine (scheme 1.9). In regards to cancer risk, it is important to consider both metabolic activation and detoxification pathways.

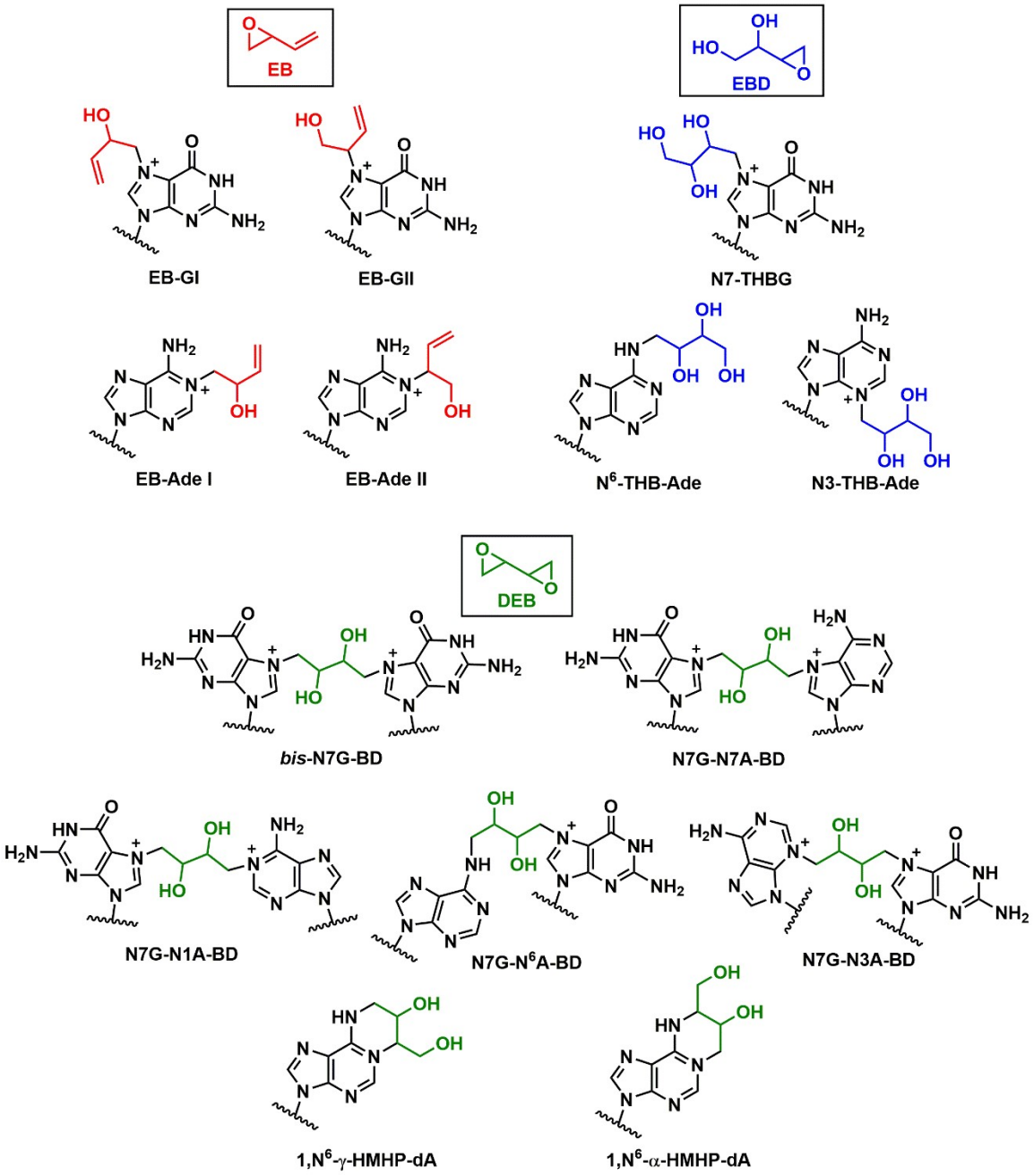
**Scheme 1.9** Detoxification pathways of 1,3-butadiene epoxides.



### 1.4.3 Biomarkers of 1,3-butadiene exposure

Many DNA adducts with EB, EBD, and DEB have been detected, as there are many nucleophilic sites within DNA where these reactive epoxides can bind. Of these, the N7 position of guanine is the most reactive site but adducts to other sites in guanine and adenine have also been observed. All of the following adducts are shown in Figure 1.2. Reported adducts from EB include N7-(2-hydroxy-3-buten-1-yl)guanine (EB-GI), N7-(1-hydroxy-3-buten-2-yl)guanine (EB-GII), N1-(2-hydroxy-3-buten-1-yl)adenine (EB-Ade I), and N1-(1-hydroxy-3-buten-2-yl)adenine (EB-Ade II).[16, 107-108] EBD has been observed to form adducts at the N7 position of guanine, N7-(2,3,4-trihydroxybut-1-yl)guanine (N7-THBG), and the $N^6$ and N3 positions of adenine: $N^6$-(2,3,4-trihydroxybut-1-yl)adenine ($N^6$-THB-Ade) and N3-(2,3,4-trihydroxybut-1-yl)adenine (N3-THB-Ade) ).[16, 107-108] While EB and EBD can only form monoadducts, DEB can form intra- and interstrand DNA-DNA crosslinks due to it being a bis-electrophile. These include 1,4-*bis*-(guan-7-yl)-2,3-butanediol (*bis*-N7G-BD), 1-(guan-7-yl)-4-(aden-1-yl)-2,3-butanediol (N7G-N1A-BD), 1-(guan-7-yl)-4-(aden-3-yl)-2,3-butanediol (N7G-N3A-BD), 1-(guan-7-yl)-4-(aden-6-yl)-2,3-butanediol (N7G-$N^6$A-BD), and 1-(guan-7-yl)-4-(aden-7-yl)-2,3-butanediol (N7G-N7A-BD).[109-110] In addition to DNA-DNA crosslinks, DEB can form exocyclic DNA adducts by reacting with two different sites on the same nucleobase. Examples of this are 1,$N^6$-(1-hydroxymethyl-2-hydroxypropan-1,3-diyl)-2-deoxyadenosine (1,$N^6$-α-HMHP-dA) and 1,$N^6$-(2-hydroxy-3-hydroxymethylpropan-1,3-diyl)-2-deoxyadenosine (1,$N^6$-γ-HMHP-dA).[111]

**Figure 1.2** Observed DNA adducts formed by EB, EBD, and DEB.



As it is the metabolically activated epoxides of BD that are responsible for toxicity and mutagenicity, rather than BD itself, BD-DNA adducts are of interest as use as biomarkers to estimate the effective dose of activated BD. Initial studies used $^{32}$P-post labeling methodologies to identify DNA adducts of EB, EBD, and DEB in mice and rats exposed to BD; however, sensitivity with these methods was limited.[112-114] Accurate quantitation of BD-DNA adducts became possible with the introduction of isotope dilution MS combined with LC-ESI$^+$-MS/MS analysis. This methodology was used to quantify EB-GI, EB-GII, and N7-THBG in the livers of mice and rats exposed to 1250 ppm BD over 10 days.[107] In our laboratory, sensitive LC-ESI$^+$-MS/MS methods have been developed for the quantitation for DEB induced DNA-DNA adducts *in vivo*. After mice were treated with 625 ppm BD for 5 days, levels of *bis*-N7G-BD adducts were measured in the liver (3.2 ± 0.4 adducts per $10^6$ dG) and lungs (1.8 ± 0.5 adducts per $10^6$ dG).[59] N7G-N1A-BD adducts were measured in the livers of BD exposed mice (3.1 ± 0.6 adducts per $10^8$ nucleotides), while levels of N7G-N3A-BD and N7G-N7A-BD were both below the limit of detection.[115] Both *bis*-N7G-BD and N7G-N1A adducts were measured in tissues of mice and rats exposed to increasing amounts of BD, where a difference in metabolism between the species was seen.[116] While adduct levels increased linearly from 0-625 ppm BD in mice, levels in rats was 4-10 fold lower and plateaued at 62.5 ppm BD.

More recently, we have developed ultra-sensitive methodologies for BD-DNA adduct analysis by employing nanoflow LC and/or HRMS$^3$ technologies. Sangaraju et al. developed a nanoLC-nanoESI$^+$-MS/MS method to quantify *bis*-N7G-BD levels in the

livers of mice exposed to sub ppm levels of BD.[71] A LC-HRMS/MS method using an Orbitrap Velos mass spectrometer was developed for the quantitation of N7-THBG in human leukocyte DNA from nonsmokers, smokers, and occupationally exposed workers (7.08 ± 5.29, 8.20 ± 5.12, and 9.72 ± 3.80 adducts per $10^9$ nucleotides, respectively).[66] Combining these two methodologies, a nanoLC-nanoESI$^+$-HRMS$^3$ method was developed and used to quantify EB-GII levels the livers of rats exposed to sub ppm levels of BD and the urine of BD exposed rats, occupationally exposed workers, and smokers.[65, 67] These studies mark the first time BD-DNA adducts have been detected and quantified in humans.

### 1.4.4 BD exposure in cigarette smoke and lung cancer risk

Lung cancer accounts for approximately 30% of cancer deaths in the US.[117] The biggest risk factor for lung cancer development is cigarette smoking, accounting for ~81% of lung cancer deaths and ~30% of all other cancer deaths.[118] Cigarette smoke contains a plethora of known human carcinogens, including heterocyclic aromatic amines, N-nitrosamines, polycyclic aromatic hydrocarbons, volatile hydrocarbons, nitrohydrocarbons, metals, phenolic compounds, formaldehyde, acetaldehyde, 1,3-butadiene (BD), isoprene, and benzene. Despite not being the most carcinogenic compound in cigarette smoke, BD is by far the most prevalent component, leading it to have one of the highest cancer risk index per cigarette per day.[119] Thus, BD exposure in cigarette smoke likely plays a role in lung cancer development.

## II. QUANTITATION OF THE FORMATION AND REPAIR OF *BIS*-N7G-BD

Adapted from:

Chesner, L. N.; Degner, A.; Sangaraju, D.; Yomtoubian, S.; Wickramaretne, S.; Malayappan, B.; Tretyakova, N.; and Campbell, C.; Cellular repair of DNA-DNA cross-links induced by 1,2,3,4-diepoxybutane. *Int. J. Mol. Sci.* **2017**, 18(5), 1086.

This work was performed in collaboration with Dr. Lisa Chesner, Dr. Dewakar Sangaraju, Shira Yomtoubian, Dr. Susith Wickramaratne, and Bhaskar Malayappan, under the direction of Dr. Natalia Tretyakova and Dr. Colin Campbell. Compounds were synthesized by Dewakar Sangaraju. Amanda Degner quantified adduct levels in cells by mass spectrometry.

and

Walker, V. E.; Degner, A.; Tretyakova, N.; Nicklas, J. A.; Walker, D. M.; and Albertini, R. J..; The t(9:22) chromosome translocation (Philadelphia Chromosome): Lack of induction by 1,2,3,4-diepoxybutane. *Chem. Biol. Interact.* Submitted June 2019.

Richard Albertini was the principal investigator for the project and developed the concept for the research and designed the initial mutagenesis experiments. Vernon Walker contributed to and optimized the cell culture experiments, prepared cell samples for other analyses, and wrote the bulk of the manuscript. Amanda Degner conducted and optimized the analyses for *bis*-N7G-BD. Janice Nicklas conducted the molecular studies for the t(9:22) translocations. Dale Walker performed statistical analyses.

## 2.1 Introduction

1,2,3,4-diepoxybutane (DEB) is considered the ultimate carcinogenic species of 1,3-butadiene (BD), a common environmental and industrial chemical present in cigarette smoke, urban air[10], and in rubber and plastic industries.[92] Long-term animal studies have demonstrated clearly that BD is a multisite carcinogen in mice and rats,[95] with vastly different cancer potencies in the two species.[91] In mice, lung tumors arose at inhaled concentrations as low as 6.25 ppm BD for two years, whereas, in rats, tumors have been induced at chronic exposures of 1000 ppm and higher (although lower exposures have not been evaluated).[91, 95] These interspecies differences in cancer susceptibility have been attributed to differences in the extent of metabolic activation of BD by cytochrome P450 monooxygenases (CYP 2E1 and CYP 2A6) to electrophilic epoxides such as 3,4-epoxy-1-butene (EB), 1,2,3,4-diepoxybutane (DEB), and 3,4-epoxy-1,2-butanediol (EBD) (Chapter 1, Scheme 1.8) and cellular repair of BD-induced DNA adducts.[120-121] In humans, occupationally exposed workers in the rubber and polymer industries have been reported to have an increased risk of developing leukemia.[98-99]

DEB is known to induce a variety of DNA lesions including nucleobase monoadducts, DNA-protein crosslinks, and both intrastrand and interstrand DNA–DNA crosslinks.[107, 122] Due to the presence of two epoxide rings in its structure, DEB sequentially alkylates guanine bases within DNA to form interstrand and intrastrand 1,4-*bis*-(guan-7-yl)-2,3-butanediol crosslinks (*bis*-N7G-BD). Previous studies

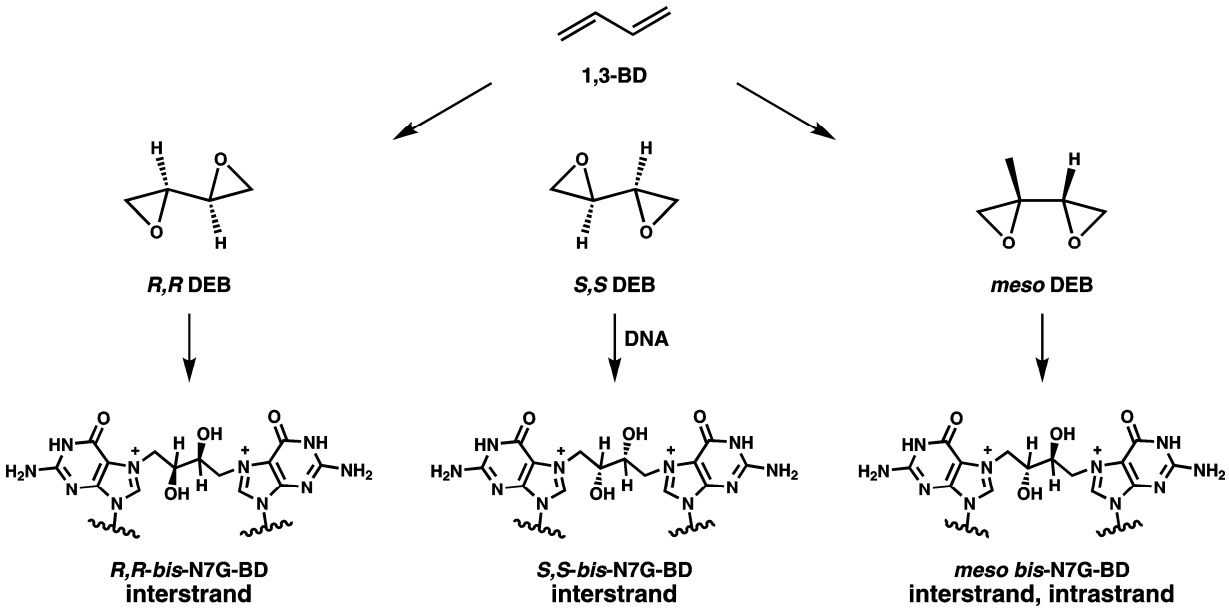
conducted in our laboratory have revealed that the DNA cross-linking specificity of DEB is dependent on its stereochemistry: while *SS*- and *RR*- DEB preferentially form interstrand crosslink adducts, the *meso*- isomer generates equal numbers of intrastrand and interstrand lesions (Scheme 2.1).[109, 123-124] Both intrastrand and interstrand *bis*-N7G-BD crosslinks are hydrolytically labile, with average half-lives of 3.5–4 days under physiological conditions.[125] *Bis*-N7G-BD adduct formation seems to differ by species. BD exposures have been shown to induce much higher levels of *bis*-N7G-BD adducts in laboratory mice as compared to rats exposed to the same levels of carcinogen,[116] which is accompanied by higher frequencies of chromosomal aberrations, sister chromatid exchanges, micronuclei (MN), and translocations.[126-128]

Epidemiological studies of BD exposure and cancer in humans have reported positive associations linked to specific sub-groups of leukemia including overall increases in chronic myeloid leukemia (CML).[99, 129-131] CML is defined by the presence of the Philadelphia chromosome, which is a translocation between chromosomes 9 and 22 (t9:22) that produces a fusion of the proto-oncogene *ABL*, located on the long arm of chromosome 9, with the *BCR* gene on chromosome 22, known as the breakpoint cluster region (bcr).[132] The *BCR-ABL* fusion oncogene, created mainly by one of the three functional t(9:22) translocations,[133] can produce somewhat different CML phenotypes by encoding a hybrid protein, Bcr-Abl tyrosine kinase, which is sufficient to initiate CML. Although RAS and other

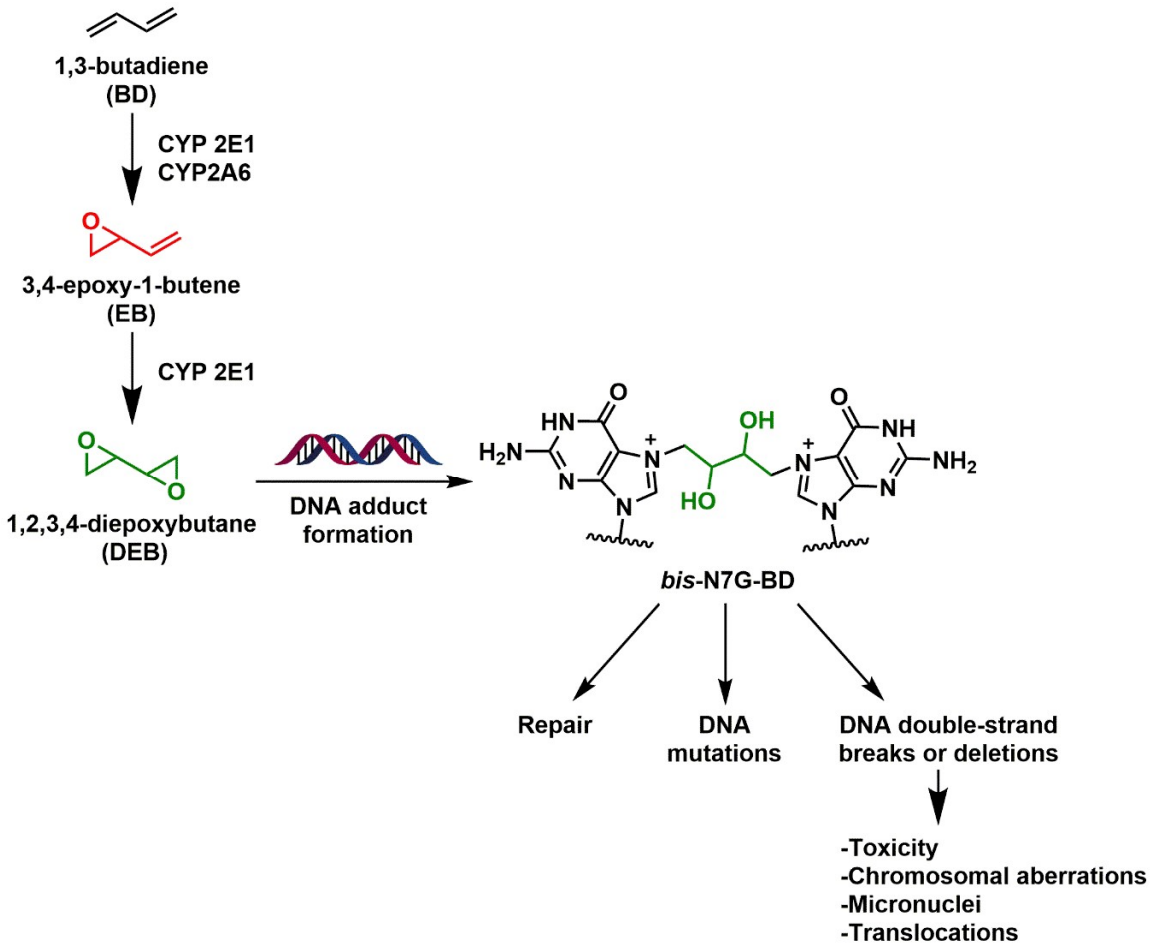
signaling pathway functions are aberrant in CML, these aberrations are triggered by downstream functions of the Bcr-Abl tyrosine kinase and are not due to primary oncogene mutations.[134-136] The ability of the Bcr-Abl tyrosine kinase to produce CML has been demonstrated by the efficacy of specific inhibitors of this kinase in inducing remissions.[137-138] Thus, no additional gene mutations are required. Reports of cancer patients receiving alkylating agents and developing CML as an adverse outcome have been questioned because of doubt that chemicals (medicinal or occupational) are capable of inducing t(9:22) translocations,[139] while CML is strongly associated with ionizing radiation and t(9:22) translocations can be induced by ionizing radiation *in vitro* in human blood cells.[139-142] A corollary to the above observations is that an agent capable of directly inducing CML also must be capable of producing pathogenic t(9:22) translocations.

Interstrand DNA-DNA crosslinks (ICLs) present a particular challenge to DNA replication, transcription, and repair because they inhibit local duplex melting at the site of adduct formation. Furthermore, both DNA strands at the site of the cross-link are damaged, potentially compromising their use as templates for accurate repair synthesis.[143] ICLs can induce point mutations, DNA double-strand breaks (DSBs), and large deletions (Scheme 2.2).

**Scheme 2.1** Metabolic activation of 1,3-butadiene to 1,2,3,4-diepoxybutane (DEB) and the induction of *bis*-N7G-BD DNA–DNA crosslinks.



**Scheme 2.2** Proposed mechanisms of genotoxicity induced by DEB.



The mechanism of ICL repair is a subject of active research. In *Escherichia coli*, elements of nucleotide excision repair (NER) and homologous recombinational repair (HRR) pathways are believed to collaborate in ICL repair.[144-146] In mammalian cells, there is evidence that ICL repair can also occur via a process that is independent of recombination. For example, NER-dependent ICL repair has been observed in G1 arrested cells, although it is not clear whether this mechanism is restricted to this phase of the cell cycle.[147-148] In addition, it is noteworthy that ICLs can be converted to DNA DSBs in yeast[149-151] and in mammalian cells[152-154]—an outcome not observed in bacteria.[146] Considerable uncertainty remains regarding the precise mechanisms through which ICLs are converted to DSBs. However, it appears likely that this process occurs when replication forks encounter ICL lesions during the S phase of the cell cycle.[152, 155-156] The identity of the nuclease(s) involved in ICL removal is yet to be conclusively established, although *XPF-ERCC1, MUS81-EME1, SLX1-SLX4, FAN1, SNM1A,* and *SNM1B* have been implicated.[157] In addition, translesion synthesis polymerases have been proposed to contribute to ICL tolerance and repair in eukaryotes.[146, 158] Xenopus egg extracts and a DNA template containing a site-specific ICL have been used as a system for investigating the molecular mechanism of replication-dependent ICL repair.[159] These studies implicate the involvement of the Fanconi anemia (FA) repair pathway in the repair of ICLS, namely through the involvement of the FANCI-FANCD2 complex[160] and RAD51.[161]

Fanconi anemia (FA) is a heterogeneous autosomal recessive disorder that predisposes individuals to cancer. Patients with FA and cells derived from FA patients display an enhanced sensitivity to DNA damaging agents that induce ICLs.[162] It has been reported that biallelic mutations in the human *XPF* gene cause FA.[163-164] The *XPF* gene encodes a protein that forms a heterodimer with the *ERCC1* gene product and participates in NER by cleaving DNA on the 5′ side of helix-distorting lesions.[165] More typically, inactivation of NER genes in humans is associated with the disorder Xeroderma Pigmentosum (XP). People with this disorder are predisposed to skin cancer, and cells derived from these individuals display hypersensitivity to ultraviolet radiation.[165] This previously unappreciated genetic connection between xeroderma pigmentosum and FA may help explain earlier observations that cells with defects in *XPF/ERCC1* are particularly sensitive to ICL-inducing agents,[166] while clones with defects in other NER genes display a more modest sensitivity to these agents.[167] This further suggests that there is crosstalk between the two DNA repair pathways,[168] and that a primary function of the FA pathway is to coordinate the cellular response to ICLs.[143, 168-169]

In this Chapter, we investigated the involvement of NER and FA repair pathways in repair of 1,3-butadiene induced *bis*-N7G-BD cross-links (Scheme 2.2) by following the formation and repair of these adducts in human dermal fibroblast cell lines with defects in the *XPA* and *FANCD2* genes (XPA and PD20 cell lines) and their isogenic corrected clones (XPAcorr and PD20corr cell lines). In the second set of experiments, we studied the relationship between *bis*-N7G-BD adduct formation, cytotoxicity, and micronuclei

formation in human promyelocytic leukemia HL60 cells treated with DEB to investigate whether BD exposure can induce t(9:22) translocations that lead to CML.

## 2.2 Materials and Methods

### *Materials:*

DEB is a known human and animal carcinogen and should be handled with adequate safety precautions in a well-ventilated fume hood strictly following its material safety data sheet.

LC-MS grade acetonitrile, methanol, and water were obtained from Fisher Scientific (Pittsburgh, PA, USA). NU7026 was from Selleckchem (Houston, TX, USA). Anti-H2AX antibody was obtained from Bethyl Laboratories (Montgomery, TX, USA), anti-cyclin B1 antibody was from Santa Cruz Biotechnology (Dallas, TX, USA). Goat anti-mouse secondary antibody (alexa fluor 488 conjugated) was from Thermo Scientific (Waltham, MA, USA), goat anti-rabbit secondary antibody (alexa fluor 633 conjugated) was from Invitrogen (Grand Island, NY, USA). *Bis*-N7G-BD and $^{15}N_{10}$-*bis*-N7G-BD (internal standard for mass spectrometry) were prepared in our laboratory as described previously,[59, 109, 125] and their concentrations were determined by UV spectrophotometry ($\varepsilon_{252}$ = 15700, pH 1). Reagents for cell culture were purchased from the following sources: Iscove's Modified Dulbecco's Medium (IMDM) with high glucose, L-glutamine, and 20 mM HEPES from the ATCC (Manassas, VA); fetal calf serum and glutamine from Hyclone Laboratories (Logan, UT); and penicillin-streptomycin from

Lonza BioWhittaker (Walkersville, MD) or MD Biomedicals (Santa Ana, CA). Cytochalasin B (>98%) was obtained from Sigma-Aldrich; nucleic acid isolation kits and PCR kits were acquired from Qiagen (Germantown, MD); and trypan blue, Quik Diff kit, methanol, agarose, DNA ladder size marker, and other miscellaneous chemicals/reagents, as well as plasticware and related supplies, were purchased from Fisher Scientific (Pittsburgh, PA). All other chemicals and solvents were obtained from Sigma-Aldrich (St. Louis, MO, USA).

***Cell culture***

A human dermal skin-derived clone[170] obtained from a patient with biallelic mutations in the *FANCD2* gene (PD20) and the corresponding retrovirally-corrected clone, PDcorr[171] were kind gifts from Dr. Alexandra Sobeck, University of Minnesota. Immortalized human dermal fibroblasts from an XP patient with inactivating mutations in the NER *XPA* gene (XPA), as well as from a gene-corrected clone derived from this line (XPAcorr), were obtained from the Coriell institute (GM04312). Human cells were cultured in Dulbecco's modified Eagle's media (Life Technologies, Grand Island, NY, USA) supplemented with 9% fetal bovine serum. Cells were maintained in a humidified atmosphere of 5% carbon dioxide, 95% air, at 37 °C.

The human promyelocytic leukemia cell line HL60 (ATCC® CCL-240™) was obtained from the ATCC and cultured according to recommended conditions with minor modifications.[133] HL60 cells were cultured in IMDM supplemented with 20% fetal calf

serum, 1 mL additional glutamine (200 mM or 100X), and 1 mL penicillin-streptomycin (5000 IU/mL) per liter of medium in a humidified incubator of 5% carbon dioxide, 95% air, at 37 °C. The Trypan Blue Exclusion Method for stained and unstained (viable) cells counted in a hemocytometer was used to determine/monitor cell growth and cell viability. Relative percentage of cell survival estimates were based on the growth curves for DEB-exposed cells versus vehicle control cells, with the average cell counts for control cells at differing time points post-exposure set at 100%.

***Cytotoxicity assays***

For direct-counting assays, cells (PD20, PD20corr, XPA, and XPAcorr) were seeded into 6 cm dishes at a density of $0.5 \times 10^6$ cells/dish. On the following day, cells were exposed to various concentrations of DEB (1–30 μM) for 3 h in serum-free media. Following treatment, the media was replaced with DEB-free normal growth media. Twenty-four hours later, cells were harvested and counted using a hemocytometer. Trypan blue exclusion confirmed that ≥99% of recovered cells were viable. Percent cell survival was calculated by dividing the number of cells obtained from DEB-treated cultures compared to the corresponding number of cells recovered from control cultures not exposed to DEB. All experiments were performed in triplicate, and the results represent the average ± the standard error of the mean from three or more independent experiments.

To examine the influence of inactivating cellular non-homologous DNA end-joining activity, cytotoxicity experiments were performed in the presence or absence of the

DNA-PK inhibitor NU7026 (Selleckchem, Houston, TX, USA). Human PD20 and PDcorr cells were exposed to DEB (5 μM) for 1 h in the presence or absence of 10 μM NU7026, in triplicate. At the end of the 1 hr incubation, drug-containing media was replaced with drug-free media, allowed to recover for 24 h, and then counted as described above. The results from these experiments represent an average ± the standard error of the mean from two independent experiments.

***Bis-N7G-BD adduct detection in DEB treated cells***

Human XPA, XPAcorr, PD20, and PDcorr cells (10 million, in triplicate) were treated with 15 μM DEB for 3 h. Following treatment, DEB-containing media was replaced with fresh media, and the cells were allowed to recover for 0, 3, 24, or 72 h to allow for adduct repair. To quantify the *bis*-N7G-BD adducts remaining at these time points, cells were harvested with 5 mL of PBS, sedimented, and stored at −20 °C until DNA extraction and adduct analysis as described below.

HL60 cells in log phase growth (15 million cells in 30 mL supplemented medium/flask) were exposed for 3 h to 0, 1.0, 1.7, 2.5, 4, 7, or 10 μM DEB ($n$ = 7 replicates/exposure level). Of these, 10 million cells/flask were centrifuged, washed with PBS, pelleted, and stored at −20 °C until DNA extraction and measurement of *bis*-N7G-BD adducts. The remaining five million cells per flask were held for determining relative cell survival and micronuclei formation.

DNA was extracted from treated cells using the Qiagen Puregene DNA extraction solution set (Qiagen, Hilden, Germany). Briefly, 3 mL of cell lysis buffer and 15 µL of proteinase K solution were added to samples of ~10 million cells. These were incubated overnight at room temperature upon slow invert mixing. The following day, RNA digestion was performed by adding 15 µL of RNAse A solution and incubating samples at room temperature for 3 hrs. Proteins were then precipitated out by adding 1.5 mL protein precipitation solution, vortexing samples for ~20 sec, and centrifuging at 2000xg for 15 min. DNA was then precipitated by adding 100 µL of 5 M $NH_4OAc$ and 4 mL ice-cold IPA to each sample and storing overnight in a -20°C freezer. Extracted DNA was then washed twice with 1 mL of 70% EtOH in water and sheered using 19 and 22 gauge needles. DNA concentrations were estimated using a nanodrop UV spectrophotometer (Thermo Scientific, Waltham, MA) based on the absorbance at 260 nm. DNA purity was assessed from $A_{260}/A_{280}$ absorbance ratios, which were typically between 1.8 and 1.9. DNA (50–150 µg) was re-suspended in water (200 µL) and spiked with $^{15}N_{10}$-*bis*-N7G-BD (50 fmol, internal standard for mass spectrometry). Samples were subjected to neutral thermal hydrolysis (70 °C for 1 h) to release *bis*-N7G-BD adducts from the DNA backbone as free base conjugates. Following ultrafiltration to remove partially depurinated DNA, *bis*-N7G-BD and $^{15}N_{10}$-*bis*-N7G-BD were isolated by offline HPLC as described elsewhere.[172] HPLC fractions corresponding to *bis*-N7G-BD and its internal standard were dried under reduced pressure, reconstituted in water (25 µL), and subjected to nanoLC-nanoESI$^+$-MS/MS analysis as described previously.[71] Selected reaction monitoring transitions used

for quantitation of *bis*-N7G-BD were *m/z* 389.1 $[M + H]^+ \rightarrow$ 238.1 $[M + H - Gua]^+$, *m/z* 152.1 $[Gua + H]^+$ and *m/z* 399.1 $[^{15}N_{10}\text{-}M + H]^+ \rightarrow$ 243.1 $[M + H - [^{15}N_5]Gua]^+$, *m/z* 157.1 $[^{15}N_5\text{-}Gua + H]^+$ for $^{15}N_{10}$-*bis*-N7G-BD. Quantitation was conducted using nanoLC-nanoESI$^+$-MS/MS areas in extracted ion chromatograms corresponding to the analyte and its internal standard.[172] Calibration curves were constructed with authentic standards of *bis*-N7G-BD and $^{15}N_{10}$-*bis*-N7G-BD.

## 2.3 Results

### 2.3.1 Cell lines

The first series of experiments was performed in human cell lines deficient in the *FANCD2* gene (PD20), the NER *XPA* gene (XPA), and on gene-corrected clones derived from these respective cell lines (PD20-corrected and XPA-corrected). FANCD2, or Fanconi anemia group D2 protein, is activated by the Fanconi anemia complex to its monoubiquitinated form in response to DNA damage, recruiting FAN1[173-175] and XPF-ERCC1,[176] nucleases that facilitate the unhooking of the ICL from DNA. Thus, deficiency of *FANCD2* should inhibit the involvement of the FA pathway in *bis*-N7G-BD repair after DEB exposure. XPA, or Xeroderma pigmentosum complementation group A protein, plays a key role in the NER pathway. XPA binds to damaged DNA strands and acts as a scaffold for transcription factor II H, XPF-ERCC1, and replication protein A (Chapter 1, Scheme 1.4), facilitating the unwinding of the DNA duplex and the unhooking of the ICL from DNA.[167] Thus, *XPA* deficiency was expected to inhibit the involvement of the NER pathway in *bis*-N7G-BD repair in cells exposed to DEB.

### 2.3.2 Human cell viability in the presence of 1,2,3,4-diepoxybutane

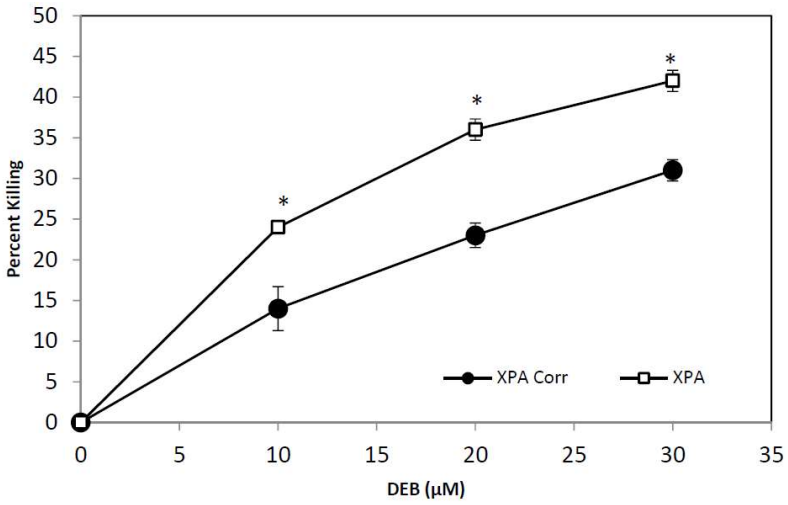
To establish the effects of DEB exposure on cell viability, we performed direct-cell counting cytotoxicity experiments in DEB treated cells. These experiments indicated both NER and FA-deficient human cell lines were significantly more sensitive to DEB than their corrected clones. *FANCD2*-deficient clones (PD20) showed increased cellular sensitivity

to DEB-induced cell death by nearly two-fold ($LC_{50}$ 24 vs. 44 μM) while *XPA*-deficient clones (XPA) were approximately 40% more sensitive ($LC_{50}$ 40 vs. 55 μM) than their gene-corrected counterparts (Figure 2.1).

Two recent papers reported that inactivation of a number of NHEJ genes can rescue, or partially rescue, the hyper-sensitivity phenotype of FA-deficient clones to DNA cross-linking agents,[177-178] suggesting that a function of the FA pathway is to prevent inappropriate end-joining repair of double strand breaks induced by cross-linking agents. To explore this possibility, we determined the influence of the DNA-PK inhibitor NU7026 on cellular survival of DEB-treated PD20 and their gene-corrected control. As the results in Table 2.1 reveal, exposure to 10 μM NU7026 had no significant effect on the sensitivity of either clone to DEB.

**Figure 2.1** Effects of exposure to DEB on cell death of (A) human wild-type (circles) and XPA (squares) and (B) human wild-type (circles) and PD20 (diamonds) cells. The influence of exposure to various concentrations of DEB for 3 h on cell survival was determined through direct cell counting of trypan blue-excluding cells. Results represent average ± the standard error of the mean, N = 3 or more *P<0.05.

A.




B.

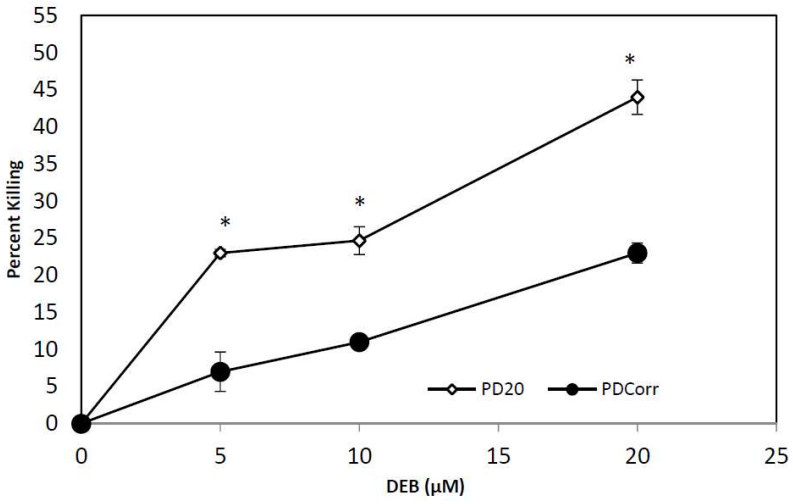


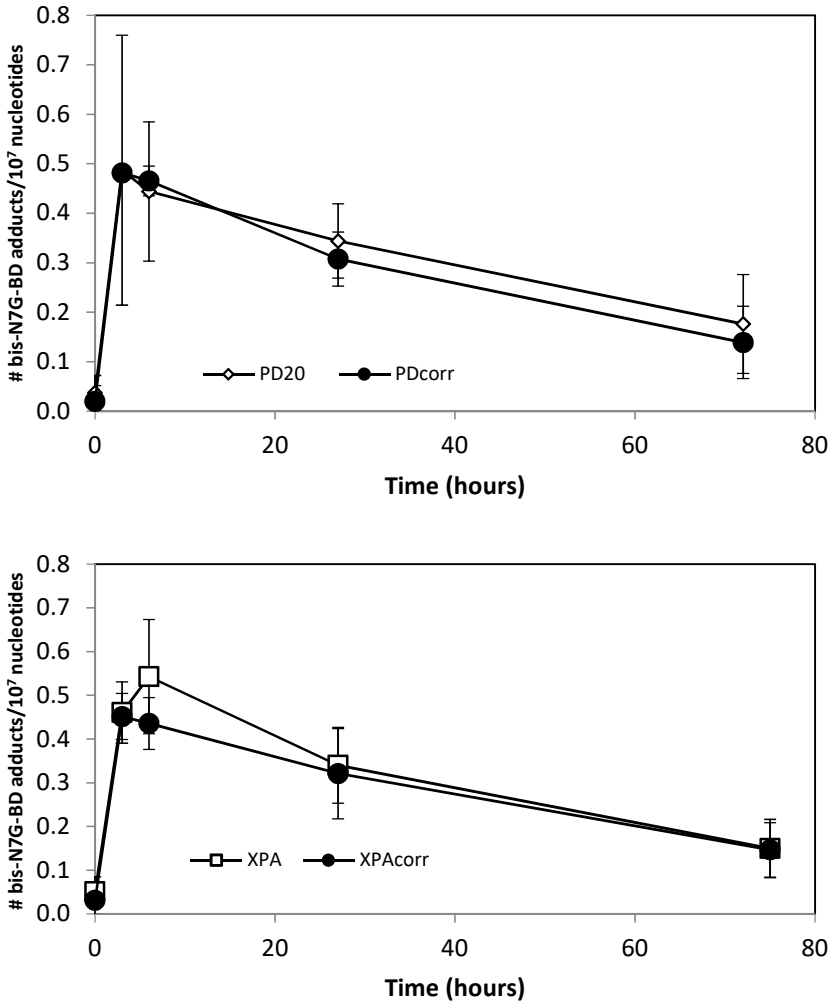
**Table 2.1** Effect of DNA-PK inhibitor NU7026 on 1,2,3,4-diepoxybutane (DEB)-induced cell death in human cells. Cells were exposed to DEB in the presence or absence of 10 μM NU7026 and the percent cell death determined as described in the Methods section.

| **Cell Line** | **PDcorr (WT)** | | **PD20 (-*FANCD2*)** | |
|---|---|---|---|---|
| NU7026 | - | + | - | + |
| Percent cell death (± SEM) | 21 (± 1.4) | 20 (± 7.8) | 40 (± 2.8) | 42 (± 8) |

### 2.3.4. *Bis*-N7G-BD adduct formation in FA or NER-deficient human cells treated with DEB

To examine ICL formation and removal in wild-type and repair-deficient human cell lines treated with DEB, PD20, PD20corr, XPA, and XPAcorr cell lines were treated with 15 μM DEB for 3 h. Following replacing the media with DEB-free media, cells were allowed to recover for 0, 3, 24, and 72 h. Cells were harvested and DNA was extracted. Samples were enriched for *bis*-N7G-BD and analyzed by nanoLC-nanoESI$^+$-MS/MS as described above. These results revealed that the kinetics of ICL formation and removal in *FANCD2*- or *XPA*-deficient clones did not significantly differ from that observed in isogenic wild-type control lines (Figure 2.2). The ICL kinetics data obtained from human DEB-treated cells are incompatible with the hypothesis that sensitivity to DEB is determined by the efficiency with which cells remove drug-induced ICLs.

**Figure 2.2** Time course for the formation and removal of bis-N7G-BD cross-links in human wild-type (circles), XPA (squares), and cells PD20 (diamonds). Cells were exposed to 15 μM DEB for 3 hours and then replaced with fresh growth media. Chromosomal DNA was isolated at the indicated times, and bis-N7G-BD adduct levels were quantified by nanoLC-ESI+-MS/MS as described in the Methods section. Results represent average ± the SEM, n = 3.



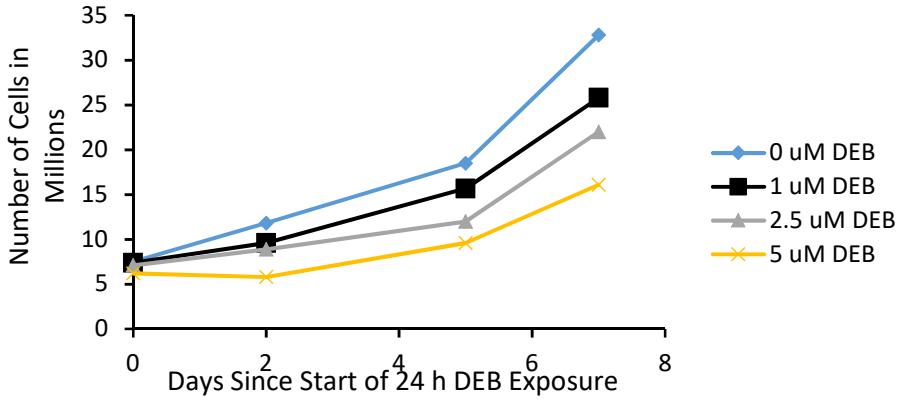
### 2.3.7 Cytotoxicity and induction of micronuclei in HL60 cells exposed to DEB

Several experiments were performed to characterize the change in cell growth and the relative cell survival over time for up to seven days after completion of 24 h exposures of HL60 cells to differing concentrations of DEB. This allows the ability of HL60 cells to recover from DEB exposure to be studied. Across these experiments, flasks of HL60 cells (5 million/flask) were exposed to 10 concentrations of DEB including 0, 0.5, 1, 2, 2.5, 4, 5, 6, 8, and 10 μM. Cell growth curves from the first experiment revealed that after a 24 h exposure to 0, 1, 2.5, or 5 μM DEB, a span of two to three days in culture was required for cytotoxicity to occur and surviving cells to recover from the treatment and return to log phase growth (Figure 2.3*A*). During this period of recovery and return to exponential cell growth, the numbers of treated cells at differing doses of DEB relative to the numbers of unexposed control cells (with the control values set at 100% at each time point) was fairly constant, as illustrated in Figure 2.3*B*. Thus, for delineating the shape of the dose-response curve for DEB-induced cytotoxicity, the relative cell survival of treated cells was measured at two to five days after completion of DEB exposure. The dose-response curve in Figure 2.4 shows that cell survival declined in a linear fashion with increasing concentration between 0 to 6 μM DEB ($R$, –0.988, Pearson's Correlation test), before yielding almost complete cell death with only 4.4, 3.3, and 1.0 per cent relative cell survival following exposures to 6, 8, and 10 μM DEB, respectively. The shape of the dose-response curve for relative cell survival in DEB-exposed HL60 cells is generally consistent with that

previously found for cell viability and growth following DEB exposures of TK6 human lymphoblastoid cells.[103, 179-180]

**Figure 2.3** Relationship between dose of DEB and the growth and relative cell survival of HL60 cells. HL60 cells in log phase growth were exposed for 24 h to 0, 1, 2.5, or 5 μM DEB (n = 3 replicates/exposure level) and cells were counted 0, 2, 5, and 7 days after completion of the exposure period. (A) Growth of HL60 cells after completion of a 24 h exposure to DEB. (B) The relative cell survival was calculated for each dose of DEB at each time point post-exposure. The plotted data show that the numbers of treated cells at differing doses of DEB relative to the numbers of sham-exposed control cells (with the control values set at 100% at each time point) was fairly constant during a 7-day period of recovery and return to exponential cell growth.

A.




B.

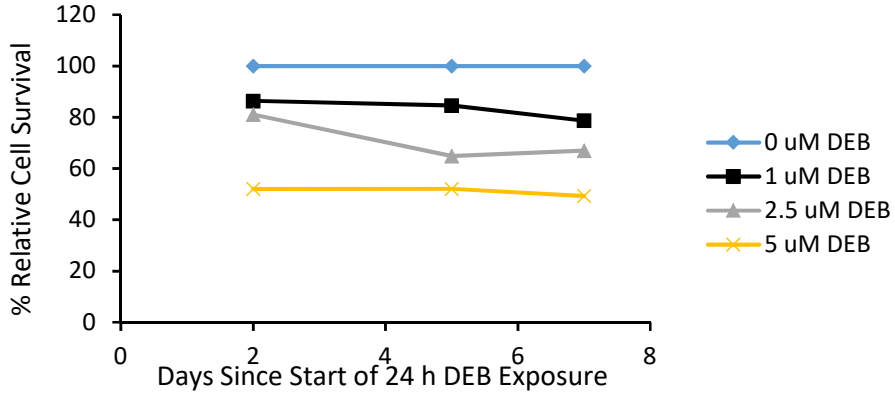
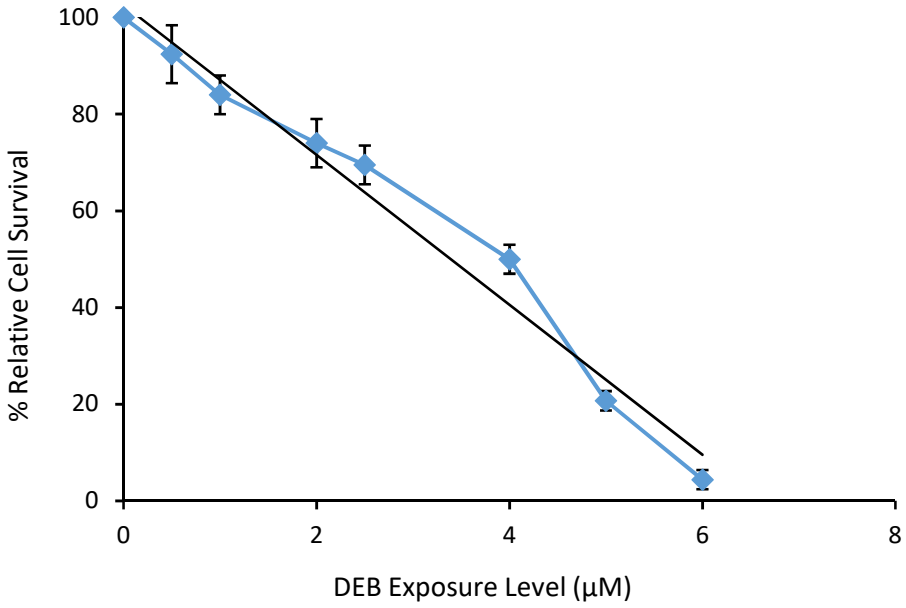


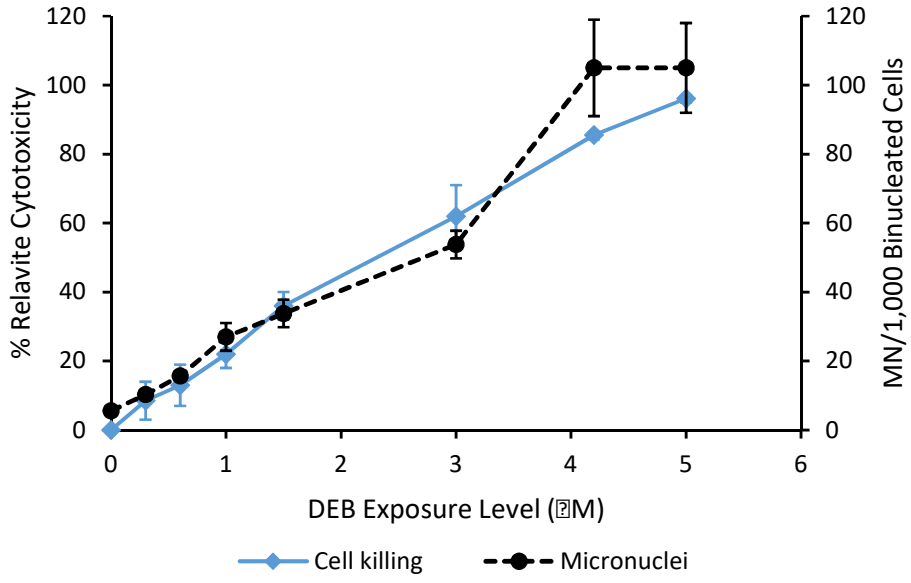
**Figure 2.4** Relationship between dose of DEB and the relative cell survival of HL60 cells. HL60 cells in log phase growth were exposed for 24 h to 0, 0.5, 1, 2, 2.5, 4, 5, or 6 μM DEB, and the relative cell survival and growth of DEB-exposed versus vehicle-exposed cell over time was used to estimate the percent survival at each dose level (n = 3 replicates/dose level/experiment; 12 total control replicates). The best fit to the data from 0 to 6 μM DEB was a linear curve with a strong inverse correlation between DEB exposure level and the percent relative cell survival (R, –0.988, Pearson's Correlation test).



Based upon data presented in Figures 2.3 and 2.4, an additional series of complementary experiments were completed to delineate the dose-response relationships between relative cell survival and formation of MN in HL60 cells exposed for 24 h across a range of DEB concentrations predicted to produce target levels of cytotoxicity that encompass the 20%, 50%, and 80% cell killing. The exposure concentrations selected for defining the dose-response curves for cytotoxicity and formation of MN included 0, 0.3, 0.6, 1.0, 1.5, 3.0, 4.2, and 5.0 μM DEB. These experiments included 0.3 μM DEB as a new lowest exposure level, while exposures to DEB levels greater than 5.0 μM were excluded from this new work because the earlier results showed that 24-h exposures to 6 to 10 μM DEB yielded >95% cell killing.

Following exposures of HL60 cells to concentrations ranging from 0.3 to 5.0 μM DEB, there was a highly significant linear dose-response for reduction in cell survival (via cytostasis and cytotoxicity) with a strong positive correlation between exposure level and the degree of cytotoxicity ($R = 0.9977$, $P = 0.002$; Pearson's Correlation test) (Figure 2.5). The exposure level of 0.6 μM DEB, yielding 13% decrease of cell survival, was the lowest concentration at which the degree of cytostasis and cytotoxicity was significant ($P = 0.05$, Student's *t*-test). The 9% drop of cell survival at the low-dose concentration of 0.3 μM DEB was not significant.

**Figure 2.5** Dose-response relationships between cell killing and formation of micronuclei (MN) in HL60 cells exposed to DEB. HL60 cells in log phase growth were exposed for 24 h to 0-5.0 μM DEB (n = 4 replicates/dose level, except n = 6 to 9/group for one experiment comparing 0, 0.3, and 0.6 μM DEB; 17 total control replicates). Cell killing relative to the survival and growth of vehicle-exposed control cells was calculated by subtracting the relative cell survival at each DEB dose level from 100%. The best fit to the data for relative cell killing was a linear curve with a highly significant positive correlation between DEB exposure level and the percent cell survival (R = 0.998, P = 0.002; Pearson's Correlation test). Likewise, there was a linear dose-response for induction of MN with a significant positive correlation between dose level and the frequency of MN per 1,000 binucleated cells as determined using the cytokinesis-block MN assay (R = 0.984, P = 0.016; Pearson's Correlation test). Error bars, SEM.



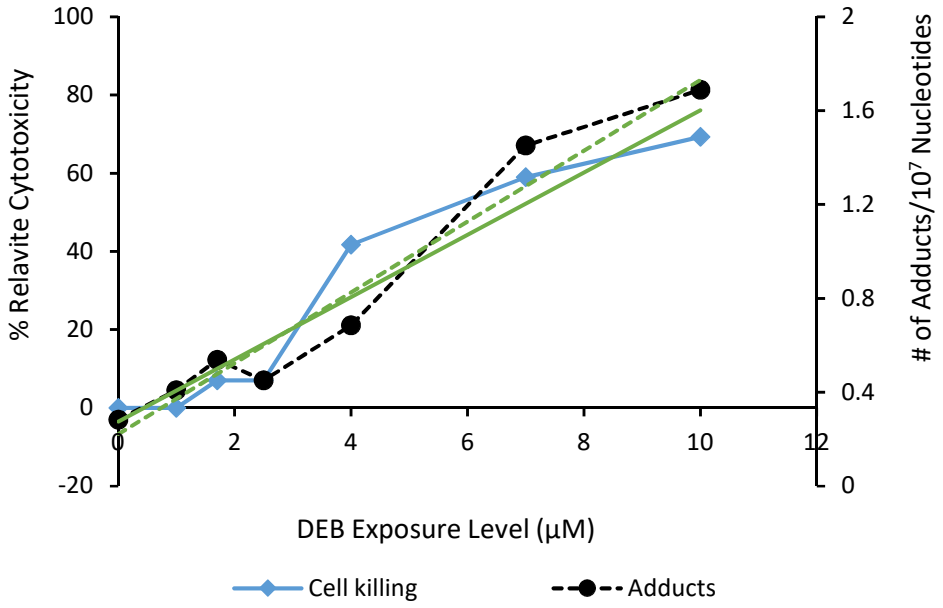
We next examined the effect of DEB on micronuclei formation. Micronuclei are used as a dosimeter of DNA DSBs, as they are formed by genotoxic events and chromosomal instability.[181-182] Following exposures of HL60 cells to concentrations ranging from 0.3 to 5.0 μM DEB, there was a linear dose-response for induction of MN with a significant positive correlation between dose level and the frequency of MN per 1,000 BNCs ($R = 0.9839$, $P = 0.016$; Pearson's Correlation test) (Figure 2.5). Spontaneous background levels of MN ranged from 1 to 10 per 1,000 BNCs across three experiments ($n = 17$ control replicates), averaging 5.6 ± 3.1 MN/1,000 BNCs. The difference between the frequencies of MN in control versus DEB-exposed cells was significant at all exposure concentrations including the low-dose level of 0.3 μM DEB, where the frequency of MN in exposed cells was elevated by 67% (mean = 10.25 MN/1,000 BNCs) over background (mean = 6.11 MN/1,000 BNCs in the relevant experiment) ($n = 9$ replicates/dose level; $P = 0.001$, Student's *t*-test). In HL60 cells exposed for 24 hrs to 4.2 and 5.0 μM DEB, the frequencies of MN were the same (105 MN/1,000 BNCs) suggesting that the formation of measurable MN in viable cells reached a plateau at ~19-fold above background levels of MN at DEB exposure concentrations yielding >85% diminution of cell survival.

### 2.3.8 Relationships between formation of *bis*-N7G-BD DNA-DNA crosslink adducts and cytotoxicity after a 3 hr exposure of HL60 cells to DEB

A single cell culture experiment was performed to characterize the dose-response relationships between formation of *bis*-N7G-BD and relative cytotoxicity in HL60 cells

exposed for 3 h across a range of DEB exposure levels that yielded approximately 20%, 35%, 50%, and 70% or greater cell killing in earlier experiments in HL60 cells exposed for 24h to DEB. For making these comparisons, a shorter DEB exposure of only 3 h was used because both intrastrand and interstrand *bis*-N7G-BD lesions are hydrolytically labile under physiological conditions ($t_{1/2}$ = 3.5 ± 0.3 days).[125] Thus, HL60 cells in log phase growth (15 million cells) were exposed for 3 h to 0, 1.0, 1.7, 2.5, 4, 7, or 10 μM DEB (*n* = 7 replicates/exposure level) and then 10 million cells peer flask were used for measuring *bis*-N7G-BD levels and 5 million cells were used for determining MN formation and relative cell survival. Cell killing relative to the survival and growth of vehicle-exposed control cells was calculated by subtracting the relative cell survival at each DEB exposure level from 100%. Following 3 h DEB exposures of HL60 cells, there was a highly significant linear dose-response for reduction in cell survival, with a strong positive correlation between exposure level and the degree of cytotoxicity ($R$ = 0.961, $P$ = 0.002; Pearson's Correlation test) (Figure 2.6). Likewise, there was a linear dose-response for induction of *bis*-N7G-BD adducts with a significant positive correlation between DEB exposure level and the frequency of adducts (Figure 2.6). Trend lines fit to the data for induction of *bis*-N7G-BD and for cell killing had similar slopes (Figure 2.6), suggesting a direct relationship between the formation of *bis*-N7G-BD cross-links and cytotoxicity of DEB.

**Figure 2.6** Dose-response relationships between cell killing and formation of 1,4-*bis*-(guan-7-yl)-2,3-butanediol (*bis*-N7G-BD) adducts in HL60 cells exposed acutely to DEB. HL60 cells in log phase growth were exposed for 3 h to 0-10 μM DEB (for cell killing analysis, n = 4 replicates/exposure level; for adduct analysis, n = 5 control replicates and 4 replicates per exposure level). Cell killing relative to the survival and growth of control cells was calculated by subtracting the relative cell survival at each DEB exposure level from 100%. The best fit to the data for relative cell killing was a linear curve with a significant positive correlation between DEB exposure level and the % cell survival (R = 0.961, P = 0.002; Pearson's Correlation test). Likewise, there was a linear dose-response for induction of bis-N7G-BD adducts with a significant positive correlation between dose level and the frequency of adducts (R = 0.977, P = 0.03; Pearson's Correlation test).



## 2.4 Discussion

ICL repair is required for cell survival because of the ability of these toxic lesions to block DNA replication and transcription. Current models of ICL repair in mammalian cells support the roles of NER and FA repair pathways.[167] For example, cells derived from Chinese hamsters or humans that are deficient in the FA pathway are hypersensitive to the cytotoxic effects of a variety of DNA cross-linking agents.[162, 166] Mammalian cell clones deficient in NER pathway members, especially *ERCC1* and *XPF*, show an enhanced sensitivity to ICL-forming agents such as cisplatin and cyclophosphamide.[167] Enoiu et al.[148] have shown that both transcription-coupled NER and FA recombinational repair pathways play a role in protecting cells from death induced by cisplatin. However, the relative contributions of the two repair pathways to ICL removal are still under debate.

Cellular sensitivity of DNA repair mutants towards ICL-inducing agents has often been used to judge the relative importance of various repair pathways to ICL removal.[162] However, it should be noted that in addition to ICL lesions, cross-linking agents such as cisplatin and nitrogen mustards induce large numbers of DNA monoadducts, which may be substrates for repair. In fact, DNA monoadducts represent 90–95% of DNA adducts formed upon treatment with *bis*-alkylating agents, while ICL adducts account for 1–3% of the total.[107, 122] Since DNA monoadducts and other types of DNA damage such as DNA-protein cross-links can contribute to the observed toxicity, the results of such cellular viability studies may be ambiguous and difficult to interpret.

Previous work has utilized mass spectrometry to quantify formation of ICL induced by psoralens and ultraviolet A (UVA) irradiation.[183-184] In the present study, we employed a mass spectrometry based methodology developed in our laboratory using a sensitive nanoLC-nanoESI$^+$-MS/MS method on a TSQ Quantiva triple quadrupole mass spectrometer (Thermo Scientific, Waltham, MA) to accurately quantify the formation and repair of *bis*-N7G-BD cross-links in DEB-treated wild-type, NER-deficient, and FA-deficient clones. Due to the fact that the vast majority (~95%) of crosslinks produced by *d,l* DEB are of the interstrand type,[59, 125] this strategy permitted us to gain a new insight into the roles of NER and FA kinetics of ICL formation and removal.

Experiments performed on human cell lines revealed that inactivation of the FA (*FANCD2*) pathway had no significant effect on the kinetics of DEB-induced ICL formation or removal (Figure 2.2). DEB-induced ICL levels in *FANCD2*-deficient PD20 cells and their gene-corrected counterpart cells reached their highest levels at the end of the 3 h treatment with DEB and then declined with linear kinetics over a subsequent three-day period. This result indicates that *FANCD2* deficiency does not prohibit the repair of *bis*-N7G-BD crosslinks.

Parallel experiments performed in *XPA*-deficient and gene-corrected human cells revealed a similar pattern, with a slight, but not statistically significant, increase in ICL levels in *XPA*-deficient cells observed 3 h following DEB removal. Again, as with the PD20 cells, ICL levels declined in both the NER-proficient and NER-deficient human with near-linear kinetics for the subsequent three-day period (Figure 2.2). With this lack of a

difference in the kinetics of DEB-induced ICL formation and removal between wild type, NER-deficient, and FA-deficient clones, it is clear that there is no obvious connection between drug sensitivity and the rate of ICL removal.

We, therefore, propose that the enhanced sensitivity to DEB-induced cell death observed in human cells deficient in the FA pathway is attributable to a defect in their ability to appropriately re-join ICL-induced DSBs. Consistent with this model, following exposure to DEB, chromosomal DSBs and γ-H2AX foci rapidly appear and disappear in the wild-type Chinese hamster cells, suggesting that DSBs formed during ICL removal are rapidly repaired.[68] In contrast, DSBs are much more abundant and persist longer in FA-deficient Chinese hamster cell lines.[68] Given the hypersensitivity of FA cells to ICL-forming agents and their relative insensitivity to X-rays, we propose that these clones suffer from a defect in re-joining DSB repair intermediates created following ICL removal.

Previous observations of elevated levels of ICLs in DEB treated NER-deficient Chinese hamster cells,[68] along with the slightly elevated ICLs observed in the human *XPA*-deficient clone (Figure 2.2), is consistent with the interpretation that NER plays a role in ICL repair.[148] Liu et al. also reported diminished repair of ICLs formed by 8-methoxypsoralen 24 h after treatment in *XPA*-deficient human skin fibroblasts.[185] However, the finding that these cells are only modestly more sensitive to DEB induced cell death than wild-type cells suggests that cells possess an NER-independent mechanism that is capable of removing these ICLs, albeit with diminished efficiency.

It has also been suggested that in the absence of functional FA DNA damage response, DSBs formed following ICL removal are directed to a "default" error-prone NHEJ pathway, which causes chromosomal rearrangements and cell death. Pace et al. reported that inactivation of the NHEJ gene *Ku70* in *FANCC*-deficient chicken DT40 cells partially rescued the latter's sensitivity to cisplatin-induced cell death.[177] Consistent with this finding, Adamo et al.[178] found that inactivation of the NHEJ gene *DNA ligase IV* in clones of *Caenorhabditis elegans* (*C. elegans*) deficient in the *FANCD2* gene restored wild-type sensitivity levels to cisplatin. These authors also observed that inhibition of the NHEJ factor DNA-dependent protein kinase (*DNA-PK*) rescued the mitomycin C hypersensitivity phenotype in human-derived HeLa or MO59K clones expressing *FANCD2* RNAi and partially rescued mitomycin C sensitivity in mouse cells null for either the *FANCA* or *FANCC* genes.[178] In contrast, murine cells lacking both *Ku80* and *FANCD2* genes exhibited greater sensitivity to mitomycin C and cisplatin than did clones deficient in either gene alone.[186] Similarly, Howard et al. reported that inactivation of the *Ku70* gene enhanced the cisplatin sensitivity of *FANCA*-deficient murine cells.[187] In a study by Pace et al.,[177] inactivation of either the *DNA ligase IV* or *DNA PK* genes influenced the cisplatin-sensitivity phenotype of *FANCC* deficient in *C. elegans*. In the present work, *DNA PK* inhibition had no effect on toxicity of DEB in either FA-proficient or FA-deficient mammalian cells (Table 2.1). Taken together, these results lead us to conclude that a complex interaction of both the HR and end-joining repair pathways ultimately determines the fate of DSBs produced during xenobiotic-induced ICL repair.

The experiments involving DEB treatment of HL60 cells were designed to test the postulation that, if BD exposure is causally related to an excess of deaths due to chronic myeloid leukemia (CML), then a genotoxic epoxide metabolite(s) of BD must be capable of producing pathogenic t(9:22) translocations encoding *BCR-ABL* fusion genes that drive CML oncogenesis. Since DEB forms promutagenic *bis*-N7G-BD crosslinks that can induce small and large scale mutations, leading to genotoxic effects two orders of magnitude greater than those produced by the BD monoepoxides, EB and EBD, complementary experiments were performed to define the relationships between DEB exposure and the formation of promutagenic *bis*-N7G-BD adducts, relative cell survival, the induction of MN as a dosimeter of DNA DSBs, and the production of t(9:22) translocations in a promyelocytic leukemia cell line, HL60, lacking the Philadelphia chromosome. In DEB-exposed HL60 cells, there were significant dose-related correlations between formation of *bis*-N7G-BD, levels of cell killing, and induction of MN. In spite of that, targeted doses of DEB yielding ~20%, 50%, and 80% cell killing did not induce increases in t(9:22) translocations over background, while treatments with x-ray as a positive control agent at doses producing similar levels of cytotoxicity caused significant dose-related increases in both MN and t(9:22) translocations (data not shown). These findings appear consistent with the current knowledge of BD metabolism and the patterns in DNA alkylation and genotoxicity of DEB.[59, 71, 116, 125]

While DEB induces dose-related formation of *bis*-N7G-BD crosslink adducts and MN as a quantitative index of DNA DSBs (Figures 2.5 and 2.6), acute exposures of HL60

promyelocytic leukemia cells to DEB at concentrations yielding ~20%, 50%, or 80% cytotoxicity do not appear to produce DNA DSBs at sites associated with the induction of t(9;22) translocations obligatory for CML (data not shown). Collectively, the far greater numbers of MN and disparate patterns in MN formation in x-ray exposed HL60 cells compared to DEB-exposed cells suggest important differences in mechanisms leading to DNA DSBs requisite for the induction of these chromosome lesions as well as t(9;22) translocations.

While the apparent incapacity of *in vitro* DEB exposures to induce t(9;22) translocations in a suitable human target cell is consistent with the failure to find significant associations between BD exposure levels, formation of promutagenic *bis*-N7G-BD adducts, and mutagenic or cytogenetic responses in BD workers,[188-190] additional research is needed to investigate other potential underlying factors that could explain the epidemiological-based association between BD exposures and the risk for CML.[131] The possibility that EB or EBD could cause t(9;22) translocations by direct interaction with DNA appears low because these BD metabolites alone have not been shown to be clastogenic agents. Although chronic low-level BD exposures have not been associated with perturbations in classical markers of mutational and cytogenetic responses,[188-190] it also would be informative to determine if low-level BD exposures over time can trigger endoplasmic reticulum, DNA damage, or other cellular stress responses and the long-term outcomes of inducting those responses.[191-193] Such further studies could help resolve the

conundrum of the conflicting epidemiological and genotoxicity data for BD-exposed humans.

## 2.5 Conclusions

The current study suggests that that the increased sensitivity of FA-deficient hamster cells towards DNA cross-linking agents such as DEB may be due to the accumulation of DNA double strand breaks, which are created upon replication of ICL containing DNA. Additionally, the NER pathway contributes to ICL recognition and removal in mammalian cells, although NER deficiency has a smaller effect on cell viability in the presence of DEB.

Additionally, in vitro exposures of HL60 cells to DEB failed to induce t(9;22) translocations. However, there were significant dose-related correlations between formation of *bis*-N7G-BD, levels of cell killing, and induction of MN. This means that further research is needed to determine the cause of the epidemiological-based association between BD exposure and CML development, as CML is usually caused by t(9:22) translocations.

# III. INTERINDIVIDUAL DIFFERENCES IN DNA ADDUCT FORMATION AND DETOXIFICATION OF 1,3-BUTADIENE DERIVED EPOXIDE IN HUMAN HAPMAP CELL LINES

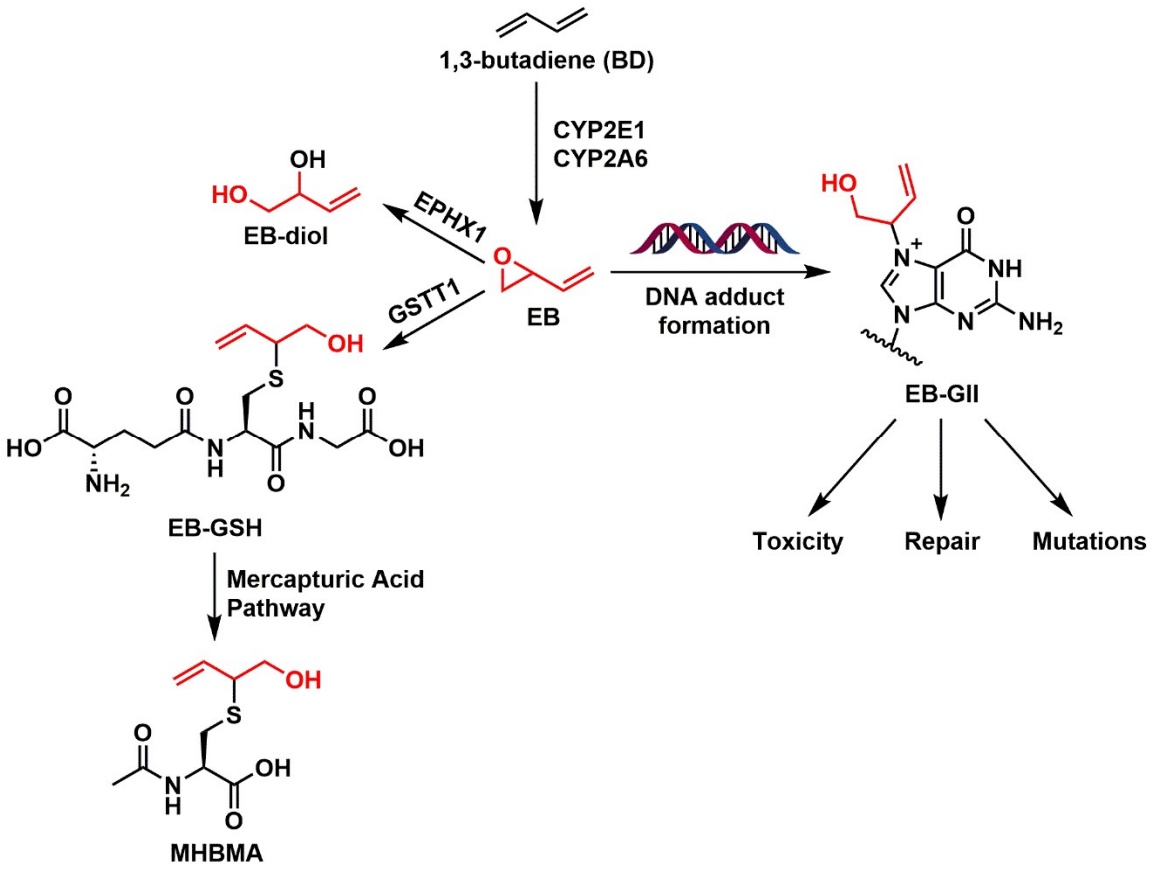
## 3.1 Introduction

Cigarette smoking is the leading cause of lung cancer worldwide, accounting for 81% of lung cancer incidence.[117-118] However, the risk of developing smoking induced lung cancer varies among different ethnic groups. After adjusting for smoking history, African American and Native Hawaiian smokers have the highest risk of developing lung cancer, while Japanese American and Latino smokers have much lower risks, and White smokers have intermediate risk.[194] While the source of these ethnic differences is not known, frequencies of genetic polymorphisms in metabolic enzymes differ between ethnic groups. This can affect the ability of enzymes to metabolize carcinogens, leading to changes in the balance between carcinogen bioactivation and detoxification. This could potentially account for these differences in lung cancer risk.

1,3-butadiene (BD) is a known human carcinogen[91] and is among the most abundant carcinogens in cigarette smoke.[92-94] It has been reported to have a high cancer risk index per cigarette per day as compared to other cigarette smoke components.[119] The genotoxic effects of 1,3-butadiene (BD) are attributed to its epoxide metabolites 3,4-epoxy-1-butene (EB), 3,4-epoxy-1,2-butanediol (EBD), and 1,2,3,4-diepoxybutane (DEB), which are formed through metabolic activation of BD by cytochrome P450 monooxygenases, namely CYP 2E1 and CYP 2A6 (Chapter 1, Scheme 1.8).[100, 102-103] BD-derived epoxides can be detoxified through enzyme catalyzed hydrolysis by epoxide hydrolases (EPHX1) or through glutathione (GSH) conjugation by glutathione S-transferase theta 1 (GSTT1). Glutathione conjugates are further metabolized through the mercapturic acid pathway,

resulting in N-acetylcysteine conjugates (mercapturic acids) that are excreted in the urine (Chapter 1, Scheme 1.9). For EB, the major routes of detoxification are hydrolysis by EPHX1 to 1-butene-3,4-diol (EB-diol) and GSH conjugation to form EB-GSH, which is further metabolized through the mercapturic acid pathway to (2-(*N*-acetyl-L-cystein-S-yl)-1-hydroxybut-3-ene and 1-(*N*-acetyl-L-cystein-S-yl)-1-hydroxybut-3-ene) (MHBMA) (Scheme 3.1).

**Scheme 3.1** Metabolic activation of BD to EB, formation of EB-GII DNA adducts, and detoxification.



If not detoxified, EB, DEB, and EBD can form covalent adduct with DNA, potentially leading to toxicity and mutagenesis (See Scheme 3.1 for EB). EB forms N7-(1-hydroxy-3-buten-2-yl)guanine (EB-GII), while EBD induces N7-(2,3,4-trihydroxybut-1-yl)guanine (N7-THBG). DEB is a bis-electrophile that can sequentially alkylate 2 DNA nucleobases, forming DNA-DNA interstrand crosslinks such as 1,4-*bis*-(guan-7-yl)-2,3-butanediol (*bis*-N7G-BD) (Chapter 1, Figure 1.2). These DNA adducts can lead to stalling of DNA polymerases and misreading during DNA replication, leading to mutations.[11] If present in a tumor suppressor gene or a proto-oncogene, such mutations can lead to unrestricted cell growth and the development of cancer.[12, 195] Indeed, BD derived epoxides are direct genotoxic agents that induce mutations in cells and animals – in studies with mice and rats, BD exposure led to the development of tumors in multiple organs.[96-97] Studies on workers who were occupationally exposed to BD also showed an increased risk of developing leukemia.[98-99]

Quantitation of BD-derived metabolites and DNA adducts can be potentially used as a biomarker of BD exposure and cancer risk. Urinary levels of MHBMA (2-(*N*-acetyl-L-cystein-S-yl)-1-hydroxybut-3-ene and 1-(*N*-acetyl-L-cystein-S-yl)-1-hydroxybut-3-ene), a mercapturic acid ultimately resulting from BD detoxification (Scheme 3.1), are associated with cigarette smoking.[196-197] MHBMA levels were measured in urine samples of White, Japanese American, and African American smokers, where they differed significantly based on ethnicity ($p = 4.0 \times 10^{-25}$).[198] African Americans excreted the highest levels of MHBMA, followed by Whites and Japanese Americans.[198] A genome-wide

association study was conducted to identify genetic polymorphisms that could account for ethnic differences in BD metabolism. Only *GSTT1* copy number was found to have a significant association to MHBMA levels ($p < 0.0001$).[198] Thus, *GSTT1* gene could potentially play a role in BD metabolism, influencing the susceptibility to developing lung cancer in smokers. Sangaraju et al. quantified urinary EB-GII adducts in White and African American smokers and observed significantly higher levels of EB-GII in Whites than in African Americans ($0.048 \pm 0.09$ vs $0.12 \pm 0.02$ pg/mg of creatinine, $p = 3.1 \times 10^{-7}$).[65] Recently, we expanded the multi-ethnic study to smokers and non-smokers from 3 ethnic groups: White, Japanese, and Native Hawaiian and observed the highest levels of EB-GII in Japanese American (5.41; 3.92-7.47 95% confidence limit) fmol EB-GII per mL urine), with lower levels in Whites (2.19; 1.62-2.96 95% confidence limit) fmol EB-GII per mL urine) and Native Hawaiians (2.04; 1.50-2.77 95% confidence limit) fmol EB-GII per mL urine). These differences did not correlate with lung cancer risk or *GSTT1* copy number (Jokipii-Krueger et al., unpublished observations).

The main goal of the current study was to examine inter-individual differences in BD bioactivation and detoxification due to variations in GSTT1 expression/activity. To do this, EB-GSH levels, EB-GII (N7-(1-hydroxy-3-buten-2-yl)guanine) DNA adduct levels, and biological outcomes (apoptosis) were compared after EB exposure in human derived B-lymphocytes from the HapMap project with *GSTT1* either null or present. The International HapMap project aimed to make a haplotype map of the human genome to observe common patterns of human genetic variation.[199] This was accomplished by

extensively genotyping B-lymphocyte cells derived from people from all over the world. These cell lines are available through the Coriell institute, and their extensive genotyping and genetic diversity makes them an ideal cell culture model for investigating ethnic differences.

## 3.2 Materials and methods

**Note:** EB is a known carcinogen and must be handled with adequate safety precautions in a well-ventilated fume hood strictly following its material safety data sheet.

LC-MS grade water and acetonitrile, and HPLC grade methanol were obtained from Fisher Scientific (Pittsburg, PA). $^{15}N_1$,$^{13}C_2$-glutathione was obtained from Toronto Research Chemicals (North York, ON, CA). EB-GII and $^{15}N_5$-EB-GII were synthesized in our laboratory as previously described.[122, 200] All other chemicals and solvents were obtained from Sigma Aldrich (St. Louis, MO).

### *Synthesis of EB-GSH and $^{15}N_1$,$^{13}C_2$-EB-GSH standards*

Isotopically labeled $^{15}N_1$,$^{13}C_2$-EB-GSH internal standard was prepared using a variation of the methodology used by Cho et al. for the enzymatic synthesis of the 1,2,3,4-diepoxybutane (DEB)-GSH conjugate.[201] In brief, 80 µg of recombinant human GSTT1 enzyme (Sigma-Aldrich, St. Louis, MO) was added to 4 µmol of glutathione in 0.1 M potassium phosphate buffer (pH 7.4) (160 µL) and the mixture was incubated in a water bath at 37°C for 15 min. 80 µmol of 3,4-epoxy-1-butene (EB) was added, and the reaction was incubated at 37°C in a water bath for 2 h (total reaction volume: 400 µL). The reaction

was then quenched with 2x volumes of 15% trichloroacetic acid solution. GSTT1 enzyme was removed from the reaction mixture by ultrafiltration using Nanosep 10K filters (Pall Life Sciences, Ann Arbor, MI, USA) centrifuged at 5000xg for 10 min. The filter was washed by adding 100 µL of water to the filter and centrifuging again at 5000xg for 10 min. This procedure was repeated for the synthesis of the isotopically labeled $^{15}N_1$,$^{13}C_2$-EB-GSH internal standard, except the unlabeled GSH was substituted for $^{15}N_1$,$^{13}C_2$-GSH, and the amount of GSTT1 enzyme, GSH, EB, and total reaction volume was scaled down to accommodate the smaller amount of the starting material (1.6 µmol).

Standards were purified by HPLC using an Agilent 1100 series HPLC system (Santa Clara, CA) with a UV-Vis variable wavelength detector. A Luna C18(2) column (250 x 4.6 mm, 5 µm) obtained from Phenomenex (Torrance, CA) was eluted with 0.01% acetic acid in water (solvent A) and 0.01% acetic acid in (95:5) acetonitrile:water (solvent B) at a flow rate of 0.92 mL/min. Solvent composition started at 2% B, and was increased to 10% B in 5 min. The gradient was further changed to 30% B over 5 min and held for 5 min before re-equilibrating the column at 2% B for 10 min. UV absorbance was measured at 215 nm. EB-GSH eluted as a sharp peak at ~10.5 min, while unreacted GSH eluted as a wide peak at ~8 min. Fractions containing EB-GSH were collected and combined and condensed under reduced pressure, resulting in a total yield of 0.4506 µmol of EB-GSH EB-GSH was quantified by $H^1$ NMR by adding a known amount of 3,4,5-trichloropyridine. $^{15}N_1$,$^{13}C_2$-EB-GSH was prepared analogously and quantified using HPLC-ESI-MS/MS by mixing with known amounts of EB-GSH.

***Cell culture and EB exposure***

Human-derived B-lymphocyte cells from the HapMap project were obtained from the NIGMS Human Genetic Cell Repository at the Coriell Institute for Medical Research (Camden, NJ). Based on previous reports on *GSTT1* polymorphisms in HapMap cell lines, eleven human HapMap cell lines were selected, six with *GSTT1* null genotype (GM12874, GM18508, GM18912, GM19128, GM19139, and GM18517) and five with *GSTT1* present (GM19130, GM12145, GM12717, GM19200, and GM12155).[202] Cells were cultured in Gibco RPMI 1640 Medium (Thermo Fisher Scientific, Waltham, MA) supplemented with 15% heat inactivated fetal bovine serum (GIBCO, Thermo Fisher Scientific, Waltham, MA) at 37°C in 5% $CO_2$ humidified atmosphere. All experiments were performed on cells at low passage number (>10).

Cells were counted using the Countess automated cell counter (Thermo Fisher Scientific, Waltham, MA), seeded in T-25 flasks, and grown at 37 °C for 24 h before EB exposure. 2 HapMap cell lines (~5 million cells) were treated 2 mM EB for 0, 2, 4, 8, or 24 hrs before being harvested in a preliminary experiment to determine the ideal EB treatment length. Additionally, 11 cell lines (~3 million cells) in triplicate were treated with 2 mM EB for 6 hrs before being harvested. The cells were then centrifuged at 300xg for 5 min. Supernatant was saved for EB-GSH measurement, while pellet was divided for DNA adduct measurement and cytotoxicity measurement.

In a separate experiment, six additional cell lines: GM12874, GM18912, and GM19128 (GSTT1(-)) and GM19130, GM19200, and GM12145 (GSTT1(+)) were treated

with 10 μM EB using the same procedure described above. Cells were centrifuged at 300xg for 5 min. The cell media was saved for EB-GSH measurement, while the cell pellet was used for DNA adduct measurements.

***Confirmation of GSTT1 transcript expression in HapMap cell lines***

The transcription of *GSTT1* in HapMap cell lines was confirmed at transcript level by quantitative RT PCR. RNA was isolated from the cell lines using RNeasy plus kit (Qiagen, Hilden, Germany) as per manufacturer's protocol. cDNA was synthesized from 1 μg RNA using RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, Waltham, MA) as per manufacturer's protocol. Real time PCR reaction was carried out using Prime PCR assay (Bio-Rad Laboratories, Hercules, CA) for GSTT1 and housekeeping control GAPDH on 250 μg cDNA as template as per manufacturer's protocol in ABI step one plus real time PCR system (Applied Biosystems, Foster City, CA). Relative expression was thus noted using standard fold change method taking GAPDH as housekeeping control and cell line GM12717 as expression control.

***Confirmation of GSTT1 protein expression in HapMap cell lines***

Total cellular protein was extracted from 10 million cells. Cells were pelleted at 300g for 5 min and washed twice with PBS. Cell pellets were re-suspended in 500 μL RIPA buffer (Thermo Fisher Scientific, Waltham, MA) and incubated on ice for 45 minutes, followed by 5 cycles of sonication of 2 mins each on ice. The lysed cells were then centrifuged at 13000xg for 15 min at 4°C and the supernatant, containing total cellular

protein, was collected. Protein concentrations were measured using the Pierce BCA protein assay kit (Thermo Fisher Scientific, Waltham, MA).

Equal amount of proteins (100 μg) were boiled in Laemmli buffer (Bio-Rad Laboratories, Hercules, CA), separated on 10% SDS- polyacrylamide gels, and transferred onto polyvinylidene fluoride (PVDF) membrane (MilliporeSigma, Burlington, MA). Bovine Serum Albumin (BSA), 5%, was used to block the PVDF membrane. This was followed by incubation with primary antibody (Sigma ABS1653 for GSTT1 and Sigma SAB 1404522 for Vinculin) for 3 h and horseradish peroxidase (HRP) conjugated secondary antibody for 45 min at room temperature with intermittent washing 3 times with 0.05 % tween-20 in PBS (v/v) at room temperature for 15 min each. Immunoreactive bands were probed with the enhanced chemiluminescence (ECL) western blot detection system (Bio-Rad Laboratories, Hercules, CA) and viewed in Image studio. The signal intensity of GSTT1 was noted relative to housekeeping control Vinculin.

***Measurement of apoptosis upon EB exposure***

For cytotoxicity (apoptosis) measurements, cell pellets were transferred to fresh media and allowed to recover for 24 h before end point measurement. Annexin V is an apoptotic marker binds to phosphatidylserine that translocates from the inner membrane of the plasma membrane to the outer membrane during apoptosis and thus gives a measure of the percentage of cells actively undergoing apoptosis.[203] Cells exposed to EB (0.5 million) were stained with allophycocyanin (APC) labeled annexin-V and PI as per the manufacturer's guidelines (eBiosciences, San Diego, CA). Population was then analyzed

for percentage of cells in healthy and apoptotic phase on a LSR II 4760 flow cytometer (Becton Dickinson, Franklin Lakes, NJ) using FACSDiva and Flowjo software (University of Minnesota flow-cytometry core).

***Sample preparation for HPLC-ESI$^+$-MS/MS analysis of EB-GSH***

EB-GSH was quantified in cell media using isotope dilution HPLC-ESI-MS/MS. Cell media from EB-treated and control cells (1 mL) was spiked with 126 pmol of $^{15}N_1$,$^{13}C_2$-EB-GSH internal standard. Samples were purified by a two-stage solid phase extraction (SPE) method: by Oasis MCX cartridges (30 mg/mL) followed by Oasis MAX cartridges (30 mg/mL). MCX cartridges were conditioned with 1 mL of 2% formic acid in water, followed by 1 mL of water. Samples were acidified by adding 20 µL of formic acid before being loaded onto the prepared cartridges and washed with 1 mL of water, followed by 2 mL of MeOH. EB-GSH and its internal standard were eluted with 0.5 mL of 30% MeOH in 2% $NH_4OH$, dried under vacuum, and reconstituted in 100 µL of 2% $NH_4OH$ in preparation for the next SPE cartridge. MAX cartridges were conditioned with 1 mL of 2% $NH_4OH$ followed by 1 mL water. Samples were then loaded onto the prepared cartridges and washed with 1 mL water followed by 2 mL MeOH. EB-GSH and its internal standard were eluted with 0.5 mL of 2% formic acid in water, dried under vacuum, and reconstituted in 20 µL of 5 mM ammonium formate (pH 5) in water in preparation for mass spectrometry analysis.

***HPLC-ESI$^+$-MS/MS analysis of EB-GSH***

HPLC-ESI$^+$-MS/MS analysis was conducted using a Dionex LC system interfaced with a TSQ Quantiva instrument (Thermo Fisher Scientific, Waltham, MA, USA). Solvent A was 5 mM ammonium formate (pH 5) in LCMS grade water, and solvent B was LCMS grade acetonitrile. Samples (1 µL) were injected onto an Acquity UPLC HSS T3 column (1 x 100 mm, 1.8 µm, Waters Corp., Milford, MA, USA) and eluted at 20 µL/min. A valve switch was set up after the column to direct the eluent to waste for the first 3 min, and then switching to the MS. Solvent gradient started at 2% B, increased to 10% B in 10 min and further up to 70% B in 1 min. Solvent composition was held at 70% B for 4 min, and the column was then re-equilibrated for 4 min at 2% B. EB-GSH and its internal standard eluted as a sharp peak at ~5.6 min.

A TSQ Quantiva triple quadrupole mass spectrometer (Thermo Scientific, Waltham, MA) was operated in the positive ion mode using Ar as a collision gas (1.5 mTorr). The MS parameters were optimized upon infusion of authentic EB-GSH solution to achieve maximum sensitivity. Quantitative analyses were conducted using selected reaction monitoring (SRM) mode, following the MS/MS transitions corresponding to the loss of glutamate from the analyte (*m/z* 378.10 $[M + H]^+$ → *m/z* 249.05 $[M + H - GLU]^+$) and the loss of both glutamate and water (*m/z* 378.10 $[M + H]^+$ → *m/z* 231.05 $[M + H - GLU - H_2O]^+$). Typical instrument settings included a spray voltage of 2.9 kV, capillary temperature of 400 °C, and collision energy of 10.25 V for both transitions. The corresponding transitions for the $^{15}N_1$,$^{13}C_2$ isotopically labeled internal standard were *m/z*

381.11 $[^{15}N_1,^{13}C_2\text{-M} + H]^+$ → *m/z* 252.06 $[^{15}N_1,^{13}C_2\text{-M} + H - GLU]^+$ and *m/z* 381.11 $[^{15}N_1,^{13}C_2\text{-M} + H]^+$ → *m/z* 234.06 $[^{15}N_1,^{13}C_2\text{-M} + H - GLU - H_2O]^+$. The peak width for both Q1 and Q3 was 0.7 amu. HPLC-$ESI^+$-MS/MS quantitation was based on the areas of the peak for the first transition in the extracted ion chromatograms corresponding to the analyte and the internal standard. Solvent blanks were injected every 4-6 samples to monitor for potential analyte carryover.

***Sample preparation for nanoLC-nanoESI$^+$-MS/MS analysis of EB-GII adducts***

DNA was extracted from treated cells using Qiagen Puregene DNA extraction solution set (Qiagen, Hilden, Germany). Briefly, 1 mL of cell lysis buffer and 3.4 µL of proteinase K solution were added to ~3 million cells. These were incubated overnight at room temperature upon slow invert mixing. The following day, RNA was removed by adding 5 µL of RNAse A solution and incubating samples at room temperature for 3 h. Proteins were then precipitated out by adding 0.5 mL protein precipitation solution, vortexing samples for ~20 sec, and centrifuging at 2000g for 15 min. DNA was then precipitated by adding 34 µL of 5 M $NH_4OAc$ and 1.5 mL ice-cold IPA to each sample and storing overnight in a -20°C freezer. Extracted DNA was then washed twice with 1 mL of 70% EtOH in water and sheered using 19 and 22 gauge needles.

DNA concentrations were estimated using a nanodrop UV spectrophotometer (Thermo Scientific, Waltham, MA, USA) based on the absorbance at 260 nm. DNA purity was assessed from $A_{260}/A_{280}$ absorbance ratios, which were between 1.8 and 1.9. Accurate

DNA amounts were confirmed by dG quantitation after enzymatic hydrolysis as described previously.[172]

DNA (10-25 μg in water) was used for quantitation of EB-GII DNA adducts, as described previously with some changes.[65, 67] Briefly, samples were spiked with 50 fmol $^{15}N_5$-EB-GII internal standard and heated at 70°C for 1 h to release EB-GII adducts from the DNA backbone. The DNA backbone was then removed by ultrafiltration with Nanosep 10K filters (Pall Life Sciences, Ann Arbor, MI, USA) at 5000xg for 10 min. The filtrate was then subjected to SPE using Strata-X polymeric reverse phase cartridges (30 mg/mL). Cartridges were conditioned with 2 mL of HPLC grade methanol, followed by 2 mL of water. Samples were then loaded onto the cartridges. They were then washed with 1 mL of water followed by 1 mL of 10% methanol in water. Finally, EB-GII and its internal standard were eluted in 1 mL of 60% methanol in water. Samples were dried under vacuum, reconstituted in 200 μL of water, and filtered using Nanosep 10K filters at 5000g for 10 min in order to remove any SPE particles. Filtered samples were then dried under vacuum and reconstituted in 15 μL of 0.01% acetic acid in water in preparation for nanoLC-nanoESI$^+$-MS/MS analysis.

***NanoLC-nanoESI$^+$-MS/MS analysis of EB-GII adducts***

The method used for nanoLC-nanoESI$^+$-MS/MS analysis of EB-GII adducts was based on our previously published method,[67] but was adapted for a TSQ Quantiva triple quadrupole MS. NanoLC-nanoESI$^+$-MS/MS analysis was conducted using a Dionex LC

system with a 5 µL loop interfaced with a TSQ Quantiva instrument (Thermo Fisher Scientific, Waltham, MA, USA). Solvent A was 0.01% acetic acid in LCMS grade water and solvent B was 0.02% acetic acid in LCMS grade acetonitrile. Samples (1 µL) were injected onto a nano-LC column with a fused-silica tipped emitter (New Objective, Woburn, MA) that was manually packed with Synergi Hydro-RP, 80Å, 4 µm chromatographic packing (Phenomenex, Torrance, CA). The solvent gradient started at 2% B at 800 nL/min, which was held for 7.5 min to allow the sample to reach the nano-LC column. The flow rate was then reduced to 300 nL/min and maintained at 2% B for 0.5 min. The gradient was then increased linearly to 25% B in 9 min and then up to 50% B in 10 min. The column was then re-equilibrated at 800 nL/min at 2% B for 7 min. EB-GII and its internal standard eluted as a sharp peak at ~16.5 min.

The triple quadrupole mass spectrometer was operated in the positive ESI mode using Ar as a collision gas (1.5 mTorr). The mass spectrometer was tuned upon infusion of authentic EB-GII solution to achieve maximum sensitivity. Quantitative analyses were conducted using the selected reaction monitoring (SRM) mode by fragmenting $[M + H]^+$ ions of EB-GII (*m/z* 222.155) to $[Gua + H]^+$ ions (*m/z* 151.84) via collision induced dissociation (CID) in Q2 using a collision energy of 13.8 V. Typical instrument settings included a spray voltage of 2.5 kV, capillary temperature of 375°C, RF lens of 64 V, and a dwell time of 100 ms. The peak width for both Q1 was 0.4 amu and Q3 was 0.7 amu. Eb-GII was quantified using the extracted ion chromatogram for *m/z* 222.155 $[M + H]^+ \rightarrow$ *m/z* 151.84 $[Gua + H]^+$. The corresponding transitions for the $^{15}N_5$ isotopically labeled internal

standard were *m/z* 227.14 [$^{15}N_5$-M + H]$^+$ → *m/z* 156.825 [$^{15}N_5$-Gua + H]$^+$. NanoLC-nanoESI$^+$-MS/MS quantitation was done by comparing the areas of the peaks for the analyte and the internal standard. Solvent blanks were injected every 4-6 samples to monitor for potential analyte carryover.

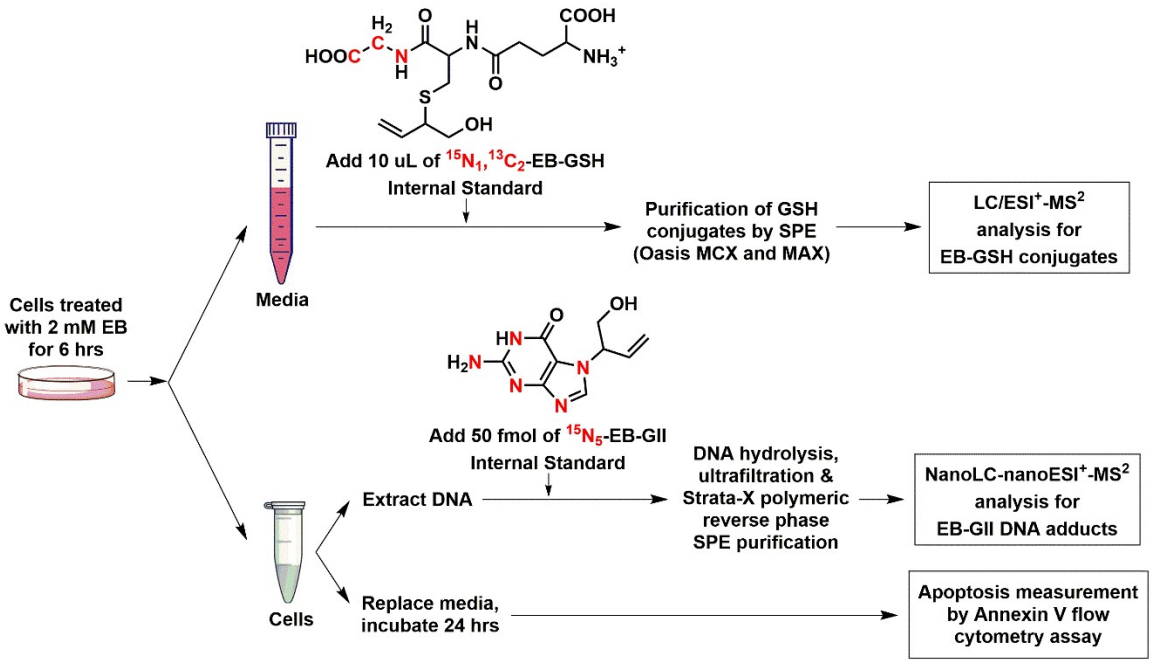
## 3.3 Results

### 3.3.1 Experimental Approach

The present study was designed to elucidate the effects of GSTT1 expression in human cells on cellular responses to treatment with the active metabolite of BD, 3,4-epoxy-1-butene (EB, Scheme 3.1). Our ultimate goal was to investigate the role of GSTT1 gene product in cellular protection against EB toxicity and carcinogenesis. For this reason, we compared the extent of GSH conjugation, DNA adduct formation, and apoptosis in a range of human HapMap cells with different levels of GSTT1 gene expression following exposure to EB (Scheme 3.2). As discussed above, human cell lines from the HapMap project were derived from human populations from all over the world and have been extensively genotyped, allowing for cell culture studies that mimic the genetic diversity of real human populations.[199]

For our study, we selected 11 cell lines from the HapMap project that either did or did not express GSTT1 protein. Flow cytometry assay using Annexin V marker was used to determine the extent of cell death via apoptosis.[203] The levels of EB-GSH conjugates were quantified to examine the role of this GST isoform in detoxification of EB. EB-GII (Scheme 3.1) was chosen as a representative EB-DNA adduct. EB-GII, along with its structural isomer EB-GI, is the most prevalent DNA adduct resulting from exposure to EB.[107]

**Scheme 3.2** Experimental procedure for the analysis of EB-GSH conjugates, EB-GII adducts, and apoptosis in cells exposed to EB.

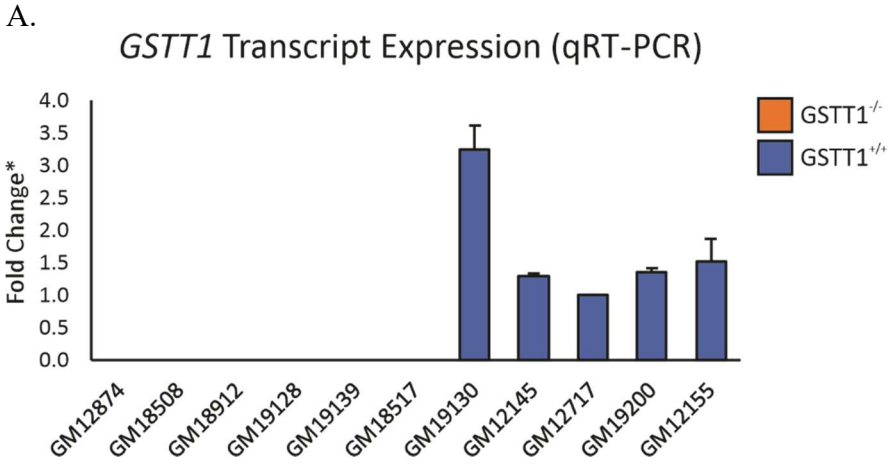


### 3.3.2 Confirmation of GSTT1 expression at transcript and protein level

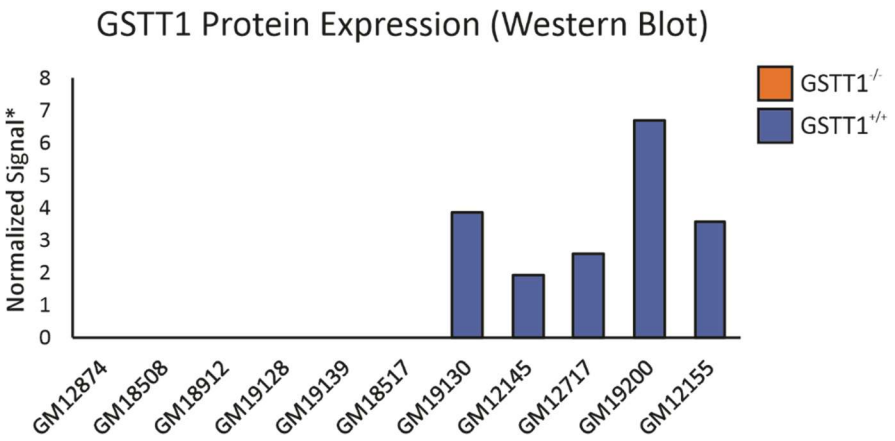
GSTT1 expression levels in 11 HapMap cell lines were confirmed by RT-PCR, and GSTT1 protein expression was confirmed by western blot. GSTT1 expression levels correlated with the published genotype[202] at both transcript and protein level in all 11 cell lines. Cell lines with GSTT1 deletion contained no GSTT1 transcripts or the corresponding protein, whereas the cell lines with at least one copy of the gene had detectable levels of transcript and protein (Figure 3.1).

**Figure 3.1** GSTT1 expression in HapMap cells. A. Expression of *GSTT1* transcript in selected HapMap cell lines, confirmed by qualitative RT PCR. *Fold change noted with respect to cell line GM12717. B. Expression of GSTT1 protein in selected HapMap cell lines, confirmed by western blot. *Signal intensity normalized to vinculin signal.

A.




B.

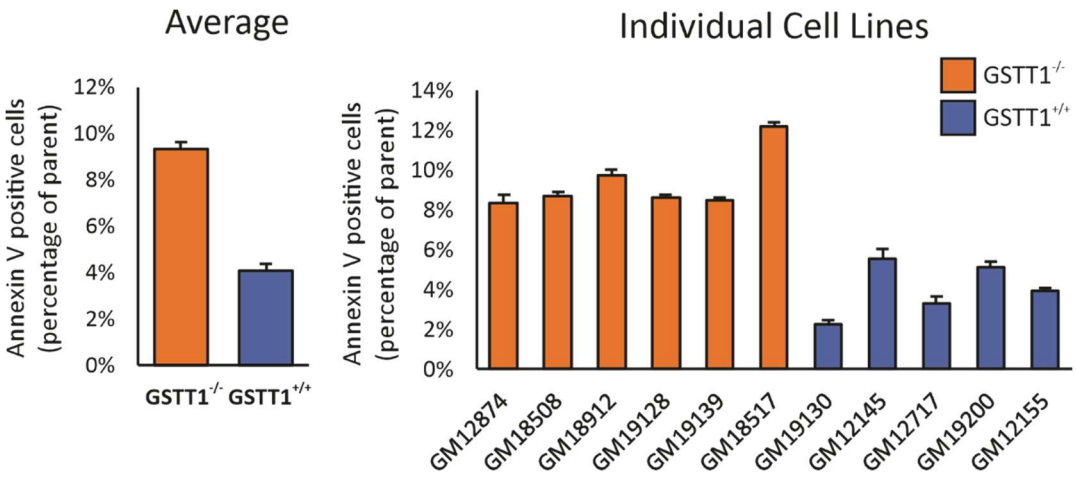


### 3.3.3 Apoptosis in cells treated with EB

In order to study the biological effects of EB treatment, a flow cytometry assay was devised. Annexin V is a cell surface marker for apoptosis, and thus by measuring the levels of this marker on the outer membrane gives the percentage of cells currently undergoing apoptosis.[203] Following EB treatment, cells were allowed to recover for 24 h before the flow cytometry assay was performed. Preliminary concentration dependence studies revealed that no apoptosis was observed in HapMap cells treated with EB concentrations lower than 1 mM.

To establish the effects of GSTT1 expression on toxicity of EB, HapMap cells (3 million) were treated with 2 mM EB for 6 h, and the fraction of cells undergoing apoptosis was established by flow cytometry. We found that the percentage of cells in apoptotic phase after EB exposure was higher in *GSTT1*(-) cell lines (9.35% ± 0.26%) as compared to those with at least one copy of *GSTT1* (4.05% ± 0.317%) (Figure 3.2).

**Figure 3.2** Apoptosis levels after exposure to 2 mM EB for 6 hrs. 3 million cells, in triplicate, were incubated for 24 h before being exposed to 2 mM EB for 6 hrs. Cells were then further incubated with fresh media for 24 hrs. Apoptosis was measured using Annexin V marker expression on cell surface.



### 3.3.4 Synthesis of EB-GSH and $^{15}N_1$,$^{13}C_2$-EB-GSH standards

EB-GSH was synthesized using recombinant human GSTT1 to conjugate GSH to EB. The resulting EB-GSH was isolated by HPLC and quantified by proton NMR using 3,4,5-trichloropyridine as an internal standard (Appendix A, Figure A1).

$^{15}N_1$,$^{13}C_2$-EB-GSH was synthesized analogously using $^{15}N_1$,$^{13}C_2$-GSH. Both standard and internal standard were purified by HPLC, and their identity was established by proton NMR and LC-MS. A stock solution of $^{15}N_1$,$^{13}C_2$-EB-GSH (25.2 pmol/µL in water) was stored at -20ºC and used as an internal standard for all HPLC-ESI$^+$-MS/MS analyses of EB-GSH.

### 3.3.5 HPLC-ESI$^+$-MS/MS method development for EB-GSH in cellular media

In order to measure the amounts of EB that undergoes conjugation with GSH in human cells, an LC-MS method for EB-GSH was developed using authentic standards prepared in our laboratory. In animals and humans that have been exposed to EB, EB-GSH is further metabolized through the mercapturic acid pathway to form MHBMA, which was employed in our previous studies.[198, 204-205] However, the enzymes responsible for further metabolism of GSH conjugates to mercapturic acids are mostly found in the liver and kidney tissues, and not in B-lymphocytes like the HapMap cell lines.[206] Thus, EB-GSH is less likely to be further metabolized in this cell line model. Therefore, an isotope dilution LC-MS/MS method for EB-GSH was developed.

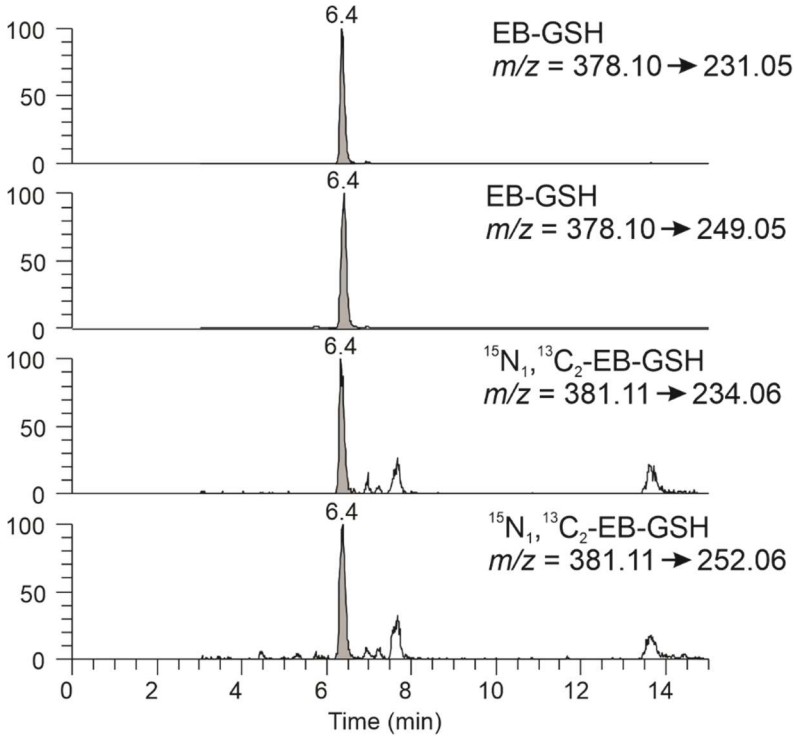
Early experiments showed that EB-GSH was rapidly excreted from HapMap cells into the media. In order to analyze these samples by LC-MS/MS, a clean-up procedure to get rid of any salts, proteins, and other interfering compounds from the media needed to be used. As glutathione can hold either a positive charge on the amine under acidic conditions or a negative charge on the carboxylic acid under basic conditions, a two-step SPE method was devised. Media samples were first acidified and subjected to mixed-mode strong cation exchange SPE (Oasis MCX cartridges, 30 mg/mL, Waters Corporation, Milford, MA) to get rid of any salts, neutrals, and negatively charged compounds. The resulting eluent was dried down, reconstituted in a basic solution, and subjected to mixed-mode strong anion exchange SPE (Oasis MAX cartridges, 30 mg/mL, Waters Corporation, Milford, MA), to remove any contaminants that cannot hold a negative charge. The resulting eluent containing EB-GSH and its internal standard was clean enough for analyses by capillary HPLC-ESI-MS/MS. A valve switch was added to the beginning of the LC-MS run, so that the first 2 min of the gradient would be sent to waste instead of the MS, as an added precaution to keep the MS system clean.

EB-GSH is a very polar molecule and retains poorly on most reverse phase HPLC columns. Several types of stationary phases were tested, including Zorbax SB-C18 (Agilent Technologies, Santa Clara, CA), Luna C18(2) and Synergi Hydro RP (Phenomenex, Torrance, CA), and Acquity UPLC HSS T3 (Waters Corporation, Milford, MA). Only the Acquity UPLC HSS T3 retained EB-GSH more than 3 min. Next, several different mobile phases and gradients were tested, and a gradient of 5 mM ammonium formate in water (pH

5) and acetonitrile afforded the best chromatography and ionization efficiency. With the final optimized chromatography, EB-GSH elutes at ~6.4 min (Figure 3.3).

Two SRM transitions for HPLC-ESI-MS/MS analysis of EB-GSH were selected corresponding to the loss of glutamate from the analyte (*m/z* 378.10 → 252.05) and a loss of both glutamate and water from the analyte (*m/z* 378.10 → 234.05). The first transition, loss of glutamate, showed higher reproducibility and sensitivity, and was used for quantitation. The second transition, loss of glutamate and water, was used for confirmation purposes. Similar transitions with added +3 mass units were used for $^{15}N_1$,$^{13}C_2$-EB-GSH (internal standard for quantification).

**Figure 3.3** HPLC-ESI$^+$-MS/MS analysis of EB-GSH in cellular media from GM12874 cell line treated with 2 mM EB for 6 h.



### 3.3.6 Quantification of EB-GSH in cellular media of HapMap cells treated with 2 mM EB

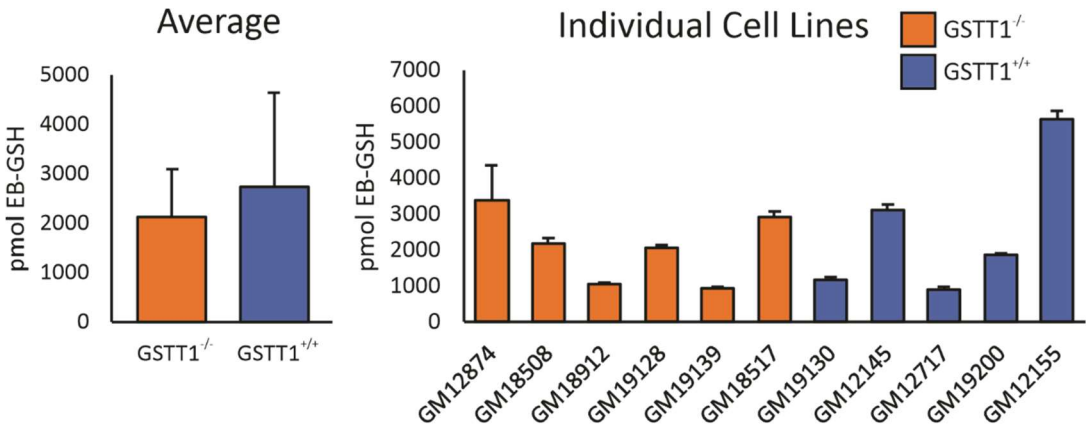
To examine the effects of *GSTT1* genotype on EB-GSH conjugate formation after EB exposure, 6 *GSTT1*$^{-/-}$ HapMap cell lines (GM12874, GM18508, GM18912, GM19128, GM19139, and GM18517) and 5 *GSTT1*$^{+/+}$ HapMap cell lines (GM19130, GM12145, GM12717, GM19200, and GM12155) were exposed to 2 mM EB for 6 h. 1 mL of media from each cell line was removed and spiked with the $^{15}N_1$,$^{13}C_2$-EB-GSH internal standard (126 pmol). The resulting samples were processed by two step SPE and analyzed by LC-ESI$^+$-MS/MS methods as described above. EB-GSH amounts in treated cells varied between 1.04 and 5.85 nmol EB-GSH per mL media. No EB-GSH was observed in control cells. While differences in EB-GSH levels between cell lines were seen, these differences did not correlate with GSTT1 expression, suggesting that in the absence of GSTT1, GSH conjugation of EB can be catalyzed by another glutathione-S transferase (Figure 3.4 A).

Potentially different results were observed in a subset of 6 cell lines (3 *GSTT1*$^{-/-}$, 3 *GSTT1*$^{+/+}$) treated with 10 μM EB for 6 h. EB-GSH adduct levels in cells treated with 10 μM EB were approximately 150-fold lower than in cells treated with 2 mM EB (15.74 ± 10.81 vs 2396 ± 1457 pmol EB-GSH per mL media). Levels of EB-GSH were higher in *GSTT1*$^{+/+}$ (22.95 ± 10.66 pmol EB-GSH per mL media) than in *GSTT1*$^{-/-}$ (8.53 ± 4.78 pmol EB-GSH per mL media) cell lines, however, this difference was not statistically significant ($P > 0.05$, Figure 3.4 B). The lack of statistical significance between *GSTT1*$^{-/-}$ and *GSTT1*$^{+/+}$ cell lines could be due to the low number of samples analyzed – only a single EB treated

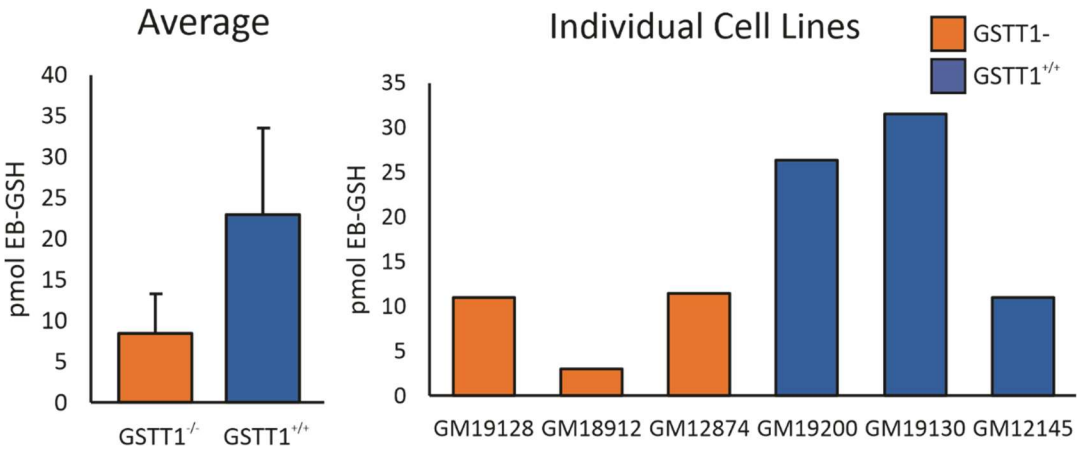
sample from each of the 6 cell lines were analyzed in time to be reported here. It is possible that with the analysis of more samples, statistical significance will be observed.

**Figure 3.4** (A) Quantitation of EB-GSH conjugates in *GSTT1*$^{-/-}$ (GM12874, GM18508, GM18912, GM19128, GM19139, and GM18517) and $^{+/+}$ cell lines (GM19130, GM12145, GM12717, GM19200, and GM12155). 3 million cells from 11 cell lines, in triplicate, were treated with 2 mM EB for 6 h. (B) 5 million cells from 3 *GSTT1*$^{-/-}$ (GM19128. GM18912, and GM12874) and 3 $^{+/+}$ cell lines (GM19200, GM19130, and GM12145) were treated with 10 μM EB for 6 h. EB-GSH levels in the media were quantified by HPLC-ESI-MS/MS from 1 mL aliquots of media from treated cells.

A. 2 mM EB




B. 10 μM EB



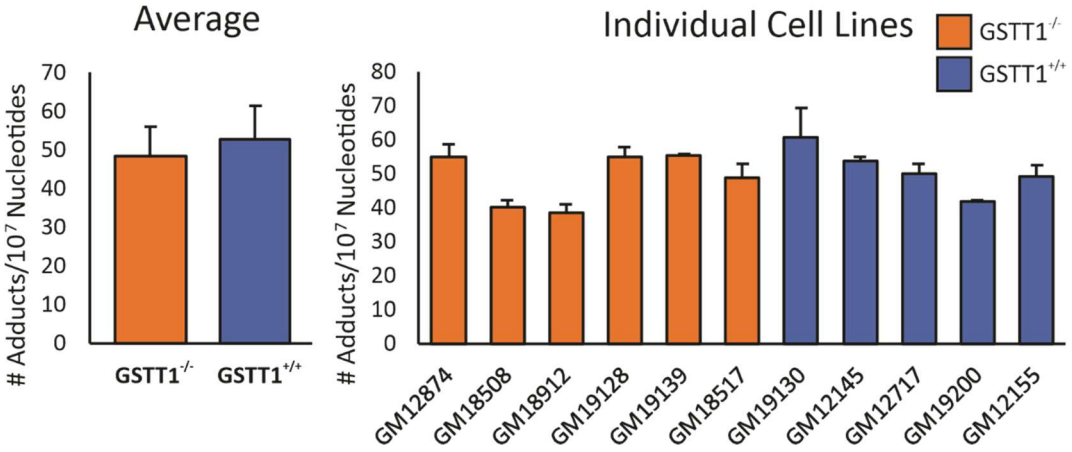
### 3.3.7 Quantification of EB-GII in cells treated with 2 mM and 10 μM EB

To examine the effects of *GSTT1* genotype on DNA adduct formation after exposure to EB, 6 *GSTT1*$^{-/-}$ HapMap cell lines (GM12874, GM18508, GM18912, GM19128, GM19139, and GM18517) and 5 *GSTT1*$^{+/+}$ HapMap cell lines (GM19130, GM12145, GM12717, GM19200, and GM12155) were exposed to 2 mM EB for 6 h. DNA was extracted and analyzed by nanoLC-nanoESI$^{+}$-MS/MS methods as described above. No EB-GII adducts were observed in control cells (not shown). Adduct levels varied between 35.0 and 70.9 adducts per $10^7$ nucleotides depending on cell line. However, differences in EB-GII adduct formation in *GSTT1*$^{-/-}$ cell lines (48.3 ± 7.6 adducts per $10^7$ nucleotides) and *GSTT1*$^{+/+}$ cell lines (52.7 ± 8.6 adducts per $10^7$ nucleotides) were not statistically significant (Figure 3.5 A).

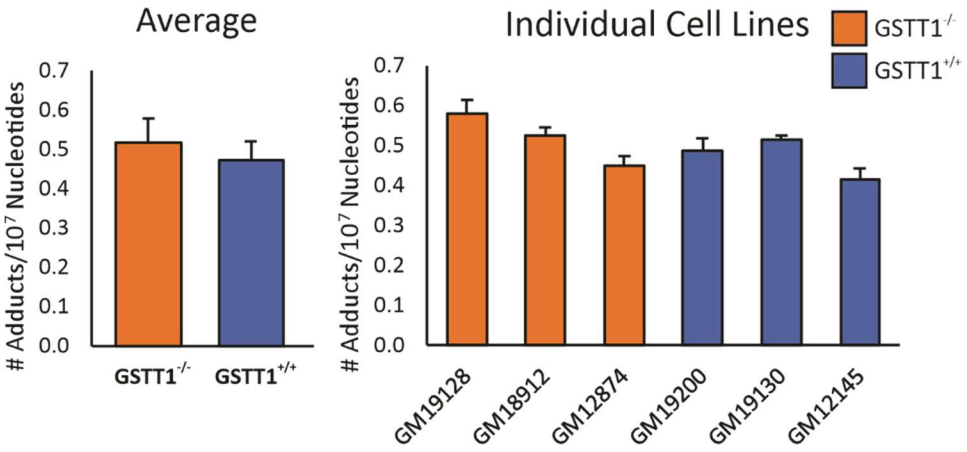
Similar results were observed in a subset of 6 cell lines (3 *GSTT1*$^{-/-}$, 3 *GSTT1*$^{+/+}$) treated with 10 μM EB for 6 h. EB-GII adduct levels in cells treated with 10 μM EB were approximately 100-fold lower than in cells treated with 2 mM EB (0.049 ± 0.0059 vs 50.5 ± 8.3 adducts per $10^7$ nucleotides), but still showed no statistically significant difference between *GSTT1*$^{-/-}$ (5.2 ± 0.62 adducts per $10^8$ nucleotides) and *GSTT1*$^{+/+}$ (4.7 ± 0.49 adducts per $10^8$ nucleotides) cell lines (figure 3.5 B).

**Figure 3.5** Quantitation of EB-GII adducts in *GSTT1*$^{-/-}$ and $^{+/+}$ cell lines. 3 million cells, in triplicate, were treated with 2 mM EB (A) and 10 µM EB (B) for 6 h. DNA was extracted and EB-GII adduct levels were quantified by MS.

A. 2 mM EB




B. 10 µM EB



### 3.4 Discussion

Glutathione S-transferase theta 1 (GSTT1) in a key enzyme in detoxification pathway of BD-derived epoxides. Previous studies have shown that urinary levels of MHBMA, the ultimate product of EB-GSH conjugation excreted in urine, was associated with cigarette smoking.[196-197] Furthermore, significant differences in EB-GSH levels were seen between White, Japanese American, and African American smokers.[198] A genome-wide association study indicated that *GSTT1* copy number was able to partially account for the differences in MHBMA levels between smokers of these ethnicities.[198]

In the present study, we investigated the effects of GSTT1 expression on detoxification of EB, EB-DNA adduct formation, and apoptotic cell death after exposure to EB. Cell lines investigated were chosen from the HapMap project, as they more closely represent the diversity of real human populations than a monoclonal cell line and have been extensively genotyped. 6 cell lines that were null for GSTT1 expression and 5 cell lines that were positive for GSTT1 expression were chosen for investigation.

As expected, *GSTT1*$^{-/-}$ cells exhibited higher sensitivity towards EB-induced apoptosis as determined by Annexin V flow cytometry assay (Figure 3.2). These results suggest that GSTT1 has a protective effect against EB toxicity. In order to quantify the amount of EB-GSH conjugate formed after EB exposure, we developed a sensitive HPLC-ESI-MS/MS methodology for measuring EB-GSH in cellular media. Since media is a highly complex biological matrix, a two-stage SPE method using both mixed-mode strong

cation exchange and mixed-mode strong anion exchange stationary phases was developed to isolate GSH conjugates prior to LC-MS analysis. An isotope dilution HPLC-ESI$^+$-MS/MS method for detecting and quantifying EB-GSH was developed on our TSQ Quantiva triple quadrupole system. EB-GSH was readily detectable in cellular media, with EB-GSH signals being well above the background (Figure 3.2). While a few contaminants were seen in the extracted ion chromatograms, they were all baseline resolved from EB-GSH and present at much lower levels, and thus did not interfere with EB-GSH quantitation. Their presence is likely due to the inherent complexity of the cellular media matrix and the limitations of the current clean-up method.

After exposure to 2 mM EB, the amount of EB-GSH in the cellular media was quantified by HPLC-ESI$^+$-MS/MS. However, similar EB-GSH amounts were found in cells that expressed GSTT1 and in *GSTT1*$^{-/-}$ cells (Figure 3.3). Furthermore, nanoLC-nanoESI$^+$-MS/MS analysis of EB-GII DNA adduct levels after 2 mM EB exposure also did not show any difference between *GSTT1*$^{-/-}$ and *GSTT1*$^{+/+}$ cell lines (Figure 3.3 A).

The high level of EB that cells were exposed to (2 mM) was chosen to be consistent with the apoptosis assay, as this concentration of EB was required to observe Annexin V signal above the background noise. This concentration of EB is much higher than the exposure concentrations used in previous cell culture studies, which is typically in the μM range.[67, 207] While enzymatic conjugation of GSH is faster, spontaneous conjugation of GSH to EB can also occur. With the high concentration of EB used for exposures, it is

possible that any GSTT1 present in these cells was saturated and non-enzymatic GSH conjugation of EB was the primary driver of EB-GSH formation.

To test whether a difference in adduct and conjugate levels would be observed at a lower EB exposure level, a subset of 6 cell lines (3 *GSTT1*$^{-/-}$ and 3 *GSTT1*$^{+/+}$) were exposed to 10 μM EB for 6 hrs. However, despite EB-GII adduct levels being approximately 100-fold lower than in 2 mM EB treated cells, once again, no difference in adduct levels was seen between *GSTT1*$^{-/-}$ and *GSTT1*$^{+/+}$ cells (figure 3.3 B). EB-GSH conjugate levels have not yet been analyzed in these samples, but likely reflect the trend seen previously. It is possible that even at 10 μM EB, GSTT1 enzymes have been fully saturated and non-enzymatic GSH conjugation is still the primary driver of EB-GSH conjugation. The cell lines used for this study were all human derived B lymphocytes, whereas hepatocytes are the primary source of GSH conjugation.[206] While the expression of GSTT1 was confirmed in all 5 *GSTT1*$^{+/+}$ cell lines, the expression levels were still very low. This inherently low level of GSTT1 enzyme may cause enzymes to be saturated unless they are exposed to even lower amounts of EB, with non-enzymatic GSH conjugation contributing to the majority of EB-GSH formation. Further studies should be conducted using more metabolically active cell lines such as hepatocytes to determine whether the lack of effect of GSTT1 expression seen in this study on EB-GSH and EB-GII formation is due to protein expression levels or our original hypothesis being incorrect.

Another possible explanation for the lack of change in conjugate and adduct formation is the ability of other GST enzymes to mediate the conjugation of GSH to EB.

As the role of GSTs in BD detoxification has not been extensively studied, it is possible that other GSTs could contribute to EB-GSH conjugation. Glutathione S-transferases are a superfamily of proteins that are involved in the detoxification of xenobiotic substrates. The cytosolic family of GSTs, one of the 3 families that make up this superfamily, contains 7 classes of GSTs, including GST theta and mu.[208] While previous studies found that GSTT1 accounted for the majority of BD detoxification, GST mu 1 (GSTM1) and GST theta 2 (GSTT2) may also be involved.[209-210] Boldry et al. found that both GSTT1 and GSTT2 can catalyze EB-GSH conjugation, with GSTT1 exhibiting a faster rate of conjugation.[198] Additionally, MHBMA levels in Whites, African Americans, and Japanese Americans were found to be associated with *GSTM1* genotype, suggesting that GSTM1 may also play a role in BD metabolism.[198] While the expression of GSTT1 in the cell lines investigated is known, the expression of GSTM1 and GSTT2, among other GSTs, is not. It is possible that in the absence of GSTT1, other GSTs may fill its role and contribute to EB-GSH conjugation. GSTs also play a role in much more than just EB detoxification, and the source of the biological effect of the lack of GSTT1, seen in the higher percentage of apoptosis in *GSTT1*$^{-/-}$ cell lines, could be less direct than the conjugation of EB-GSH.

Another possibility is that EB is not a major contributor to the development of smoking-induced cancer, and thus would not play a large role in the difference in cancer risk between different ethnicities. EB is not nearly as cytotoxic or mutagenic as 1,2,3,4-diepoxybutane (DEB), another reactive epoxide metabolite of BD. DEB is thought to be the major contributor to BD carcinogenesis, as it is a bis-electrophile that can form

extremely toxic DNA-DNA crosslinks.[103-104] A parallel study to this one, focusing on the effects of GSTT1 expression on the formation of DEB-GSH conjugates and *bis*-N7G-BD DNA adducts, is currently ongoing.

In summary, GSTT1 appears to protect HapMap cells against EB-mediated apoptosis. In contrast, GSTT1 expression level had no effect on the formation of EB-GSH conjugates and EB-GII DNA adducts, suggesting that another GST enzyme can catalyze the same reaction in the absence of GSTT1. Therefore, GSTT1 is not essential for EB detoxification following exposure to 1,3-butadiene. Furthermore, another biological mechanism not related to its role in carcinogen metabolism must be responsible for the observed protective effects of GSTT1 in human HapMap cells.

# IV. DISCOVERY OF NOVEL 4-HYDROXYBENZYL-VALINE HEMOGLOBIN ADDUCTS IN HUMAN BLOOD

Reprinted with permission from:

Degner, A.; Carlsson, H.; Karlsson, I.; Eriksson, J.; Pujari, S. S.; Tretyakova, N. Y.; and Törnqvist, M.; Discovery of novel 4-hydroxybenzyl-valine hemoglobin adducts in human blood. *Chem. Res. Tox.* 2018, 31(12) 1305-1314. 

This work was performed in collaboration with Dr. Henrik Carlsson, Dr. Isabella Karlsson, Johan Eriksson, and Dr. Suresh Pujari, under the direction of Dr. Natalia Tretyakova and Dr. Margareta Törnqvist. 4-QM precursor was synthesized by Suresh Pujari. Benzaldehyde compounds were synthesized and purified by Johan Eriksson. Compound identification spectra including NMR and high-resolution mass spectra were collected by Amanda Degner, Suresh Pujari, and Isabella Karlsson. Adductomics screening, LCMS method development, hypothesis of adduct structure, and synthesis, derivatization, and LCMS analysis of adduct derived from 2-QM were done by Henrik Carlsson. Synthesis, derivatization, and LCMS analysis of adduct derived from 4-QM was done by Amanda Degner with help from Isabella Karlsson. Derivatization, LCMS analysis. and quantitation of benzaldehyde derived valine adducts was done by Isabella Karlsson. Amanda Degner wrote the manuscript.

## 4.1 Introduction

DNA and protein adducts are formed when endogenous and exogenous electrophilic compounds react with nucleophilic sites of cellular biomolecules.[4] Levels of formed adducts increase with the dose of electrophiles *in vivo*, which is associated with an increased risk of health effects, as has been demonstrated for genotoxic compounds/metabolites in animal cancer tests.[211-212] The formation of covalent adducts, especially with DNA, is associated with an increased risk of cancer[5-7] and other chronic diseases. As such, there is increasing evidence that endogenous and exogenous exposures, which also includes electrophilic compounds/intermediates, are equally important to disease development as genetic factors.[76-77, 213] Since many electrophilic compounds that form adducts with nucleophilic sites in DNA can also form adducts with hemoglobin (Hb), which is highly abundant and more available, electrophile-induced Hb adducts have been used as biomarkers for electrophile-induced DNA adduct formation.[4, 47-48]

In previous studies, sensitive mass spectrometry (MS)-based targeted screening methods have been developed to detect structurally defined albumin and Hb adducts in human blood, making it possible to characterize and measure human exposures to known electrophiles.[4, 47-48] In recent years, such methodology has been expanded to allow for screening of unknown adducts to human serum albumin and Hb via liquid chromatography–tandem mass spectrometry (LC-MS/MS)-based untargeted adductomics, making it possible to characterize the part of the human exposome containing electrophilic species.[79, 83] Unlike targeted analyses, untargeted screening does not require *a priori*

knowledge of adduct structures, allowing for the detection and identification of unknown exposures.

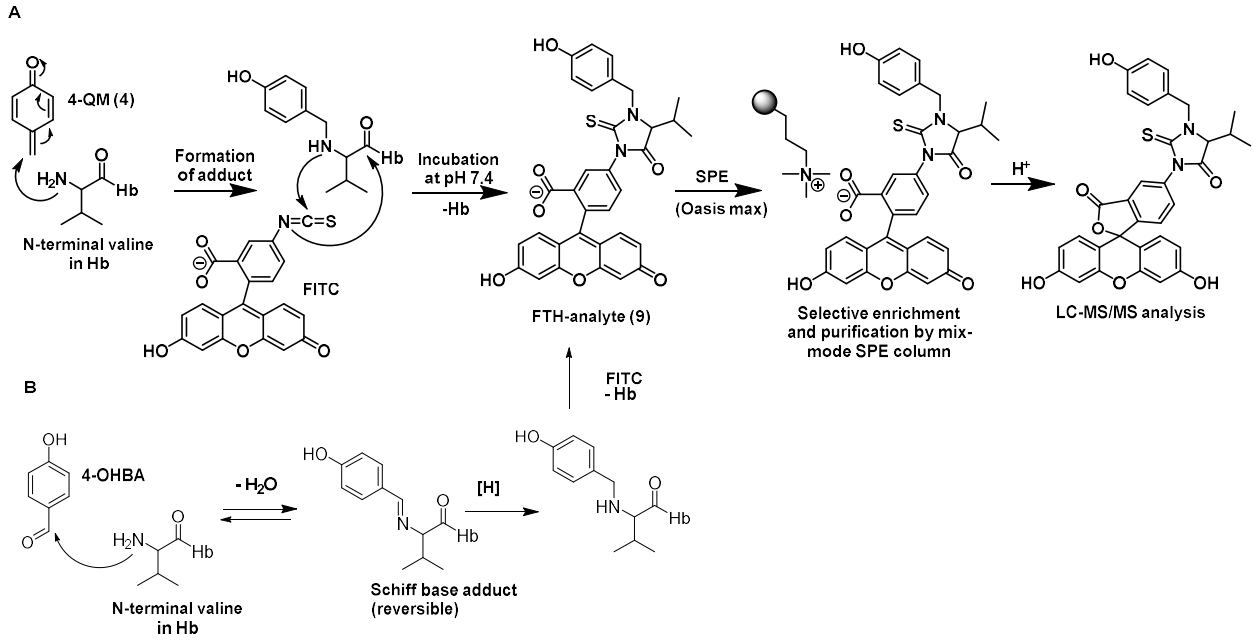
The screening methodology developed by us is based on the FI*R*E procedure, which employs modified Edman degradation with fluorescein isothiocyanate (FITC) to selectively detach and derivatize the adducted N-terminal amino acid of Hb.[89] The resulting fluorescein isothiohydantoin (FTH)-Val derivatives are analyzed by LC-MS/MS. In our first untargeted adductomic screening of Hb adducts in human blood, in addition to known adducts, 19 unknown adducts were detected, which initiated the challenging work of structurally identifying these adducts and their precursors. We have previously reported the identification of six of these adducts with human Hb and their probable precursors: ethylation, ethyl-vinyl ketone, acrylic acid, glyoxal, methylglyoxal, and 1-octen-3-one.[83-85] The most abundant N-terminal Hb adduct in human blood measured was the previously identified methyl modification, followed by carboxymethylation.[214] High levels of protein methylation in human blood can be explained by the endogenous methylating agent *S*-adenosylmethionine.[215] The carboxymethyl adducts are formed as a result of degradation of glycated Hb (HbA1c),[216-217] or from reaction with glyoxal, which comes mostly from dietary sources.[218] However, many of the unknown adducts have yet to be identified.

The third most abundant N-terminal Hb adduct in human blood samples was an unknown adduct with *m*/*z* of 595.153 of the FTH derivative, which corresponds to an added mass of 106.042 Da from the adduct to Val,[83, 85] that is, an elemental composition of $C_7H_6O$. To accommodate this molecular formula, the only option is an aromatic adduct

(see ref. 83 for the strategy used to generate hypotheses on adduct identities using adductome LC-MS/MS data). Considering all possible molecular structures available, we hypothesized that the most likely precursors of this adduct were 2- or 4-quinone methide (2-QM and 4-QM, respectively, Scheme 4.1A). QMs are conjugated organic compounds that contain a cyclohexadiene moiety with a carbonyl and a conjugated exocyclic methylene group. QMs have been implicated as intermediates in a variety of biological processes[219] and are responsible for the toxicity of many drugs. They are highly electrophilic Michael acceptors containing α,β-unsaturated carbonyls, which readily react with nucleophiles such as amino acid side chains of proteins[84, 220] and thus are anticipated to readily form adducts to N-terminal amino groups of Hb. Alternative precursors of the unknown adduct with added mass of *m*/*z* 106.042 include hydroxybenzaldehydes, which may form reversible Schiff base conjugates with the N-terminal amino group, followed by reduction to produce structurally identical hydroxybenzyl adducts (Scheme 4.1B).

In the present work, we describe complete structural identification of the unknown Hb adduct with an added mass of 106.042 Da as *N*-(4-hydroxybenzyl)valine (4-OHBn). This structural assignment was made using liquid chromatography (LC) - high-resolution mass spectrometry (HRMS) on an Orbitrap mass spectrometer and comparisons with authentic reference adducts synthesized in our laboratory. The adduct was quantified in blood of smokers and nonsmokers, with no significant differences between the two groups. Possible precursors of the novel adduct in humans are discussed.

**Scheme 4.1** Proposed formation of N-terminal Valine-quinone methide adducts in Hb followed by derivatization via FI*R*E procedure[83] and cleavage during acid workup. A) 4-quinone methide forms an adduct at the N-terminal valine in Hb. Derivatization with FITC cleaves the adducted valine from Hb, forming the FTH-analyte (9). FTH-analytes are enriched by MAX SPE and cyclized by acid work-up. B) 4-OHBA forms an adduct at the N-terminal valine in Hb, resulting in a Schiff base. Reduction of this adduct forms the stable 4-OHBn-Val-Hb adduct. This then undergoes FITC derivatization, SPE, and acid work-up as shown in A.



## 3.2 Materials and Methods

### *Materials*

**Caution:** Fluorescein isothiocyanate (FITC) is toxic and should be handled with care.

**Chemicals.** Fluorescein-5-isothiocyanate (isomer I, Reagent grade, FITC) was obtained from Karl Industries (Aurora, OH, USA). L-Valine p-nitroanilide hydrochloride (H-ValpNA) was obtained from Bachem (Bubendorf, Switzerland). D,L-Valine, sodium cyanoborohydride, benzaldehyde (BA), 2-hydroxybenzaldehyde (2-OHBA), 3-hydroxybenzaldehyde (3-OHBA), and 4-hydroxybenzaldehyde (4-OHBA) were purchased from Sigma-Aldrich Sweden AB. ortho-Quinone methide precursor (2-QMP, compound **4.5**, see Scheme 4.2) was generously provided by Professor Steven Rokita (John Hopkins University).[221-223] Acetonitrile (ACN), dimethylformamide (DMF), methanol, water and formic acid were obtained from VWR Chemicals. All solvents were of HPLC grade. All other chemicals and solvents were of analytical grade or higher.

### *Instruments and Equipment*

NMR Instrumentation and Mode of Analysis. $^1$H and HSQC spectroscopy characterization and quantification studies were performed on either a Bruker 500-MHz or a Bruker 700-MHz spectrometer using $CDCl_3$, $D_2O$, $CD_3OD$, or DMSO-$d_6$ as solvent.

Equipment for Incubation, Derivatization and Cleanup of Blood Samples. A thermomixer comfort and a 5804 R centrifuge with rotor F-45-30-11 (Eppendorf Nordic, Denmark) were

used for incubations and derivatization of blood samples. Oasis Max solid-phase extraction (SPE) cartridges (3 cc, 60 mg, 60 μm; mixed mode anion exchange) were obtained from Waters (Milford, MA, USA). The Hb analyzer (Hb 201+) was obtained from HemoCue (Ängelholm, Sweden).

Analytical HPLC System A. Our analytical HPLC system A consisted of an Agilent Technologies HPLC System (1100 model) equipped with a UV detector and an autosampler. Chromatographic separation was performed using a Luna $C_{18}$ (2) column (150 × 4.6 mm, 5 μm) and a gradient of 0.1% formic acid in water (A) and ACN (B) at a flow rate of 1 mL/min. Solvent composition was initially held at 10% B for 5 min, followed by a linear increase to 56% B in 13 min, followed by an increase to 95% B in 3 min, which was held for another 2 min before re-equilibration of the column for 6 min. UV absorbance was monitored at 254 nm.

Analytical HPLC System B. The analytical HPLC system B consisted of a Shimadzu LC-system with two pumps (LC 10AD), an auto injector (SIL HTC), and a UV detector (SPD 10A). The column was a $C_{18}$ reversed-phase column (Ace 5 $C_{18}$, 250 × 2.1 mm, 5 μm particles) from Advanced Chromatography Technologies, and the detection wavelength was set at 274 nm. The HPLC mobile phases consisted of ACN and water containing 0.1% formic acid, and the flow rate was 0.3 mL/min. The program was isocratic for 0.5 min with 5% ACN, followed by a linear increase to 90% ACN over 18 min, after which it was lowered back to 5% ACN in 2 min and kept at 5% ACN until the program was stopped after 25 min.

Semipreparative HPLC. The analytical HPLC system B, with exception of the column, was also used for the semipreparative purifications. The column used was a semipreparative Hichrom column (KR100-5C18-25098). The compounds were eluted isocratically with 35% ACN in water with 0.2% formic acid at a flow rate of 4 mL/min.

HPLC-MS. The HPLC-ESI-HRMS system consisted of a Dionex UltiMate 3000 LC system interfaced to an Orbitrap Q Exactive HF mass spectrometer (Thermo Fisher Scientific, MA, USA). Mobile phase A consisted of 0.1% formic acid in $H_2O$/ACN (95:5) and mobile phase B consisted of 0.1% formic acid in $H_2O$/ACN (5:95). The MS instrument settings were the same as described previously,[90] except that the *m*/*z* range of 550–600 was used in the data-independent acquisition DIA mode instead of the full 500–700 *m*/*z* range used previously.[90] The target compounds (*m*/*z* 595) were further studied using the parallel reaction monitoring (PRM) mode.

*Method 1.* Chromatographic separation was performed using a Discovery HS $C_{18}$ column (3.0 μm, 2.1 × 150 mm) with a Discovery HS $C_{18}$ guard column (3.0 μm, 2.1 × 20 mm) from Supelco Analytical. The gradient used started at 20% B and increased to 100% B in 25 min. It was thereafter held at 100% B for another 5 min, followed by re-equilibration at 20% B for 5 min. The flow rate used was 120 μL/min, and the injection volume was 20 μL.

*Method 2.* To improve the chromatographic separation and at the same time decrease the analysis time, a new method was developed using an Aquity UPLC HSS $C_{18}$ column (2.1

× 100 mm, 1.8 μm) from Waters. The gradient started at 20% B for 0.5 min, proceeded with an increase to 50% B in 3.5 min, then by an increase to 70% B in 3.5 min, and finally increased to 100% B in 1.5 min. Solvent composition was held at 100% B for 1 min before re-equilibration for 2.5 min. The flow rate used was 300 μL/min, and the injection volume was 20 μL.

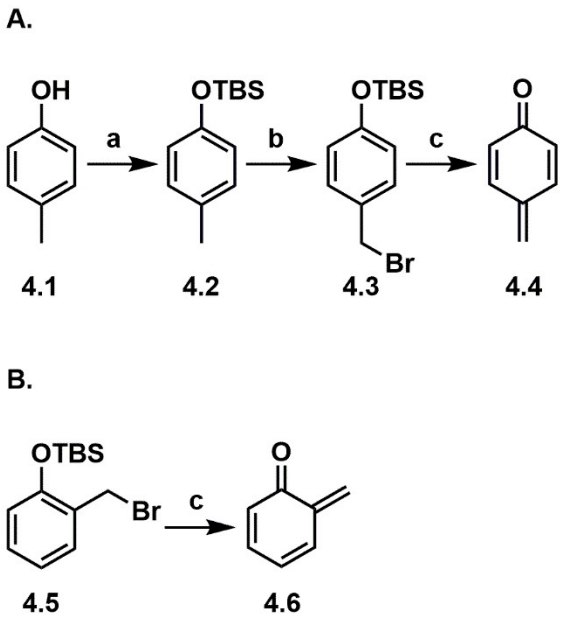
***Synthesis of tert-butyldimethyl(p-tolyloxy)silane (compound 4.2 in Scheme 4.2)***

4-Cresol (compound **4.1**, 2.0 g, 18.5 mmol) was dissolved in DMF (10 mL) under nitrogen. *tert*-Butyldimethylsilyl chloride (6.9 g, 46.2 mmol) and imidazole (6.2 g, 92.4 mmol) were added, and the reaction mixture stirred at room temperature overnight. After completion of the reaction (as monitored by TLC), the reaction mixture was quenched by addition of $H_2O$ (50 mL). The mixture was extracted with ethyl acetate (3 × 40 mL) and the combined organic phases were washed with brine, dried over $Na_2SO_4$, and concentrated under reduced pressure to yield a crude product as a colorless oil. The desired compound was purified by silica gel flash column chromatography (hexane/ethyl acetate step gradient of 99:1, 49:1, and finally 19:1) and isolated as a colorless oil (3.97 g, 96%). $^1$H-NMR ($CDCl_3$, 500 MHz): δ 0.00 (s, 6H), 0.80 (s, 9H), 2.09 (s, 3H), 6.54-6.55 (d, J = 5 Hz, 2H), 6.82-6.84 (d, J = 10 Hz, 2H).

***Synthesis of 4-(bromomethyl)phenoxy)(tert-butyl)dimethylsilane (compound 4.3 in Scheme 4.2)***

*N*-bromosuccinimide (0.80 g, 4.5 mmol) was added to a silyl ether-4-methyl-O-tert-butyldimethylsilylphenol (compound **4.2**, 1.0 g, 4.5 mmol) solution in $CCl_4$ (10 ml). The solution was heated to reflux for 10 min and 2,2'-azobis(2-methylpropionitrile) (AIBN) (0.026 g, 0.04 mmol) was added. The reaction mixture was further refluxed for 35 min, cooled to room temperature, and filtered. The filtrate was washed with water, dried with $Na_2SO_4$, and concentrated under reduced pressure to yield compound **4.3** as a colorless oil (0.72 g, 53%). $^1$H-NMR (Appendix A, Figure A2) ($CDCl_3$, 500 MHz): δ 0.01 (s, 6H), 0.75 (s, 9H), 4.49 (s, 2H), 6.62-6.64 (d, J = 10 Hz, 2H), 7.14-7.15 (d, J = 5 Hz, 1H).

**Scheme 4.2** (A) Synthesis of 4-QM Precursor and Its Activation in the Presence of KF and (B) Activation of 2-QM Precursor in the Presence of KF. (A) (a) TBS-Cl, imidazole, in DMF, under nitrogen, at room temperature overnight; (b) NBS, AIBN, in CCl4, 85°C, 35 min; (c) KF, in phosphate buffer (pH 7.4) and ACN, room temperature, 30 min. (B) (a) KF, in phosphate buffer (pH 8.4) and water, 37°C, 30 min.



***Synthesis of 2-OHBn-ValpNA (compound 4.7 in Scheme 4.3)***

Stock solutions of ValpNA (Scheme 4.3, 183 μM), 2-QMP (compound **4.5** in Scheme 3.4, 83 mM), and KF (33 mM) were made in water. A solution of 50 μL of 183 μM ValpNA, 50 μL of 33 mM KF, 50 μL of 100 mM phosphate buffer (pH 8.4), and 300 μL water was prepared. To start the reaction, 50 μL of 83 mM 2-QMP was added and the solution was incubated at 37°C for 30 min with continuous mixing. The presence of the 2-OHBn-ValpNA adduct was confirmed by LC-MS (full scan in positive mode).

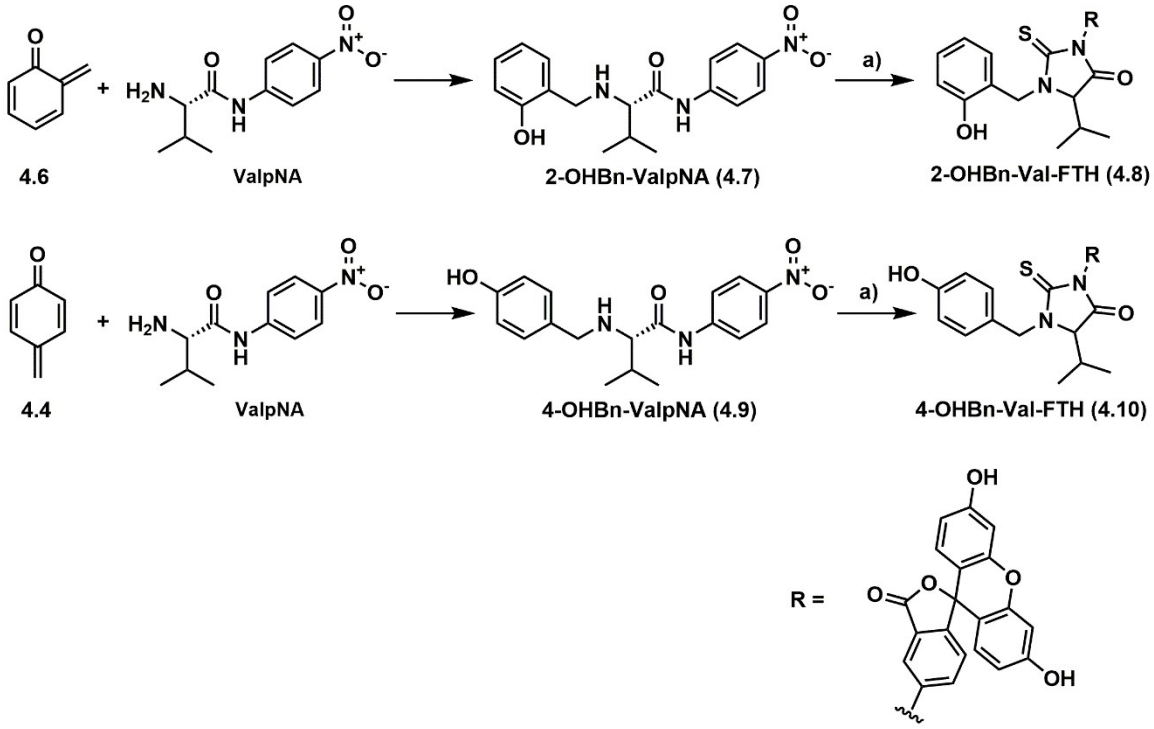
***Synthesis of 4-OHBn-ValpNA (compound 4.9 in Scheme 4.3)***

To synthesize the 4-OHBn-ValpNA, two different synthetic procedures were used, one small scale and one larger scale. For the small-scale reaction, ValpNA (Scheme 4.3, 33.2 nmol) and KF (33.2 nmol) were dissolved in 78 μL of phosphate buffer (pH 7.4) and added to 622 μL of ACN. To start the reaction, 0.1 mg of 4-QMP (compound **4.3**, 332 nmol) dissolved in 800 μL of ACN was added to this solution. The presence of the 4-OHBn-ValpNA adduct was confirmed by LC-MS (full scan in positive mode).

For the larger scale reaction, the amount of $H_2O$ in the reaction mixture was greatly reduced in order to minimize hydrolysis of the activated quinone methide. 4-QMP (compound **4.3** in Scheme 4.2A, 46.5 μmol, 10 μL) and ValpNA (4.6 μmol) were dissolved in anhydrous DMSO (77 μL). To start the reaction, 18 μL of KF (4.6 μmol) in $H_2O$ was added to this solution. The reaction was allowed to proceed for 1 h, after which the presence of the 4-OHBn-ValpNA adduct was confirmed by LC-MS (full scan in positive mode).

The reaction mixture was diluted with $H_2O$, filtered using a Costar Spin-X filter, and separated using HPLC system A. The 4-OHBn-ValpNA adduct eluted as a small peak at ~12.0 min, which was collected and solvent evaporated to dryness. The purity of this peak was confirmed by MS infusion and $^1H$ and HSQC NMR (Appendix A, Figure A3). 4-OHBn-ValpNA: $^1H$ NMR (DMSO-$d_6$, 700 MHz): δ 0.96–0.97 (m, 3H), 0.99–1.0 (m, 3H), 2.26 (br s, 1H), 3.73 (s, 1H), 4.00 (br s, 1H), 4.10 (br s, 1H), 6.75 (d, 2H), 7.26 (d, 2H), 7.81 (d, 2H), 8.26 (d, 2H), 9.13–9.27 (m, 2H), 9.72 (br s, 1H), 11.03 (s, 1H).

**Scheme 4.3** Synthesis of ortho-quinone methide-ValpNA adduct (2-OHBn-ValpNA, compound 4.7) and para-quinone methide-ValpNA adduct (4-OHBn-ValpNA, compound 4.9) and their corresponding FTH derivatives (compounds 4.9 and 4.10) from their respective precursor quinone methides, 2-QM (4.6) and 4-QM (4.4). A) FIRE derivatization: 15 µL of 1M KHCO3 and 5 mg of FITC added to 250 µL human whole blood samples spiked with compound 7 or 8, incubated overnight in a Thermomixer (37°C, 750 rpm).



***N-benzylvaline, N-(2-hydroxybenzyl)valine, N-(3-hydroxybenzyl)valine and N-(4-hydroxybenzyl)valine (compounds 4.11-4.14 in Scheme 4.4).***

Approximately 2.0 mmol of the respective benzaldehydes (BA, 2-, 3- and 4-OHBA, respectively), 2.2 mmol valine, and 2.5 mmol of sodium cyanoborohydride together with 100 mL methanol were placed in a 250 mL round-bottom flask. The reaction solution was kept at a temperature of 40 °C with stirring, and the progress of the respective reaction was followed by HPLC system B. When all benzaldehyde had reacted (after 4–48 h), the methanol in the reaction solution was removed with a rotary evaporator. Recrystallization from an ethanol–water mixture produced the *N*-benzylvalines as white crystals in yields ranging from 50 to 60%. The purity and structures of these compounds were confirmed by $^1$H and HSQC NMR (Appendix A, Figures A4–7).

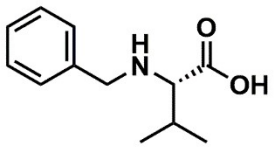
*N-Benzylvaline (Bn-Val):* $^1$H NMR ($D_2O$, 500 MHz): δ 7.40–7.43 (m, 5H), 4.22 (d, $J$ = 13 Hz, 1H), 4.08 (d, $J$ = 13 Hz, 2H), 3.34 (d, $J$ = 4.5 Hz, 1H), 2.06–2.13 (m, 1H), 0.93 (d, $J$ = 7 Hz, 3H), 0.87 (d, $J$ = 7 Hz, 3H).

*N-(2-Hydroxybenzyl)valine (2-OHBn-Val):* $^1$H NMR ($CD_3OD$, 500 MHz): δ 0.25–0.31 (m, 6H), 1.50–1.53 (m, 1H), 2.52–2.57 (m, 1H), 3.41 (d, $J$ = 13 Hz, 1H), 3.48 (d, $J$ = 13 Hz, 1H), 6.15–6.19 (m, 2H), 6.54–6.59 (m, 2H).

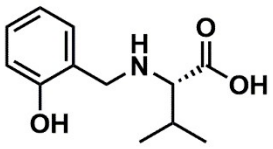
*N-(3-Hydroxybenzyl)valine (3-OHBn-Val):* $^1$H NMR (DMSO-d6, 500 MHz): δ 0.88–0.92 (m, 6H), 1.90–1.94 (m, 1H), 2.86 (d, $J$ = 5 Hz, 1H), 3.57 (d, $J$ = 13.5 Hz, 1H), 3.80 (d, $J$ = 13.5 Hz, 1H), 6.66 (d, $J$ = 8 Hz, 1H), 6.75–6.79 (m, 2H), 7.10–7.13 (m, 1H).

*N-(4-Hydroxybenzyl)valine (4-OHBn-Val):* $^{1}$H NMR ($D_2O$, 500 MHz): δ 0.74 (d, $J$ = 7 Hz, 3H), 0.79 (d, $J$ = 7 Hz, 3H), 1.76–1.80 (m, 1H), 2.91 (d, $J$ = 5 Hz, 1H), 3.47 (d, $J$ = 13 Hz, 1H), 3.67 (d, $J$ = 12.5 Hz, 1H), 6.56 (d, $J$ = 8 Hz, 2H), 7.02 (d, $J$ = 8 Hz, 2H).

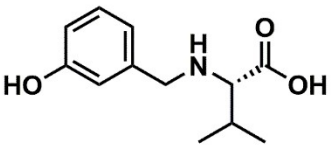
**Scheme 4.4** Structures of valine adducts formed from benzaldehyde and hydroxybenzaldehydes.



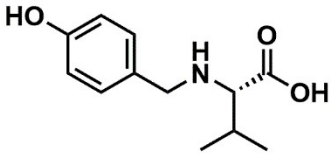

N-benzylvaline (4.11)
(Bn-Val)




N-(2-hydroxybenzyl)valine (4.12)
(2-OHBn-Val)




N-(3-hydroxybenzyl)valine (4.13)
(3-OHBn-Val)




N-(4-hydroxybenzyl)valine (4.14)
(4-OHBn-Val)

### *Synthesis of FTH derivatives of benzylvalines*

Approximately 0.12 mmol of the respective *N*-benzylvalines and 0.10 mmol of FITC were dissolved in 20 ml of phosphate buffer (0.1 M, pH 7.4). The solutions were kept at 60 °C and the progress of the reactions was followed by HPLC system B. When all FITC had reacted, the reaction mixtures were evaporated to dryness in a rotary evaporator. Thereafter the crude products were dissolved in methanol and purified by semi-preparative HPLC, which afforded the FTH of the *N*-benzylvalines as orange-brown powders in yields of 30 – 40 %. FTH derivatives were confirmed by HRMS mass. Benzylvaline FTH (Bn-Val-FTH): HRMS (Orbitrap ESI) calculated for $C_{33}H_{27}N_2O_6S^+$ 579.1584, found 579.1585. *N*-(2-hydroxybenzyl)valine (2-OHBn-Val-FTH): HRMS (Orbitrap ESI) calculated for $C_{33}H_{27}N_2O_7S^+$ 595.1534, found 595.1538. *N*-(3-hydroxybenzyl)valine: FTH (3-OHBn-Val-FTH): HRMS (Orbitrap ESI) calculated for $C_{33}H_{27}N_2O_7S^+$ 595.1534, found 595.1536. *N*-(4-hydroxybenzyl)valine FTH (4-OHBn-Val-FTH): HRMS (Orbitrap ESI) calculated for $C_{33}H_{27}N_2O_7S^+$ 595.1534, found 595.1535.

### *Quantitative and Qualitative Analysis of Hb Adducts in Human Blood*

#### *Blood samples and study population*

Commercial human blood used for incubation experiments were obtained from Komponentlab at Karolinska University Hospital Huddinge (Stockholm, Sweden). Bovine blood (with citrate) was obtained from Håtunalab AB (Bro, Sweden). Blood samples from six smokers and six nonsmokers, collected with approval from the Regional Ethical Review

Board in Stockholm, Sweden (no 96-312), were also analyzed. Upon arrival, all blood samples, except the fresh whole blood, were separated into red blood cells (RBCs) and plasma. The RBCs were washed with 0.9% (w/v) sodium chloride and lysed with distilled water. The hemolyzed RBC samples were stored at -20 °C until the day of analysis.

***FIRE method for derivatization and clean-up of blood samples***

According to the FI*R*E procedure[89, 224] blood samples (250 µL each) were derivatized by addition of $KHCO_3$ (1M, 15 µL) and FITC (5 mg) to each sample, followed by incubation overnight in a Thermomixer (37°C, 750 rpm). A relevant internal standard was added when appropriate. Thereafter, ACN (1.4 mL) was added to precipitate out the proteins, and the samples were centrifuged (10 min at 11,000 rpm). The SPE columns were conditioned with ACN (one column volume) and 0.01 M ammonium hydroxide (1/3 column volume). Next, 25 µL of 1 M ammonium hydroxide was added to the supernatant to alkalize them before transferring them to the conditioned SPE columns. After loading, columns were washed with ACN and $H_2O$ (one column volume each) followed by 0.25% (w/v) cyanoacetic acid in $H_2O$ (1/2 column volume). Analytes were then eluted with 1.2 mL of 0.25% (w/v) cyanoacetic acid in ACN. Samples were dried under nitrogen and then reconstituted in 100 µL of $H_2O$/ACN (3:2, v/v) in preparation for HPLC-ESI-HRMS analysis.

***Spiking of blood samples with 4-OHBn-ValpNA and 2-OHBn-ValpNA***

Aliquots of the 4-OHBn-ValpNA (compound **4.9**) reaction mixture were added (4.4, 22, and 88 µL, respectively) to three human whole blood samples. Aliquots of the 2-OHBn-ValpNA reaction mixture (compound **4.7**, 100 µL and 200 µL) were added to another two human whole blood samples. Three additional human whole blood samples and one bovine RBC sample were used as controls. All seven samples were treated according to the general procedure and analyzed with LC-HRMS according to Method 1.

***Incubation of fresh whole blood with BA, 2-OHBA, 3-OHBA, and 4-OHBA***

Aliquots of stock solutions of BA, 2-OHBA, 3-OHBA, and 4-OHBA (50 mM in ethanol (95%), 10 µL) were added to fresh human whole blood (990 µL), resulting in final concentrations of electrophiles of 0.50 mM. Two control human whole blood (990 µL) samples, to which ethanol only was added (95%, 10 µL), were prepared. The samples were incubated overnight in a Thermomixer (37°C, 500 rpm). The following day, the tubes were centrifuged (2000×g, 4°C, 15 min) and the top-layers, i.e. the plasma, were removed. The remaining bottom fraction in each blood sample, i.e. the RBCs, were washed with 0.9% (w/v) sodium chloride by gentle mixing followed by centrifugation (2000×g, 4°C, 10 min). The wash procedure was repeated twice. Thereafter, all six samples (incubated samples and controls) were derivatized according to the general procedure described above with the exception that all volumes and amounts were up-scaled to accommodate the larger blood volume used. The samples were analyzed with LC-HRMS according to Method 2.

***Analysis of background levels of 2-OHBn-Val and 4-OHBn-Val in blood samples from smokers and nonsmokers***

The levels of 2-OHBn and 4-OHBn adducts to N-terminal valine in Hb were measured in blood samples from six smokers and six nonsmokers using the general procedure. Prior to derivatization with FITC, Hb contents were measured in the samples. After incubations with FITC but before clean-up, 3-OHBn-Val-FTH (0.1 pmol/μL, 35 uL) was added as internal standard. The samples were analyzed with LC-HRMS according to Method 2.

## 4.3 Results

### 4.3.1 Structural Identification of the Unknown Hb Adduct

The unknown N-terminal Val Hb adduct derivative with molecular ion $[M + H]^+$ at *m/z* 595.153 was detected in untargeted adductomics screening of N-terminal valine adducts in hemoglobin using the FI*R*E procedure.[83, 90] In our HPLC-ESI-MS/MS experiments, this unknown eluted at 17.75 min (Figure 4.1A) and exhibited major MS/MS fragments at *m/z* 489.1121, 445.0492, 390.0433, 373.0821 (Figure 4.2A). Based on the accurate mass of 595.153 (corresponding to a mass addition of 106.042 Da to N-terminal valine) and retention time, the adduct was proposed to be (hydroxybenzyl)-valine (OHBn-Val) derived from 2- or 4-QM (Scheme 4.1).

2- and 4-QM are extremely reactive molecules that are readily hydrolyzed at ambient conditions.[222, 225] Our initial efforts to generate the 2-QM in situ through oxidation of o-cresol using a previously published method[226] failed, with no adduct formation seen by LC-MS infusion. Therefore, we employed the elegant strategy developed by the Rokita lab to generate chemically stable QM precursors (QMPs).[222] QMP activation to quinone methides is achieved by the addition of KF.[222]

**Figure 4.1.** Mass chromatograms showing the exact mass (*m/z* = 595.15181-595.15539) of the FTH derivative of the identified adduct in (**A**) a background sample of human blood, (**B**) a human blood sample spiked with 4-OHBn-ValpNA (compound **4.9**), and (**C**) a human blood sample spiked with 2-OHBn-ValpNA (compound **4.7**).

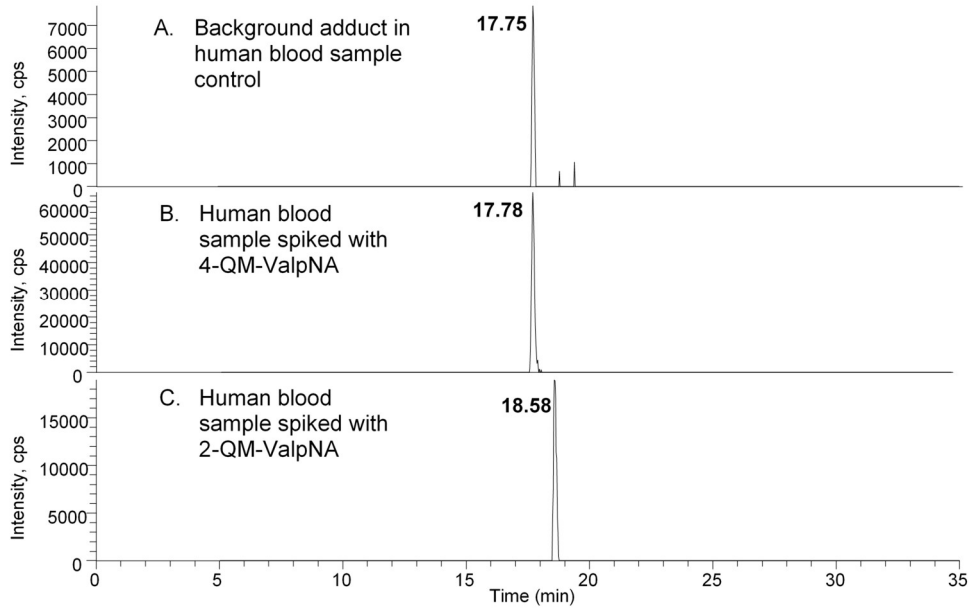


**Figure 4.2.** $MS^2$ spectra of (**A**) background adduct in human blood sample, (**B**) authentic standard 4-OHBn-Val-FTH (compound **4.10**), and (**C**) authentic standard 2-OHBn-Val-FTH (compound **4.8**).

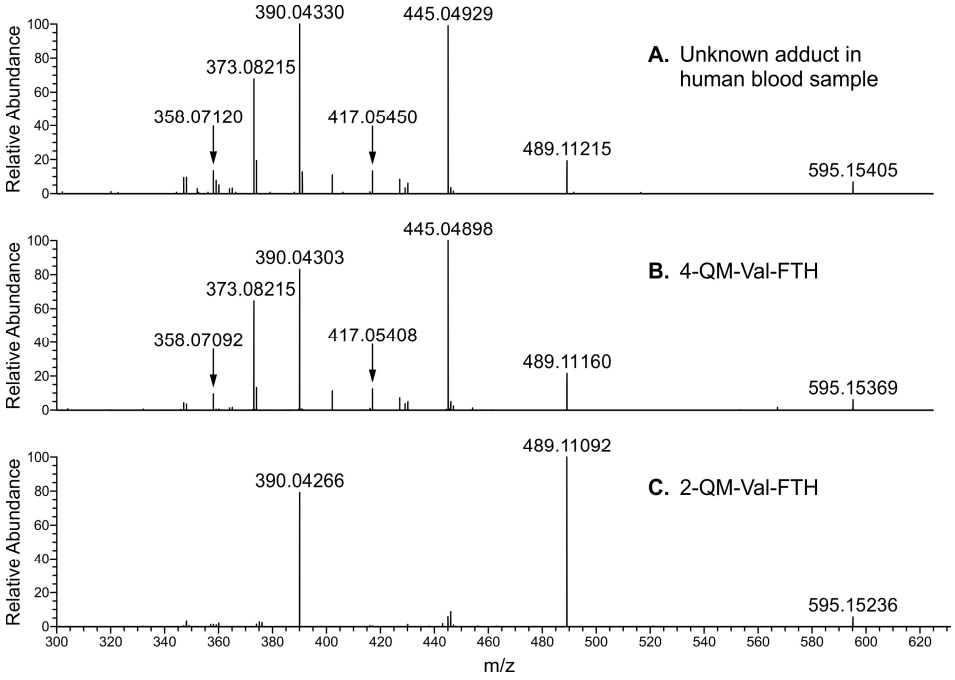


**Table 4.1** Comparison of Fragmentation Patterns of Adduct 595 for (**A**) the background analyte, (**B**) the synthesized 4-OHBn-Val in an incubated blood sample, and (**C**) the synthesized 2-OHBn-Val in an incubated blood sample.

| | | Relative Intensity of Fragments *m/z* (%) | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | RT (min) | 358 | 373 | 390 | 417 | 445 | 489 |
| A | 17.75 | 16 | 65 | 100 | 13 | 98 | 21 |
| B | 17.78 | 17 | 48 | 100 | 9 | 73 | 16 |
| C | 18.58 | 2 | 0 | 59 | 0 | 6 | 100 |

Authentic 2- or 4-QM Val adducts, i.e. 2- and 4-OHBn-ValpNA, were generated using the *p*-nitrophenyl derivative of Val (ValpNA, Scheme 4.3) as described in our previous publications.[84] The o-phenyl group of ValpNA is cleaved to reveal free Val during the FI*R*E procedure (Scheme 4.3). These standards of 2-OHBn-ValpNA and 4-OHBn-ValpNA (compounds **4.7** and **4.9** in Scheme 4.3, respectively) were prepared using 2-QMP (compound **4.6**) and 4-QMP (compound **4.4**) and ValpNA (Scheme 4.3). QMPs were deprotected using KF to form the active quinone methides in the presence of ValpNA, where it immediately reacted to form the adducts. 2-OHBn-ValpNA (compound **4.7**) and 4-OHBn-ValpNA (compound **4.9**) were isolated by HPLC, and their structures were verified by LC-MS/MS, $^{1}$H and HSQC NMR (Appendix A, Figure A3).

To examine potential formation of QM-Val adducts in human blood, authentic 2-OHBn-ValpNA (compound **4.7**, derived from 2-QM) and 4-OHBn-ValpNA (compound **4.9**, derived from 4-QM) (Scheme 4.3) were spiked into human blood. These, along with unspiked human blood samples, underwent derivatization with FITC followed by LC-HRMS analysis. The accurate mass, retention times, and fragmentation patterns of 2-OHBn-Val-FTH and 4-OHBn-Val-FTH from the spiked blood samples were compared to the unknown adduct seen in the unspiked blood samples. While the accurate mass of the 2-OHBn-Val-FTH adduct (compound **4.8**, *m/z* 595.1524) matched the unknown (*m/z* 595.1540), the retention times differed (17.75 vs 18.58 min, see Figure 4.1A,C). Furthermore, the fragmentation patterns of the unknown adduct at *m/z* 595.1540 and authentic 2-OHBn-Val-FTH different significantly, with fragments at m/z 445.0492 and

373.0821 not observed for 2-OHBn-Val-FTH (Figures 4.2A,C). Authentic 4-OHBn-Val-FTH (compound **4.10**) eluted at the same retention time as the unknown adduct (17.7 min, Figure 4.1A,B) and displayed the same accurate mass (*m/z* 595.1537) and MS/MS fragmentation pattern (Figure 4.2A,B). Based on this evidence, the unknown adduct was identified as *N*-(4-hydroxybenzyl)valine (4-OHBn-Val) (Figure 4.1).

### 3.3.2 Alternative Formation of N-(4-hydroxybenzyl)valine Adducts from 4-OH-benzaldehyde

In addition to the 4-QM-mediated mechanism shown in Scheme 4.1A, the *N*-(4-hydroxybenzyl)valine adduct may also originate from reaction with 4-OH-benzaldehyde, followed by reduction (Scheme 4.1B). To explore this possibility, valine was reacted with BA, 2-OHBA, 3-OHBA, and 4-OHBA, followed by reduction with sodium cyanoborohydride and derivatization with FITC. The structures of these standard compounds were confirmed by $^{1}$H and HSQC NMR (Appendix A, Figures A4-7). As expected, NMR spectrum, accurate mass, and MS/MS fragmentation of the FTH derivative of 4-OHBA-Val adduct matched those of 4-OHBn-Val-FTH derived from 4-QM, suggesting that 4-OHBA could be another potential precursor to the identified 4-OHBn-Val-FTH adduct (Scheme 4.1).

To determine whether the Schiff base formed between 4-OHBA and the N-terminal valine of Hb can be reduced in vivo, fresh human blood was incubated with BA, 2-OHBA, 3-OHBA, and 4-OHBA, followed by incubation with FITC and LC-HRMS analysis. All four benzaldehyde derivatives (Scheme 4.4) were detected and formed the same adducts

as those produced synthetically using sodium cyanoborohydride. Therefore, the reductive capacity of blood appears to be sufficient for 4-OHBn-Val to be formed from the Schiff base of 4-OHBA to N-terminal valine (Scheme 4.1B). By comparison of the spiked blood with the non-spiked blood, relatively large amounts of the 4-OHBn-Val-FTH and minor amounts of 2-OHBn-Val-FTH could be detected in the non-spiked blood, whereas 3-OHBn-Val-FTH and Bn-Val-FTH could not be detected in the non-spiked blood.

### 3.3.3 Quantitation of N-(4-hydroxybenzyl)valine Adduct Levels in Human Blood

The 4-OHBn-Val adduct was quantified in Hb samples from six nonsmoker and six smoker blood samples. As no background levels were detected of 3-OHBn-Val-FTH (Scheme 3.5) this compound was used as an internal standard. For quantification, calibration curves were established from synthesized standards of 3-OHBn-Val-FTH ($R^2$ =0.984, five concentrations (0.001 – 0.005 pmol/μL) in duplicate) and 4-OHBn-Val-FTH ($R^2$ = 0.987, five concentrations (0.01 – 0.05 pmol/μL) in duplicate). The adduct levels in human samples ranged from 140 to 650 pmol/g Hb, with an average adduct level of 380 ± 160 pmol/g Hb (Table 4.2). There was no significant difference between the smoker and nonsmoker samples (*p* value for two-sided *t* test was 0.48).

**Table 4.2** Quantified Adduct Levels of 4-OHBn-Val and 2-OHBn-Val Adducts in Nonsmokers and Smokers.

| | | Nonsmokers (pmol/g Hb) ($n$ = 6) | | Smokers (pmol/g Hb) ($n$ = 6) | |
|---|---|---|---|---|---|
| Adduct | RT (min) | Range | Mean ± SD | Range | Mean ± SD |
| 4-OHBA | | 143–618 | 374 ± 179 | 247–651 | 379 ± 149 |
| 2-OHBA | | 2.0–6.7 | 4.7 ± 1.6 | 3.6–6.1 | 5.3 ± 1.3 |

Quantification of 2-OHBn-Val isomer was also attempted in the same blood samples with the use of a calibration curve established by synthesized 2-OHBn-Val-FTH ($R^2$ = 0.983, five concentrations (0.001–0.005 pmol/μL) in duplicate). The estimated level of 2-OHBn-Val adducts was much lower than 4-OHBn-Val, ranging from 2.0 to 6.7 pmol/g Hb, with an average adduct level of 5.0 ± 1.4 pmol/g Hb (Table 4.2). The levels are below the lowest point in the calibration curve used so they are rough estimations only.

## 4.4 Discussion

### 4.4.1 Identification and Quantification of the 4-OHBn Adduct in Human Blood

The FTH-derivative of the unknown adduct with $[M + H]^+$ m/z 595.153 found in human blood samples was identified as 4-OHBn-Val-FTH using authentic standards prepared independently. Both the retention times and the accurate mass for synthetic 4-OHBn-Val-FTH and the FTH-derivative of the adduct present in human samples matched (Figure 4.1, Figure 4.2, Table 4.1). Adduct identity was further confirmed by comparing MS/MS fragmentation patterns. In contrast, HPLC retention time and fragmentation of the background adduct differed from that of the synthetic 2-OHBn-Val-FTH isomer, despite having the same exact mass.

Using a calibration curve generated from the synthesized standard matching the FTH of the unknown adduct, 4-OHBn-Val-FTH, and synthesized 3-OHBn-Val-FTH as an internal standard, we quantified the level of the background adduct at m/z 595.153 in human blood to a mean of approx. 380 pmol/g Hb. This adduct level corresponds to approximately seven OHBn-Val adducts per $10^6$ Hb molecules, indicating the importance of the FI*R*E procedure to enrich these adducts. Additionally, the mean adduct level and the large range in adduct level (143 – 651 pmol/g Hb) is similar to that of our previous studies with semi-quantification without synthesized standard and a less suitable internal standard of this unknown adduct.[83, 90]

### 4.4.2 Potential Sources for the Formation of the 4-OHBn and 2-OHBn Adducts

Once the identity of this unknown adduct was confirmed, we sought to identify its possible sources. In addition to the human blood sample controls that were investigated, a bovine blood sample was also analyzed. While the level of the 4-OHBn adduct detected in the bovine sample was far lower than the level found in human samples, its presence shows that this adduct is not unique to humans (as indicated earlier).[83] The adduct levels quantified in smokers and nonsmokers were not significantly different, thus the source of the adduct is highly unlikely to come from smoking. Additionally, this adduct was present in all blood samples analyzed in this study and in our previous studies using the FI*R*E procedure, indicating that the source may be endogenous or present in common foods. Independently, Rappaport's research group has observed an adduct of the same added mass to Cys34 in human serum albumin in adductomics screening of different sample materials. It has been suggested to correspond to an adduct from benzaldehyde,[80] and just recently quinone methide was also suggested as a precursor.[81] These suggestions were not verified through comparison with reference compounds, but indicate that the same adduct has been observed to another nucleophilic site in serum albumin. Considering the relatively high abundance of this adduct observed to Hb and the nucleophilic strength of the respective nucleophilic sites, modification of serum albumin by the same precursor seems likely.

QMs can be formed *in vivo* through two electron oxidation reactions of 4-alkyl-substituted phenols catalyzed by cytochrome P450 enzymes or peroxidases.[227] Substitutions, especially at the 2- and 6-positions of the phenol ring, strongly influence the

reactivity of the QM. In general, the larger the substitutions, the lower the reactivity of the QM, due to sterics.[226, 228] This would imply that a completely unsubstituted QM would have the highest reactivity.[229] However, due to this high reactivity, unsubstituted QMs have been more difficult to study than more substituted ones.

QMs can be formed *in vivo* through two electron oxidation reactions of 4-alkyl-substituted phenols catalyzed by cytochrome P450 enzymes or peroxidases.[227] Substitutions, especially at the 2- and 6-positions of the phenol ring, strongly influence the reactivity of the QM. In general, the larger the substitutions, the lower the reactivity of the QM, due to steric effects.[226, 228] This would imply that a completely unsubstituted QM would have the highest reactivity.[229] However, due to this high reactivity, unsubstituted QMs have been more difficult to study than more substituted ones. The Rokita lab has previously shown that unsubstituted ortho-quinone methides can form adducts at nucleophilic positions in DNA, including the N1 and $N^6$ position of dA, N3 position of dC, and N1, $N^2$, and N7 positions of dG.[26] They also found that these adducts are reversible,[230-231] though the stability of these adducts was highly dependent on the substituents on the quinone methide.[232] Additionally, Bolton et al. found that substituted *para*-quinone methides of differing reactivity all formed adducts with the α-amino groups of lysine, histidine, tyrosine, and serine, and with the thiol of cysteine.[233] The most electrophilic of these substituted quinone methides was also able to alkylate the side chain nitrogens of lysine and histidine.[233] Due to the high electrophilicity of the unsubstituted 4-QM, it is

likely capable of not only alkylating proteins at positions other than N-terminal valine but also nucleophilic positions in DNA.

This high reactivity of QMs was supported by our experiments examining ValpNA reactivity towards 4-QM, as the reaction was complete within an hour. However, *para*-quinone methides are also rapidly hydrolyzed, making it unlikely that this adduct is formed by free 4-QM in human blood. In order to form a large enough amount of 4-OHBn-ValpNA adduct to be quantified by LC-MS, the reactions had to be carried out in the smallest amount of aqueous solvent as possible in order to limit the hydrolysis of 4-QM. Reactions performed under physiological conditions only produced barely detectable amounts of Val adducts (results not shown). Thus, we hypothesize that there must be an alternative origin of the adduct.

One hypothesis is that this adduct is formed as a byproduct of tyrosine degradation by the [FeFe] hydrogenase maturase, HydG.[234] HydG degrades tyrosine in order to form CO and $CN^-$ ligands for the enzyme's [FeFe] cluster, resulting in the generation of 4-oxidobenzyl radical as a byproduct. While this radical is usually further degraded to form 4-cresol as the end product, it is possible that it could instead react to form the 4-OHBn adduct we observed.

Another possible source of this adduct is 4-OHBA. NMR of the synthesized reaction product between 4-OHBA and valine showed that the same 4-OHBn-Val adduct was formed as for 4-QM and valine. However, in order to form the 4-OHBn-Val adduct

from 4-OHBA, the (reversible) Schiff base adduct, which is formed first, must be reduced (and stabilized). To explore if this may happen in vivo, 4-OHBA was incubated with fresh whole blood overnight, and indeed the reduced Schiff base adduct (4-OHBn-Val adduct) was detected (after derivatization with the FI*R*E procedure). Thus, it appears as if 4-OHBA could be a possible precursor for the detected 4-OHBn-Val adduct. 4-OHBA has previously been shown to be formed naturally: for instance it is an active component of *Gastrodia elata*, a Chinese herbal medicine used to treat headaches and migraines, and of *Dendrocalamus asper* bamboo shoots.[235-236] In addition, 4-OHBA could be formed in natural processes from tyrosine,[237] including in the Maillard reaction that occurs during cooking of food.[238] It is also a precursor to vanillin that is naturally found in vanilla beans,[239] and is an important volatile component of vanilla aroma and flavor.[240-241] However, further work must be done to identify the source of this adduct in humans.

2-Hydroxybenzaldehyde (salicylaldehyde, 2-OHBA), which can form an adduct with the same structure as the adduct from 2-QM (2-OHBn-Val), is a naturally occurring component in buckwheat,[242] cinnamon, and tea and is found in many foods. It is also structurally related to salicylic acid, one of the most commonly ingested analgesics and anti-inflammatory agents. Low levels of 2-OHBn-Val adduct were detected in both smokers and nonsmokers in this study (5.0 ± 1.4 pmol/g Hb).

## 4.5 Conclusions

The unknown human adduct with m/z 595.153 was identified as an adduct corresponding to the addition of a 4-hydroxybenzyl group to N-terminal valine in Hb. By reaction with FITC, according to the FI*R*E procedure, *N*-(4-hydroxybenzyl)-5-isopropyl-2-thioxo-imidazolidin-4-one (4-OHBn-Val-FTH) is formed. Two possible precursors have been identified; one being *para*-quinone methide (4-QM) which can form the 4-OHBn-Val adduct via a Michael addition with the N-terminal valine, the second possible precursor would be 4-hydroxybenzaldehyde (4-OHBA), which can form the detected adducted via a Schiff base reaction with the N-terminal valine followed by reduction. It is essential to establish which of the two probable precursors is the main contributor to the observed adduct to assess its toxicological relevance.

# V. SUMMARY AND CONCLUSIONS

Humans are exposed to a wide range of electrophilic agents from exogenous sources such as air or food and endogenous sources such as metabolism byproducts.[1-4] These electrophiles can form adducts at nucleophilic sites in biomolecules including DNA and proteins such as hemoglobin (Hb). The formation of DNA adducts in particular is associated with an increased risk of adverse health effects, such as cancer, aging, and chronic diseases.[5-7]

One such electrophile, 1,3-butadiene (BD), is a known human carcinogen[91] found in cigarette smoke, automobile exhaust, wood burning fires, and the rubber and polymer industry.[92-95] BD has a high cancer risk index compared to other carcinogens present in cigarette smoke[119] and is likely to contribute to the development of smoking-induced lung cancer.[118] BD is metabolically activated by cytochrome P450 monooxygenases CYP2E1 and CYP2A6 to for three reactive epoxide metabolites: 3,4-epoxy-1-butene (EB),[100] 1,2,3,4-diepoxybutane (DEB),[102] and 3,4-epoxy-1,2-butanediol (EBD)[103] (Chapter 1, Scheme 1.8). These reactive epoxides are responsible for the genotoxicity and carcinogenicity of BD, through their ability to form covalent adducts with nucleophilic bases in DNA. These epoxides can be detoxified through hydrolysis or glutathione conjugation by glutathione S-transferase (GST) enzymes to form BD-mercapturic acids that are excreted in the urine (Chapter 1, Scheme 1.9). Both BD-DNA adducts and BD-mercapturic acids have been previously used as biomarkers of human exposure to BD.

In chapter 2 of this thesis, the formation and repair of the DEB-induced DNA-DNA interstrand crosslink (ICL) *bis*-N7G-BD was quantified using nanoLC-nanoESI$^+$-MS/MS

analyses. This analysis method was applied to two different studies – to investigate the role nucleotide excision repair (NER) and Fanconi Anemia (FA) repair pathways in the repair of ICLs such as *bis*-N7G-BD, and to investigate the ability of DEB to form ICLs, micronuclei, and t(9:22) translocations associated with chronic myeloid leukemia (CML). We employed an offline HPLC analyte enrichment procedure followed by sensitive nanoLC-nanoESI$^+$-MS/MS analysis on a TSQ Quantiva triple quadrupole mass spectrometer (Thermo Scientific, Waltham, MA) using a method previously developed in our lab.[71] We successfully quantified *bis*-N7G-BD in DNA from human fibroblast cell lines deficient in *FANCD2*, a component of the FA pathway (PD20 cells) and in *XPA*, a component of NER (XPA cells), and their isogenic corrected controls (PD20corr and XPAcorr), that were treated with DEB for 3 h and then allowed to recover for up to 72 h. Cytotoxicity studies showed that both FA and NER deficient cells were more sensitive to DEB treatment then their isogenic controls (Chapter 2, Figure 2.1). Deficiency in the FA protein FANCD2 (PD20 cells) had no significant effect on the kinetics of *bis*-N7G-BD crosslink formation and repair (Chapter 2, Figure 2.2), thus indicating that *FANCD2* deficiency does not prohibit the repair of *bis*-N7G-BD crosslinks. Deficiency in the NER protein XPA (XPA cells) showed a slight, but not statistically significant, increase in *bis*-N7G-BD levels after DEB treatment compared to its control, but both deficient and corrected cell lines showed similar kinetics of repair (Chapter 2, Figure 2.2). This indicates that like *FANCD2*, *XPA* deficiency does not prohibit the repair of *bis*-N7G-BD crosslinks. This is possibly due to the large amount of overlap between these DNA repair pathways,

with a functional repair pathway taking over repair when another is nonfunctional. It is also possible that the induction of DNA double strand breaks (DSBs) by ICLs such as *bis*-N7G-BD is responsible for the increased sensitivity of repair-deficient cells to DEB treatment. *Bis*-N7G-BD was also successfully quantified in human promyelotic leukemia cells (HL60) after treatment with low levels of DEB. This study was designed to test whether the epidemiological-based association of BD exposure to the development of CML could be attributed to the production of t(9:22) translocations. t(9:22) translocations are essential to CML development, as they encode for the *BCR-ABL* fusion gene that drives CML oncogenesis. In order for DEB to lead to the development of CML, it must be able to produce these translocations. While DEB treatment was found to induce a dose-related formation of *bis*-N7G-BD and micronuclei as a quantitative index of DSBs, and a dose-dependent increase in cytotoxicity (Chapter 2, Figures 2.5 and 2.6), it did not induce t(9:22) translocations. Further investigation is required to determine the cause of the association between BD exposure and CML development.

In chapter 3 of this thesis, we investigated the effects of GSTT1 expression on the detoxification of EB, EB-DNA adduct formation, and apoptotic cell death after exposure to EB. After adjusting for smoking history, the risk of developing smoking-induced lung cancer differs by ethnicity: African American and Native Hawaiian smokers have the highest risk, while Japanese American and Latino smokers have much lower risks, and White smokers have intermediate risk.[194] While the source of these ethnic differences is not known, frequencies of genetic polymorphisms in metabolic enzymes differ between

ethnic groups, which can lead to changes in the balance between carcinogen bioactivation and detoxification. Previously, the levels of MHBMA, a metabolite of the glutathione conjugate of EB, were measured in urine samples of White, Japanese American, and African American smokers, where they differed significantly based on ethnicity ($p = 4.0 \times 10^{-25}$) and correlated to smoking-induced lung cancer risk.[198] A genome-wide association study was conducted to identify genetic polymorphisms that could account for ethnic differences in BD metabolism. Only *GSTT1* copy number was found to have a significant association to MHBMA levels ($p < 0.0001$).[198] GSTT1 is known to catalyze glutathione conjugation to BD and its metabolite, and thus the *GSTT1* gene could potentially play a role in BD metabolism, influencing the susceptibility to developing lung cancer in smokers. In this study, we developed a HPLC-ESI$^+$-MS/MS for the quantitation of EB-GSH in cellular media. We used this methodology, along with previously developed methodologies for the EB-DNA adduct EB-GII and apoptosis, to quantify EB conjugate formation, EB-DNA adduct formation, and apoptosis after EB exposure in cells that were either did or did not express GSTT1. HapMap cell lines were chosen for this study as they are human derived B-lymphocytes derived from people from all over the world that have been extensively genotyped, making them more representative of interindividual differences than a single monoclonal cell line. 6 cell lines that did not express GSTT1 (GM12874, GM18508, GM18912, GM19128, GM19139, and GM18517) and 5 cell lines that did express GSTT1 (GM19130, GM12145, GM12717, GM19200, and GM12155) were treated with 2 mM EB for 6 h. *GSTT1*$^{-/-}$ cells exhibited higher sensitivity towards EB-

induced apoptosis as determined by Annexin V flow cytometry assay, suggesting that GSTT1 had a protective effect against EB toxicity (Chapter 3, Figure 3.2). However, there was no statistical difference for EB-GSH conjugates or EB-GII adduct levels between *GSTT1*$^{-/-}$ and *GSTT1*$^{+/+}$ cell lines (Chapter 3, Figures 3.3 A and 3.4 A). We hypothesized that this was due to the high concentration of EB we treated with, so we repeated the cell treatment on 3 *GSTT1*$^{-/-}$ and 3 *GSTT1*$^{+/+}$ cell lines with 10 μM EB. Again, no difference in EB-GII adduct levels was seen between *GSTT1*$^{-/-}$ and *GSTT1*$^{+/+}$ cell lines. However, EB-GSH levels were higher in *GSTT1*$^{+/+}$ cells than *GSTT1*$^{-/-}$, though this was not statistically significant due to the low number of samples analyzed. The lack of difference seen between *GSTT1*$^{-/-}$ and *GSTT1*$^{+/+}$ cell lines is possibly due to the low amount of GSTT1 that is expressed in B-lymphocytes or the influence of other GSTs in the conjugation of EB. While GSTT1 accounts for the majority of EB-GSH formation, it is possible that other GSTs, such as GSTT2 and GSTM1, catalyze the reaction in the absence of GSTT1.[198, 209-210] The expression of these enzymes in the HapMap cell lines used in this study is not currently known, and it is possible that in the absence of GSTT1, these enzymes take over its role in EB-GSH conjugation.

In chapter 4 of this thesis, we identified a 4-OHBn adduct to N-terminal valine in hemoglobin (Hb) that had been previously found in adductomics screens. Hb is a highly abundant blood protein with nucleophilic sites, including the N-terminal valine, that can form covalent adducts with electrophiles. Thus, electrophile-induced Hb adducts have been used as biomarkers for electrophile exposures, which due to their reactivity are difficult to

detect before they have formed adducts.[4, 47-48] An adductomics screening methodology developed by the Törnqvist laboratory, known as the FI*R*E procedure, employs a modified Edman degradation with fluorescein isothiocyanate (FITC) to selectively detach and derivatize the adducted N-terminal amino acid of Hb, and the resulting FTH-Val derivatives are analyzed by LC-MS/MS.[89] Adductomic screens using this methodology were performed on human blood samples to detect electrophile induced adducts to N-terminal val, and 19 unknown adducts were detected. 6 of these have since then been identified.[83-85] The third most abundant N-terminal Hb adduct in human blood samples was an unknown adduct with *m*/*z* of 595.153 of the FTH derivative, which corresponds to an added mass of 106.042 Da from the adduct to Val.[83, 85] We hypothesized that the precursor of this adduct was likely either 2- or 4-quinone methide (QM), which are highly electrophilic. We synthesized authentic standards for the adducts corresponding to 2- and 4-QM, 2- and 4-OHBn-Val, using a quinone methide precursor that could be deprotected and immediately reacted with val to avoid hydrolysis of the QM (Chapter 4, Schemes 4.2 and 4.3). These standards were added to human blood samples and derivatized using the FI*R*E method before HPLC-ESI$^+$-HRMS analysis on a Q Exactive HF mass spectrometer (Thermo Scientific, Waltham, MA). Accurate mass of both 2- and 4-OHBn-Val-FTH matched the unknown adduct, but only the HPLC retention time and fragmentation patterns from 4-OHBn-Val-FTH matched the unknown, whereas the retention time and fragmentation pattern of 2-OHBn-Val-FTH matched another unknown adduct present at much lower levels. 3-OHBn-Val-FTH was synthesized and used as an internal standard for

the quantitation of background levels of the 4-OHBn-Val adducts in 12 human blood samples (380 ± 160 pmol/g Hb). In addition to 4-QM, it was found that 4-hydroxybenzaldehyde (4-OHBA) can also form 4-OHBn-Val after reduction. Further studies are needed to quantify the contributions from identified possible precursor electrophiles to the observed 4-OHBn-Val adduct seen in humans.

In summary, the research described in this thesis has employed LC-ESI-MS/MS methodologies to investigate the role of DNA repair pathways in the removal of *bis*-N7G-BD crosslinks, the relationship between *bis*-N7G-BD adduct formation, cytotoxicity, and micronuclei formation, the role of GSTT1 in EB detoxification, and the identification of a novel electrophile-induced Hb adduct. +New HPLC-ESI$^+$-MS/MS methodology for the quantitation of EB-GSH in EB exposed cells was developed. Other methodologies, including *bis*-N7G-BD and EB-GII were adapted for use on a sensitive TSQ Quantiva mass spectrometer. The methods described in this thesis can be used in future studies to investigate BD exposure in cell culture, laboratory animals, and humans.

# VI. FUTURE DIRECTIONS

## 6.1 Further investigation of the nucleotide excision repair (NER) and Fanconi Anemia (FA) repair pathways in the repair of DEB induced DNA-DNA crosslinks

In chapter 2, we investigated the formation and repair of DEB induced DNA-DNA crosslink, *bis*-N7G-BD in repair deficient human cell lines. No differences in formation and repair of *bis*-N7G-BD was seen in these cell lines versus their isogenic corrected controls, implying that the knockout of *XPA* or *FANCD2* does not impede *bis*-N7G-BD repair. However, these are just two proteins involved in the complex NER are FA repair pathways, and it is possible that these pathways could still function without them present.

XPA, Xeroderma Pigmentosa complementation group A protein, is recruited to the site of the DNA damage with transcription factor II H (TFIIH) as it unwinds the double-stranded DNA in both transcription-coupled (TC-NER) and global genomic (GG-NER) NER pathways. XPA binds to the damaged single-stranded DNA (ssDNA), while replication protein A (RPA) binds to the undamaged ssDNA. XPA and RPA act as scaffolds to recruit and properly position the endonulclease XPF-ERCC1, which makes an incision 5' to the damaged DNA.[243] As such, it is possible that in the absence of XPA, XPF-ERCC1 could still be recruited, allowing repair to continue unimpeded. To confirm whether or not NER is involved in the repair of *bis*-N7G-BD, the repair experiments in chapter 2 should be repeated in cell lines deficient in other components of the NER pathway, such as subunits of TFIIH and XPF-ERCC1. It may also be of interest to investigate the role of TC-NER vs. GG-NER in *bis*-N7G-BD by investigating repair in cell lines deficient in XPC, part of the XPC-RAD23B complex that recognizes helix distortions

leading to GG-NER. Several human fibroblast cell lines deficient in specific NER proteins are available from the Coriell Institute, examples of which are listed in Table 6.1.

FANCD2, Fanconi Anemia complementation group D2, is activated by the Fanconi Anemia complex in response to DNA damage, causing it to coordinate the recruiting of proteins which facilitate the unhooking of the ICL from DNA, allowing the DNA to be repaired. These include the nucleases FAN1 and XPF-ERCC1 and FANCS (BRCA1) and FANCD2 (BRCA2), which help promote strand invasion in homologous recombination repair. While it is unlikely that FA repair can proceed without the presence of FANCD2, there seems to be enough crossover between these two repair pathways to potentially bypass this deficiency. Thus, it could be of interest to investigate *bis*-N7G-BD repair in cell lines deficient in other components of the FA repair pathway, such as subunits of the FA complex that recognizes ICLs (FANCA, B, C, E, F, and G), and FANCS (BRCA1) and FANCD1 (BRCA2). Examples of human cell lines deficient in FA repair proteins are listed in Table 6.1.

Due to the large amount of crossover between these two repair pathways in the repair of ICLs, the lack of one protein may not be enough to affect the repair of *bis*-N7G-BD. In this case, multiple genes involved in these pathways could be silenced with short interfering RNA (siRNA). For example, this could be used to knock out all subunits of the FA complex to definitively inhibit ICL recognition by the FA pathway.

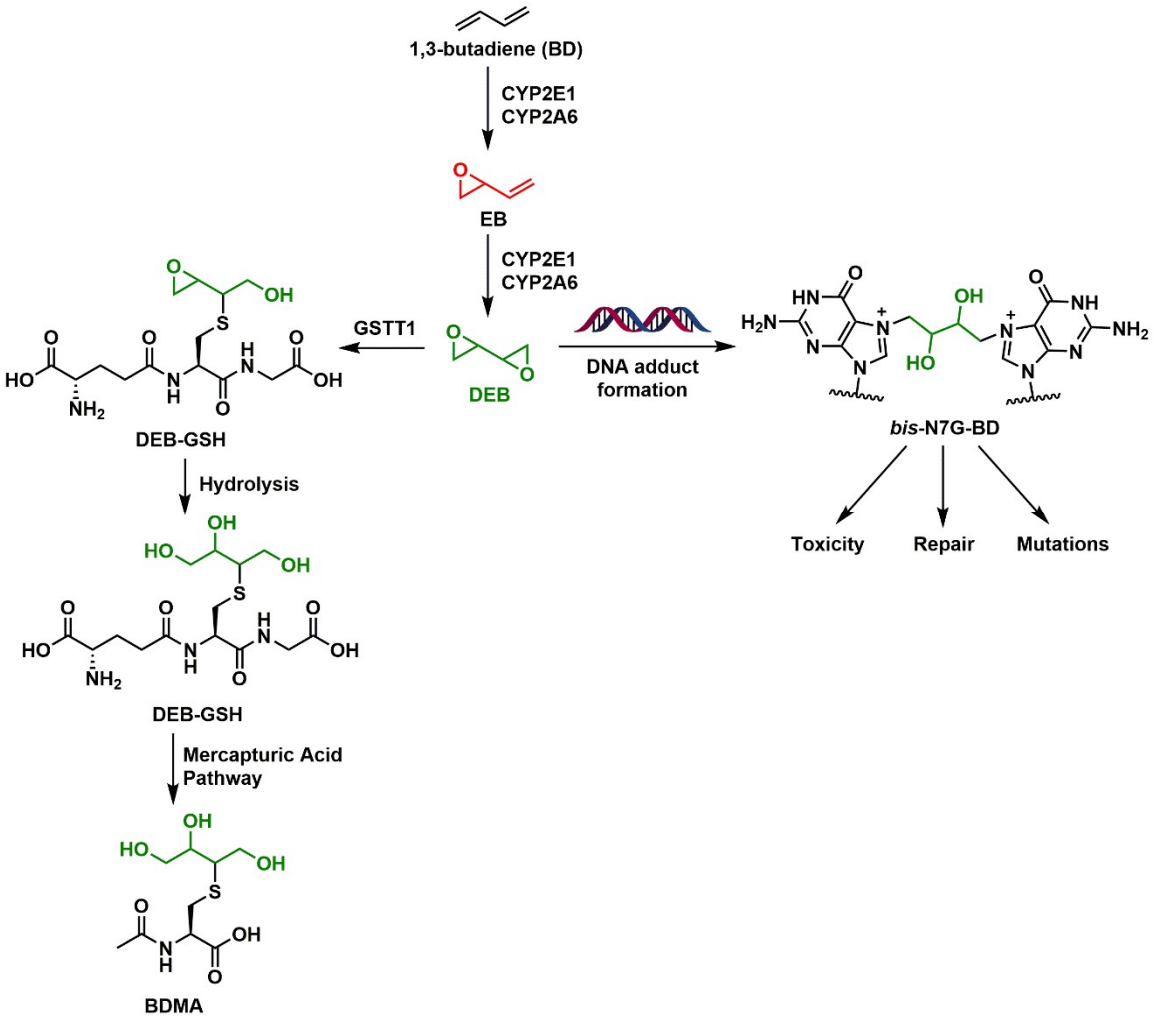
**Table 6.1** Human cell lines deficient in DNA repair pathways.

| Catalog ID | Description | Human Cell Type |
|---|---|---|
| Involved in NER only | | |
| GM03615 | XPD (subunit of TFIIH) | Fibroblast |
| GM13025 | XPB (subunit of TFIIH) | Fibroblast |
| GM03176 | XPC (subunit of XPC-RAD23B complex, damage detection in GG-NER) | Fibroblast |
| Involved in FA repair only | | |
| GM00449 | FANCC (subunit of FA complex) | Fibroblast |
| GM16632 | FANCA (subunit of FA complex) | Fibroblast |
| GM16757 | FANCF (subunit of FA complex) | B-Lymphocyte |
| GM13023 | FANCD1 (BRCA2) | B-Lymphocyte |
| Involved in both NER and FA repair | | |
| GM04313 | XPF | Fibroblast |

## 6.2 Influence of GSTT1 expression in the formation of DEB-induced DNA adducts and glutathione conjugates

In chapter 3, the relationship between GSTT1 expression, DNA adduct formation, detoxification by glutathione conjugation, and biological outcomes after 3,4-epoxy-1-butene (EB) was studied in a cell culture model. However, the presence or absence of GSTT1 did not have a statistically significant effect on adduct or conjugate formation, though GSTT1 did seem to have a slight protective effect against apoptosis after EB treatment. As stated in chapter 3, the risk of developing cigarette-induced lung cancer differs between different ethnicities. A previous genome wide association study found that genetic polymorphisms in *GSTT1* could partially account for the difference in MHBMA, a BD detoxification metabolite, in the urine of smokers of different ethnicities. One explanation for the lack of effect of GSTT1 is the relatively low toxicity of EB. EB-GII DNA adducts are readily hydrolyzed and repaired, and are not highly mutagenic. However, EB is not the only, nor is it the most biologically important, epoxide metabolite of BD. 1,2,3,4-dipoxybutane (DEB) accounts for most of BD's genotoxic and carcinogenic effects due to its ability to form DNA-DNA and DNA-protein crosslinks (Scheme 6.1). Like EB, DEB can be detoxified by GSTT1 to the glutathione conjugate DEB-GSH. Due to the increased toxicity of DEB induced adducts, it is possible that GSH conjugation by GSTT1 may play a more pronounced role in decreasing DEB induced genotoxicity than EB.

**Scheme 6.1** Metabolism and detoxification of DEB.



The experiments from chapter 3, measuring DNA adducts, GSH conjugates, and apoptosis after EB exposure in HapMap cell lines that are either null or positive for GSTT1, will be repeated using DEB (Scheme 6.1). Currently, labeled and unlabeled standards for DEB-GSH are being synthesized and purified in our lab for the development of a HPLC-$ESI^+$-MS/MS for the quantitation of DEB-GSH in cells and cellular media after DEB exposure. While EB-GSH is highly polar, DEB-GSH is even more polar. Retention of DEB-GSH on reverse phase HPLC columns has proven to be extremely difficult. A method using a hypercarb column (Thermo Scientific, Waltham, MA) has shown the most promising results for DEB-GSH retention, with retention of DEB-GSH >7 min.

### 6.3 Further investigation of sources of 4-hydroxybenzylvaline hemoglobin adducts

In chapter 4, a 4-hydroxybenzyl adduct to N-terminal valine in Hb, originally found in adductomic screens, was identified. Two possible precursors were also identified – 4-quinone methide (4-QM) and 4-hydroxybenzaldehyde (4-OHBA). However, further work must be done to confirm which of these adducts, or another precursor altogether, is the source for this adduct in humans. In chapter 4, several possible sources of 4-QM and 4-OHBA were discussed. One possible source of 4-QM is the amino acid tyrosine. Recently, a paper published by the Schöneich lab confirmed that 4-QM is generated as a byproduct of tyrosine photodegradation under UV light.[244] In order to test whether or not this could be a possible source for the 4-OHBn adduct, this photodegradation experiment should be repeated in the presence of hemoglobin. N-terminal valine Hb adducts can then be derivatized and enriched using the FI*R*E method as described in chapter 4 in order to

determine whether 4-OHBn adduct levels increase after photoirradiation. In addition, 4-OHBn adducts at nucleophilic sites in Hb other than N-terminal valine can be investigated through proteomic experiments. 4-QMP can be added to a sample of Hb and deprotected to the active 4-QM, forming 4-OHBn adducts. The Hb sample can then be incubated with trypsin in order to cleave proteins to peptides, which can be analyzed by LC-HRMS/MS. Data can be analyzed using proteomic software such as Proteome Discoverer (Thermo Scientific, Waltham, MA) in order to identify all sites within Hb where this adduct is found. Amino acid residues with nucleophilic side chains, such as lysine and cysteine, are expected to react with 4-QM to form 4-OHBn adducts. These experiments can be repeated using different concentrations of 4-QM or 4-OHBA in order to differentiate the affinities of different Hb residues for forming 4-OHBn adducts. The prevalence of 4-OHBn adducts at these sites in Hb samples that have not been treated with any precursors can be compared to the binding profiles for 4-QM and 4-OHBA, potentially implicating one, both, or neither of these precursors in the formation of 4-OHBn adducts in Hb.

## BIBLIOGRAPHY


1. Lewtas, J., Airborne carcinogens. *Pharmacol Toxicol* **1993,** *72 Suppl 1*, 55-63.
2. Abnet, C. C., Carcinogenic food contaminants. *Cancer Invest* **2007,** *25* (3), 189-96.
3. Hussain, S. P.; Harris, C. C., Molecular epidemiology and carcinogenesis: endogenous and exogenous carcinogens. *Mutation research* **2000,** *462* (2-3), 311-22.
4. Törnqvist, M.; Fred, C.; Haglund, J.; Helleberg, H.; Paulsson, B.; Rydberg, P., Protein adducts: quantitative and qualitative aspects of their formation, analysis and applications. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2002,** *778* (1-2), 279-308.
5. Tretyakova, N.; Villalta, P. W.; Kotapati, S., Mass spectrometry of structurally modified DNA. *Chem Rev* **2013,** *113* (4), 2395-436.
6. Boysen, G.; Pachkowski, B. F.; Nakamura, J.; Swenberg, J. A., The formation and biological significance of N7-guanine adducts. *Mutation research* **2009,** *678* (2), 76-94.
7. Hwa Yun, B.; Guo, J.; Bellamri, M.; Turesky, R. J., DNA adducts: Formation, biological effects, and new biospecimens for mass spectrometric measurements in humans. *Mass spectrometry reviews* **2018**.
8. Reiner, B.; Zamenhof, S., Studies on the chemically reactive groups of deoxyribonucleic acids. *The Journal of biological chemistry* **1957,** *228* (1), 475-86.
9. Rajski, S. R.; Williams, R. M., DNA Cross-Linking Agents as Antitumor Drugs. *Chem Rev* **1998,** *98* (8), 2723-2796.
10. Cochrane, J. E.; Skopek, T. R., 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* **1994,** *15* (4), 719-23.
11. Minca, E. C.; Kowalski, D., Replication fork stalling by bulky DNA damage: localization at active origins and checkpoint modulation. *Nucleic acids research* **2011,** *39* (7), 2610-23.
12. Hemminki, K., DNA adducts, mutations and cancer. *Carcinogenesis* **1993,** *14* (10), 2007-12.
13. Rundle, A., Carcinogen-DNA adducts as a biomarker for cancer risk. *Mutation research* **2006,** *600* (1-2), 23-36.
14. Essigmann, J. M.; Croy, R. G.; Nadzan, A. M.; Busby, W. F., Jr.; Reinhold, V. N.; Buchi, G.; Wogan, G. N., Structural identification of the major DNA adduct formed by aflatoxin B1 in vitro. *Proceedings of the National Academy of Sciences of the United States of America* **1977,** *74* (5), 1870-4.
15. Qian, G. S.; Ross, R. K.; Yu, M. C.; Yuan, J. M.; Gao, Y. T.; Henderson, B. E.; Wogan, G. N.; Groopman, J. D., A follow-up study of urinary markers of aflatoxin exposure and liver cancer risk in Shanghai, People's Republic of China. *Cancer Epidemiol Biomarkers Prev* **1994,** *3* (1), 3-10.

16. Koc, H.; Tretyakova, N. Y.; Walker, V. E.; Henderson, R. F.; Swenberg, J. A., Molecular dosimetry of N-7 guanine adduct formation in mice and rats exposed to 1,3-butadiene. *Chemical research in toxicology* **1999,** *12* (7), 566-74.
17. Zhang, S.; Balbo, S.; Wang, M.; Hecht, S. S., Analysis of acrolein-derived 1,N2-propanodeoxyguanosine adducts in human leukocyte DNA from smokers and nonsmokers. *Chemical research in toxicology* **2011,** *24* (1), 119-24.
18. Geacintov, N. E.; Yoshida, H.; Ibanez, V.; Jacobs, S. A.; Harvey, R. G., Conformations of adducts and kinetics of binding to DNA of the optically pure enantiomers of anti-benzo(a)pyrene diol epoxide. *Biochem Biophys Res Commun* **1984,** *122* (1), 33-9.
19. Villalta, P. W.; Hochalter, J. B.; Hecht, S. S., Ultrasensitive High-Resolution Mass Spectrometric Analysis of a DNA Adduct of the Carcinogen Benzo[a]pyrene in Human Lung. *Analytical chemistry* **2017,** *89* (23), 12735-12742.
20. Peterson, L. A., Formation, repair, and genotoxic properties of bulky DNA adducts formed from tobacco-specific nitrosamines. *J Nucleic Acids* **2010,** *2010*.
21. Balbo, S.; Brooks, P. J., Implications of acetaldehyde-derived DNA adducts for understanding alcohol-related carcinogenesis. *Adv Exp Med Biol* **2015,** *815*, 71-88.
22. Hemminki, K., DNA-binding products of nornitrogen mustard, a metabolite of cyclophosphamide. *Chemico-biological interactions* **1987,** *61* (1), 75-88.
23. Hall, A. G.; Tilby, M. J., Mechanisms of action of, and modes of resistance to, alkylating agents used in the treatment of haematological malignancies. *Blood Rev* **1992,** *6* (3), 163-73.
24. Dasari, S.; Tchounwou, P. B., Cisplatin in cancer therapy: molecular mechanisms of action. *Eur J Pharmacol* **2014,** *740*, 364-78.
25. Malayappan, B.; Johnson, L.; Nie, B.; Panchal, D.; Matter, B.; Jacobson, P.; Tretyakova, N., Quantitative high-performance liquid chromatography-electrospray ionization tandem mass spectrometry analysis of bis-N7-guanine DNA-DNA cross-links in white blood cells of cancer patients receiving cyclophosphamide therapy. *Analytical chemistry* **2010,** *82* (9), 3650-8.
26. Fichtinger-Schepman, A. M.; van der Veer, J. L.; den Hartog, J. H.; Lohman, P. H.; Reedijk, J., Adducts of the antitumor drug cis-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. *Biochemistry* **1985,** *24* (3), 707-13.
27. Perera, F. P.; Motzer, R. J.; Tang, D.; Reed, E.; Parker, R.; Warburton, D.; O'Neill, P.; Albertini, R.; Bigbee, W. L.; Jensen, R. H.; et al., Multiple biological markers in germ cell tumor patients treated with platinum-based chemotherapy. *Cancer research* **1992,** *52* (13), 3558-65.
28. Guengerich, F. P., Metabolism of chemical carcinogens. *Carcinogenesis* **2000,** *21* (3), 345-51.
29. Sancar, A.; Lindsey-Boltz, L. A.; Unsal-Kacmaz, K.; Linn, S., Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem* **2004,** *73*, 39-85.
30. Coste, F.; Ober, M.; Le Bihan, Y. V.; Izquierdo, M. A.; Hervouet, N.; Mueller, H.; Carell, T.; Castaing, B., Bacterial base excision repair enzyme Fpg recognizes bulky N7-

substituted-FapydG lesion via unproductive binding mode. *Chem Biol* **2008,** *15* (7), 706-17.
31. Lindahl, T., An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues. *Proceedings of the National Academy of Sciences of the United States of America* **1974,** *71* (9), 3649-53.
32. Preston, T. J.; Henderson, J. T.; McCallum, G. P.; Wells, P. G., Base excision repair of reactive oxygen species-initiated 7,8-dihydro-8-oxo-2'-deoxyguanosine inhibits the cytotoxicity of platinum anticancer drugs. *Mol Cancer Ther* **2009,** *8* (7), 2015-26.
33. Robertson, A. B.; Klungland, A.; Rognes, T.; Leiros, I., DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci* **2009,** *66* (6), 981-93.
34. Srivastava, D. K.; Berg, B. J.; Prasad, R.; Molina, J. T.; Beard, W. A.; Tomkinson, A. E.; Wilson, S. H., Mammalian abasic site base excision repair. Identification of the reaction sequence and rate-determining steps. *The Journal of biological chemistry* **1998,** *273* (33), 21203-9.
35. Spivak, G., Nucleotide excision repair in humans. *DNA repair* **2015,** *36*, 13-18.
36. Sancar, A., Excision repair in mammalian cells. *The Journal of biological chemistry* **1995,** *270* (27), 15915-8.
37. You, J. S.; Wang, M.; Lee, S. H., Biochemical analysis of the damage recognition process in nucleotide excision repair. *The Journal of biological chemistry* **2003,** *278* (9), 7476-85.
38. Hanawalt, P. C., Subpathways of nucleotide excision repair and their regulation. *Oncogene* **2002,** *21* (58), 8949-56.
39. Zhang, N.; Liu, X.; Li, L.; Legerski, R., Double-strand breaks induce homologous recombinational repair of interstrand cross-links via cooperation of MSH2, ERCC1-XPF, REV3, and the Fanconi anemia pathway. *DNA repair* **2007,** *6* (11), 1670-8.
40. Marcon, E.; Moens, P. B., The evolution of meiosis: recruitment and modification of somatic DNA-repair proteins. *Bioessays* **2005,** *27* (8), 795-808.
41. Chang, H. H. Y.; Pannunzio, N. R.; Adachi, N.; Lieber, M. R., Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* **2017,** *18* (8), 495-506.
42. Peng, G.; Lin, S. Y., Exploiting the homologous recombination DNA repair network for targeted cancer therapy. *World J Clin Oncol* **2011,** *2* (2), 73-9.
43. Moldovan, G. L.; D'Andrea, A. D., How the fanconi anemia pathway guards the genome. *Annu Rev Genet* **2009,** *43*, 223-49.
44. Dunn, J.; Potter, M.; Rees, A.; Runger, T. M., Activation of the Fanconi anemia/BRCA pathway and recombination repair in the cellular response to solar ultraviolet light. *Cancer research* **2006,** *66* (23), 11140-7.
45. Auerbach, A. D., Fanconi anemia and its diagnosis. *Mutation research* **2009,** *668* (1-2), 4-10.
46. Alter, B. P.; Greene, M. H.; Velazquez, I.; Rosenberg, P. S., Cancer in Fanconi anemia. In *Blood*, United States, 2003; Vol. 101, p 2072.

47. Rubino, F. M.; Pitton, M.; Di Fabio, D.; Colombi, A., Toward an "omic" physiopathology of reactive chemicals: thirty years of mass spectrometric study of the protein adducts with endogenous and xenobiotic compounds. *Mass spectrometry reviews* **2009,** *28* (5), 725-84.
48. Sabbioni, G.; Turesky, R. J., Biomonitoring Human Albumin Adducts: The Past, the Present, and the Future. - PubMed - NCBI. *Chemical research in toxicology* **2017,** *30* (1), 332-366.
49. Preston, G. W.; Phillips, D. H., Protein Adductomics: Analytical Developments and Applications in Human Biomonitoring. *Toxics* **2019,** *7* (2).
50. Scheepers, P. T., The use of biomarkers for improved retrospective exposure assessment in epidemiological studies: summary of an ECETOC workshop. In *Biomarkers*, England, 2008; Vol. 13, pp 734-48.
51. Osterman-Golkar, S. M.; Moss, O.; James, A.; Bryant, M. S.; Turner, M.; Bond, J. A., Epoxybutene-hemoglobin adducts in rats and mice: dose response for formation and persistence during and following long-term low-level exposure to butadiene. *Toxicol Appl Pharmacol* **1998,** *150* (1), 166-73.
52. Balbo, S.; Turesky, R. J.; Villalta, P. W., DNA adductomics. *Chemical research in toxicology* **2014,** *27* (3), 356-66.
53. Liu, S.; Wang, Y., Mass spectrometry for the assessment of the occurrence and biological consequences of DNA adducts. *Chem Soc Rev* **2015,** *44* (21), 7829-54.
54. Tretyakova, N.; Goggin, M.; Sangaraju, D.; Janis, G., Quantitation of DNA adducts by stable isotope dilution mass spectrometry. *Chemical research in toxicology* **2012,** *25* (10), 2007-35.
55. Ciccimaro, E.; Blair, I. A., Stable-isotope dilution LC-MS for quantitative biomarker analysis. *Bioanalysis* **2010,** *2* (2), 311-41.
56. Chen, L.; Wang, M.; Villalta, P. W.; Luo, X.; Feuer, R.; Jensen, J.; Hatsukami, D. K.; Hecht, S. S., Quantitation of an acetaldehyde adduct in human leukocyte DNA and the effect of smoking cessation. *Chemical research in toxicology* **2007,** *20* (1), 108-13.
57. Stepanov, I.; Muzic, J.; Le, C. T.; Sebero, E.; Villalta, P.; Ma, B.; Jensen, J.; Hatsukami, D.; Hecht, S. S., Analysis of 4-hydroxy-1-(3-pyridyl)-1-butanone (HPB)-releasing DNA adducts in human exfoliated oral mucosa cells by liquid chromatography-electrospray ionization-tandem mass spectrometry. *Chemical research in toxicology* **2013,** *26* (1), 37-45.
58. Talaska, G.; Schamer, M.; Skipper, P.; Tannenbaum, S.; Caporaso, N.; Unruh, L.; Kadlubar, F. F.; Bartsch, H.; Malaveille, C.; Vineis, P., Detection of carcinogen-DNA adducts in exfoliated urothelial cells of cigarette smokers: association with smoking, hemoglobin adducts, and urinary mutagenicity. *Cancer Epidemiol Biomarkers Prev* **1991,** *1* (1), 61-6.
59. Goggin, M.; Loeber, R.; Park, S.; Walker, V.; Wickliffe, J.; Tretyakova, N., HPLC-ESI+-MS/MS analysis of N7-guanine-N7-guanine DNA cross-links in tissues of mice exposed to 1,3-butadiene. *Chemical research in toxicology* **2007,** *20* (5), 839-47.

60. Tarun, M.; Rusling, J. F., Measuring DNA nucleobase adducts using neutral hydrolysis and liquid chromatography-mass spectrometry. *Crit Rev Eukaryot Gene Expr* **2005,** *15* (4), 295-316.
61. Upadhyaya, P.; Lindgren, B. R.; Hecht, S. S., Comparative levels of O6-methylguanine, pyridyloxobutyl-, and pyridylhydroxybutyl-DNA adducts in lung and liver of rats treated chronically with the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Drug Metab Dispos* **2009,** *37* (6), 1147-51.
62. Goggin, M.; Seneviratne, U.; Swenberg, J. A.; Walker, V. E.; Tretyakova, N., 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. *Chemical research in toxicology* **2010,** *23* (4), 808-12.
63. Schumacher, F.; Herrmann, K.; Florian, S.; Engst, W.; Glatt, H., Optimized enzymatic hydrolysis of DNA for LC-MS/MS analyses of adducts of 1-methoxy-3-indolylmethyl glucosinolate and methyleugenol. *Anal Biochem* **2013,** *434* (1), 4-11.
64. Koc, H.; Swenberg, J. A., Applications of mass spectrometry for quantitation of DNA adducts. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2002,** *778* (1-2), 323-43.
65. Sangaraju, D.; Boldry, E. J.; Patel, Y. M.; Walker, V.; Stepanov, I.; Stram, D.; Hatsukami, D.; Tretyakova, N., Isotope Dilution nanoLC/ESI(+)-HRMS(3) Quantitation of Urinary N7-(1-Hydroxy-3-buten-2-yl) Guanine Adducts in Humans and Their Use as Biomarkers of Exposure to 1,3-Butadiene. *Chemical research in toxicology* **2017,** *30* (2), 678-688.
66. Sangaraju, D.; Villalta, P.; Goggin, M.; Agunsoye, M. O.; Campbell, C.; Tretyakova, N., Capillary HPLC-accurate mass MS/MS quantitation of N7-(2,3,4-trihydroxybut-1-yl)-guanine adducts of 1,3-butadiene in human leukocyte DNA. *Chemical research in toxicology* **2013,** *26* (10), 1486-97.
67. Sangaraju, D.; Villalta, P. W.; Wickramaratne, S.; Swenberg, J.; Tretyakova, N., NanoLC/ESI+ HRMS3 quantitation of DNA adducts induced by 1,3-butadiene. *J Am Soc Mass Spectrom* **2014,** *25* (7), 1124-35.
68. Chesner, L. N.; Degner, A.; Sangaraju, D.; Yomtoubian, S.; Wickramaratne, S.; Malayappan, B.; Tretyakova, N.; Campbell, C., Cellular Repair of DNA-DNA Cross-Links Induced by 1,2,3,4-Diepoxybutane. *Int J Mol Sci* **2017,** *18* (5).
69. Page, J. S.; Kelly, R. T.; Tang, K.; Smith, R. D., Ionization and transmission efficiency in an electrospray ionization-mass spectrometry interface. *J Am Soc Mass Spectrom* **2007,** *18* (9), 1582-90.
70. Ho, C. S.; Lam, C. W.; Chan, M. H.; Cheung, R. C.; Law, L. K.; Lit, L. C.; Ng, K. F.; Suen, M. W.; Tai, H. L., Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin Biochem Rev* **2003,** *24* (1), 3-12.
71. Sangaraju, D.; Goggin, M.; Walker, V.; Swenberg, J.; Tretyakova, N., NanoHPLC-nanoESI(+)-MS/MS quantitation of bis-N7-guanine DNA-DNA cross-links in tissues of B6C3F1 mice exposed to subppm levels of 1,3-butadiene. *Analytical chemistry* **2012,** *84* (3), 1732-9.

72. Horgan, R. P.; Kenny, L. C., 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics - Horgan - 2011 - The Obstetrician & Gynaecologist - Wiley Online Library. *Obstet Gynaecol* **2011,** *13*, 189-195.
73. Wild, C. P., Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. In *Cancer Epidemiol Biomarkers Prev*, United States, 2005; Vol. 14, pp 1847-50.
74. Wild, C. P., The exposome: from concept to utility. *Int J Epidemiol* **2012,** *41* (1), 24-32.
75. Manolio, T. A.; Collins, F. S.; Cox, N. J.; Goldstein, D. B.; Hindorff, L. A.; Hunter, D. J.; McCarthy, M. I.; Ramos, E. M.; Cardon, L. R.; Chakravarti, A.; Cho, J. H.; Guttmacher, A. E.; Kong, A.; Kruglyak, L.; Mardis, E.; Rotimi, C. N.; Slatkin, M.; Valle, D.; Whittemore, A. S.; Boehnke, M.; Clark, A. G.; Eichler, E. E.; Gibson, G.; Haines, J. L.; Mackay, T. F.; McCarroll, S. A.; Visscher, P. M., Finding the missing heritability of complex diseases. *Nature* **2009,** *461* (7265), 747-53.
76. Lichtenstein, P.; Holm, N. V.; Verkasalo, P. K.; Iliadou, A.; Kaprio, J.; Koskenvuo, M.; Pukkala, E.; Skytthe, A.; Hemminki, K., Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *The New England journal of medicine* **2000,** *343* (2), 78-85.
77. Rappaport, S. M., Genetic Factors Are Not the Major Causes of Chronic Diseases. *PloS one* **2016,** *11* (4), e0154387.
78. Carlsson, H.; Rappaport, S. M.; Tornqvist, M., Protein Adductomics: Methodologies for Untargeted Screening of Adducts to Serum Albumin and Hemoglobin in Human Blood Samples. *High Throughput* **2019,** *8* (1).
79. Grigoryan, H.; Edmands, W.; Lu, S. S.; Yano, Y.; Regazzoni, L.; Iavarone, A. T.; Williams, E. R.; Rappaport, S. M., Adductomics Pipeline for Untargeted Analysis of Modifications to Cys34 of Human Serum Albumin. *Analytical chemistry* **2016,** *88* (21), 10504-10512.
80. Lu, S. S.; Grigoryan, H.; Edmands, W. M.; Hu, W.; Iavarone, A. T.; Hubbard, A.; Rothman, N.; Vermeulen, R.; Lan, Q.; Rappaport, S. M., Profiling the Serum Albumin Cys34 Adductome of Solid Fuel Users in Xuanwei and Fuyuan, China. *Environ Sci Technol* **2017,** *51* (1), 46-57.
81. Liu, S.; Grigoryan, H.; Edmands, W. M. B.; Dagnino, S.; Sinharay, R.; Cullinan, P.; Collins, P.; Chung, K. F.; Barratt, B.; Kelly, F. J.; Vineis, P.; Rappaport, S. M., Cys34 Adductomes Differ between Patients with Chronic Lung or Heart Disease and Healthy Controls in Central London. *Environ Sci Technol* **2018,** *52* (4), 2307-2313.
82. Grigoryan, H.; Edmands, W. M. B.; Lan, Q.; Carlsson, H.; Vermeulen, R.; Zhang, L.; Yin, S. N.; Li, G. L.; Smith, M. T.; Rothman, N.; Rappaport, S. M., Adductomic signatures of benzene exposure provide insights into cancer induction. *Carcinogenesis* **2018,** *39* (5), 661-668.
83. Carlsson, H.; von Stedingk, H.; Nilsson, U.; Törnqvist, M., LC-MS/MS screening strategy for unknown adducts to N-terminal valine in hemoglobin applied to smokers and nonsmokers. *Chemical research in toxicology* **2014,** *27* (12), 2062-70.

84. Carlsson, H.; Motwani, H. V.; Osterman Golkar, S.; Törnqvist, M., Characterization of a Hemoglobin Adduct from Ethyl Vinyl Ketone Detected in Human Blood Samples. *Chemical research in toxicology* **2015,** *28* (11), 2120-9.
85. Carlsson, H.; Törnqvist, M., Strategy for identifying unknown hemoglobin adducts using adductome LC-MS/MS data: Identification of adducts corresponding to acrylic acid, glyoxal, methylglyoxal, and 1-octen-3-one. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* **2016,** *92*, 94-103.
86. Degner, A.; Carlsson, H.; Karlsson, I.; Eriksson, J.; Pujari, S. S.; Tretyakova, N. Y.; Tornqvist, M., Discovery of Novel N-(4-Hydroxybenzyl)valine Hemoglobin Adducts in Human Blood. *Chemical research in toxicology* **2018,** *31* (12), 1305-1314.
87. Tornqvist, M.; Mowrer, J.; Jensen, S.; Ehrenberg, L., Monitoring of environmental cancer initiators through hemoglobin adducts by a modified Edman degradation method. *Anal Biochem* **1986,** *154* (1), 255-66.
88. Rydberg, P.; Luning, B.; Wachtmeister, C. A.; Eriksson, L.; Tornqvist, M., Applicability of a modified Edman procedure for measurement of protein adducts: mechanisms of formation and degradation of phenylthiohydantoins. *Chemical research in toxicology* **2002,** *15* (4), 570-81.
89. von Stedingk, H.; Rydberg, P.; Törnqvist, M., A new modified Edman procedure for analysis of N-terminal valine adducts in hemoglobin by LC-MS/MS. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2010,** *878* (27), 2483-90.
90. Carlsson, H.; Aasa, J.; Kotova, N.; Vare, D.; Sousa, P. F. M.; Rydberg, P.; Abramsson-Zetterberg, L.; Törnqvist, M., Adductomic Screening of Hemoglobin Adducts and Monitoring of Micronuclei in School-Age Children. *Chemical research in toxicology* **2017,** *30* (5), 1157-1167.
91. *IARC Monographs on the Evaluation of Carcinogenic Risk to Humans, No. 100F. Chemical Agents and Related Occupations*. IARC: Lyon, France, 2012.
92. White, W. C., Butadiene production process overview. *Chemico-biological interactions* **2007,** *166* (1-3), 10-4.
93. Kagawa, J., Health effects of diesel exhaust emissions--a mixture of air pollutants of worldwide concern. *Toxicology* **2002,** *181-182*, 349-53.
94. Brunnemann, K. D.; Kagan, M. R.; Cox, J. E.; Hoffmann, D., Analysis of 1,3-butadiene and other selected gas-phase components in cigarette mainstream and sidestream smoke by gas chromatography-mass selective detection. *Carcinogenesis* **1990,** *11* (10), 1863-8.
95. Himmelstein, M. W.; Acquavella, J. F.; Recio, L.; Medinsky, M. A.; Bond, J. A., Toxicology and epidemiology of 1,3-butadiene. *Crit Rev Toxicol* **1997,** *27* (1), 1-108.
96. Melnick, R. L.; Huff, J.; Chou, B. J.; Miller, R. A., Carcinogenicity of 1,3-butadiene in C57BL/6 x C3H F1 mice at low exposure concentrations. *Cancer research* **1990,** *50* (20), 6592-9.

97. Owen, P. E.; Glaister, J. R.; Gaunt, I. F.; Pullinger, D. H., Inhalation toxicity studies with 1,3-butadiene. 3. Two year toxicity/carcinogenicity study in rats. *Am Ind Hyg Assoc J* **1987,** *48* (5), 407-13.
98. Delzell, E.; Sathiakumar, N.; Hovinga, M.; Macaluso, M.; Julian, J.; Larson, R.; Cole, P.; Muir, D. C., A follow-up study of synthetic rubber workers. *Toxicology* **1996,** *113* (1-3), 182-9.
99. Sathiakumar, N.; Graff, J.; Macaluso, M.; Maldonado, G.; Matthews, R.; Delzell, E., An updated study of mortality among North American synthetic rubber industry workers. *Occup Environ Med* **2005,** *62* (12), 822-9.
100. Duescher, R. J.; Elfarra, A. A., Human liver microsomes are efficient catalysts of 1,3-butadiene oxidation: evidence for major roles by cytochromes P450 2A6 and 2E1. *Arch Biochem Biophys* **1994,** *311* (2), 342-9.
101. Krause, R. J.; Sharer, J. E.; Elfarra, A. A., Epoxide hydrolase-dependent metabolism of butadiene monoxide to 3-butene-1,2-diol in mouse, rat, and human liver. *Drug Metab Dispos* **1997,** *25* (8), 1013-5.
102. Csanady, G. A.; Guengerich, F. P.; Bond, J. A., Comparison of the biotransformation of 1,3-butadiene and its metabolite, butadiene monoepoxide, by hepatic and pulmonary tissues from humans, rats and mice. *Carcinogenesis* **1992,** *13* (7), 1143-53.
103. Cochrane, J. E.; Skopek, T. R., 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* **1994,** *15* (4), 713-7.
104. Recio, L.; Steen, A. M.; Pluta, L. J.; Meyer, K. G.; Saranko, C. J., Mutational spectrum of 1,3-butadiene and metabolites 1,2-epoxybutene and 1,2,3,4-diepoxybutane to assess mutagenic mechanisms. *Chemico-biological interactions* **2001,** *135-136*, 325-41.
105. Loecken, E. M.; Guengerich, F. P., Reactions of glyceraldehyde 3-phosphate dehydrogenase sulfhydryl groups with bis-electrophiles produce DNA-protein cross-links but not mutations. *Chemical research in toxicology* **2008,** *21* (2), 453-8.
106. Groehler, A. t.; Degner, A.; Tretyakova, N. Y., Mass Spectrometry-Based Tools to Characterize DNA-Protein Cross-Linking by Bis-Electrophiles. *Basic & clinical pharmacology & toxicology* **2017,** *121 Suppl 3*, 63-77.
107. Tretyakova, N.; Chiang, S. Y.; Walker, V. E.; Swenberg, J. A., Quantitative analysis of 1,3-butadiene-induced DNA adducts in vivo and in vitro using liquid chromatography electrospray ionization tandem mass spectrometry. *J Mass Spectrom* **1998,** *33* (4), 363-76.
108. Tretyakova, N.; Lin, Y. P.; Upton, P. B.; Sangaiah, R.; Swenberg, J. A., Macromolecular adducts of butadiene. *Toxicology* **1996,** *113* (1-3), 70-6.
109. Park, S.; Anderson, C.; Loeber, R.; Seetharaman, M.; Jones, R.; Tretyakova, N., Interstrand and intrastrand DNA-DNA cross-linking by 1,2,3,4-diepoxybutane: role of stereochemistry. *Journal of the American Chemical Society* **2005,** *127* (41), 14355-65.

110. Park, S.; Hodge, J.; Anderson, C.; Tretyakova, N., Guanine-adenine DNA cross-linking by 1,2,3,4-diepoxybutane: potential basis for biological activity. *Chemical research in toxicology* **2004,** *17* (12), 1638-51.
111. Seneviratne, U.; Antsypovich, S.; Goggin, M.; Dorr, D. Q.; Guza, R.; Moser, A.; Thompson, C.; York, D. M.; Tretyakova, N., Exocyclic deoxyadenosine adducts of 1,2,3,4-diepoxybutane: synthesis, structural elucidation, and mechanistic studies. *Chemical research in toxicology* **2010,** *23* (1), 118-33.
112. Koivisto, P.; Adler, I. D.; Sorsa, M.; Peltonen, K., Inhalation exposure of rats and mice to 1,3-butadiene induces N6-adenine adducts of epoxybutene detected by 32P-postlabeling and HPLC. *Environ Health Perspect* **1996,** *104 Suppl 3*, 655-7.
113. Koivisto, P.; Adler, I. D.; Pacchierotti, F.; Peltonen, K., DNA adducts in mouse testis and lung after inhalation exposure to 1,3-butadiene. *Mutation research* **1998,** *397* (1), 3-10.
114. Koivisto, P.; Kilpelainen, I.; Rasanen, I.; Adler, I. D.; Pacchierotti, F.; Peltonen, K., Butadiene diolepoxide- and diepoxybutane-derived DNA adducts at N7-guanine: a high occurrence of diolepoxide-derived adducts in mouse lung after 1,3-butadiene exposure. *Carcinogenesis* **1999,** *20* (7), 1253-9.
115. Goggin, M.; Anderson, C.; Park, S.; Swenberg, J.; Walker, V.; Tretyakova, N., 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. *Chemical research in toxicology* **2008,** *21* (5), 1163-70.
116. Goggin, M.; Swenberg, J. A.; Walker, V. E.; Tretyakova, N., 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 research* **2009,** *69* (6), 2479-86.
117. Siegel, R. L.; Miller, K. D.; Jemal, A., Cancer statistics, 2019. *CA Cancer J Clin* **2019,** *69* (1), 7-34.
118. American Cancer Society Facts & Figures 2019. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2019.html# (accessed July 18th, 2019).
119. Fowles, J.; Dybing, E., Application of toxicological risk assessment principles to the chemical constituents of cigarette smoke. *Tob Control* **2003,** *12* (4), 424-30.
120. Kirman, C. R.; Albertini, R. J.; Sweeney, L. M.; Gargas, M. L., 1,3-Butadiene: I. Review of metabolism and the implications to human health risk assessment. *Crit Rev Toxicol* **2010,** *40 Suppl 1*, 1-11.
121. Swenberg, J. A.; Bordeerat, N. K.; Boysen, G.; Carro, S.; Georgieva, N. I.; Nakamura, J.; Troutman, J. M.; Upton, P. B.; Albertini, R. J.; Vacek, P. M.; Walker, V. E.; Sram, R. J.; Goggin, M.; Tretyakova, N., 1,3-Butadiene: Biomarkers and application to risk assessment. *Chemico-biological interactions* **2011,** *192* (1-2), 150-4.
122. Tretyakova, N.; Lin, Y.; Sangaiah, R.; Upton, P. B.; Swenberg, J. A., Identification and quantitation of DNA adducts from calf thymus DNA exposed to 3,4-epoxy-1-butene. *Carcinogenesis* **1997,** *18* (1), 137-47.

123. Park, S.; Tretyakova, N., Structural characterization of the major DNA-DNA cross-link of 1,2,3,4-diepoxybutane. *Chemical research in toxicology* **2004,** *17* (2), 129-36.
124. Lawley, P. D.; Brookes, P., Interstrand cross-linking of DNA by difunctional alkylating agents. *J Mol Biol* **1967,** *25* (1), 143-60.
125. Goggin, M.; Sangaraju, D.; Walker, V. E.; Wickliffe, J.; Swenberg, J. A.; Tretyakova, N., Persistence and repair of bifunctional DNA adducts in tissues of laboratory animals exposed to 1,3-butadiene by inhalation. *Chemical research in toxicology* **2011,** *24* (6), 809-17.
126. Adler, I. D.; Filser, J.; Gonda, H.; Schriever-Schwemmer, G., Dose response study for 1,3-butadiene-induced dominant lethal mutations and heritable translocations in germs cells of male mice. *Mutation research* **1998,** *397* (1), 85-92.
127. Cunningham, M. J.; Choy, W. N.; Arce, G. T.; Rickard, L. B.; Vlachos, D. A.; Kinney, L. A.; Sarrif, A. M., In vivo sister chromatid exchange and micronucleus induction studies with 1,3-butadiene in B6C3F1 mice and Sprague-Dawley rats. *Mutagenesis* **1986,** *1* (6), 449-52.
128. Jacobson-Kram, D.; Rosenthal, S. L., Molecular and genetic toxicology of 1,3-butadiene. *Mutation research* **1995,** *339* (2), 121-30.
129. Cheng, H.; Sathiakumar, N.; Graff, J.; Matthews, R.; Delzell, E., 1,3-Butadiene and leukemia among synthetic rubber industry workers: exposure-response relationships. *Chemico-biological interactions* **2007,** *166* (1-3), 15-24.
130. Delzell, E.; Sathiakumar, N.; Graff, J.; Macaluso, M.; Maldonado, G.; Matthews, R., An updated study of mortality among North American synthetic rubber industry workers. *Res Rep Health Eff Inst* **2006,** (132), 1-63; discussion 65-74.
131. Group, I. W., 1,3-butadiene, ethylene oxide and vinyl halides (vinyl fluoride, vinyl chloride and vinyl bromide). *IARC Monogr Eval Carcinog Risks Hum 97* **2008**, 3-471.
132. Barnes, D. J.; Melo, J. V., Cytogenetic and molecular genetic aspects of chronic myeloid leukaemia. *Acta Haematol* **2002,** *108* (4), 180-202.
133. Albertini, R. J.; Carter, E. W.; Nicklas, J. A.; Vacek, P. M.; Walker, V. E., 1,3-Butadiene, CML and the t(9:22) translocation: A reality check. *Chemico-biological interactions* **2015,** *241*, 32-9.
134. Collins, S. J.; Howard, M.; Andrews, D. F.; Agura, E.; Radich, J., Rare occurrence of N-ras point mutations in Philadelphia chromosome positive chronic myeloid leukemia. *Blood* **1989,** *73* (4), 1028-32.
135. Kantarjian, H. M.; Giles, F.; Quintas-Cardama, A.; Cortes, J., Important therapeutic targets in chronic myelogenous leukemia. *Clin Cancer Res* **2007,** *13* (4), 1089-97.
136. Danisz, K.; Blasiak, J., Role of anti-apoptotic pathways activated by BCR/ABL in the resistance of chronic myeloid leukemia cells to tyrosine kinase inhibitors. *Acta Biochim Pol* **2013,** *60* (4), 503-14.
137. Iqbal, N., Imatinib: a breakthrough of targeted therapy in cancer. *Chemother Res Pract* **2014,** *2014*, 357027.
138. Trela, E.; Glowacki, S.; Blasiak, J., Therapy of chronic myeloid leukemia: twilight of the imatinib era? *ISRN Oncol* **2014,** *2014*, 596483.

139. Lichtman, M. A., Is there an entity of chemically induced BCR-ABL-positive chronic myelogenous leukemia? *Oncologist* **2008,** *13* (6), 645-54.
140. Deininger, M. W.; Bose, S.; Gora-Tybor, J.; Yan, X. H.; Goldman, J. M.; Melo, J. V., Selective induction of leukemia-associated fusion genes by high-dose ionizing radiation. *Cancer research* **1998,** *58* (3), 421-5.
141. Ito, T.; Seyama, T.; Mizuno, T.; Hayashi, T.; Iwamoto, K. S.; Dohi, K.; Nakamura, N.; Akiyama, M., Induction of BCR-ABL fusion genes by in vitro X-irradiation. *Jpn J Cancer Res* **1993,** *84* (2), 105-9.
142. Spencer, A.; Granter, N., Leukemia patient-derived lymphoblastoid cell lines exhibit increased induction of leukemia-associated transcripts following high-dose irradiation. *Exp Hematol* **1999,** *27* (9), 1397-401.
143. Deans, A. J.; West, S. C., DNA interstrand crosslink repair and cancer. *Nature reviews. Cancer* **2011,** *11* (7), 467-80.
144. Cole, R. S., Repair of DNA containing interstrand crosslinks in Escherichia coli: sequential excision and recombination. *Proceedings of the National Academy of Sciences of the United States of America* **1973,** *70* (4), 1064-8.
145. Dronkert, M. L.; Kanaar, R., Repair of DNA interstrand cross-links. *Mutation research* **2001,** *486* (4), 217-47.
146. Lehoczky, P.; McHugh, P. J.; Chovanec, M., DNA interstrand cross-link repair in Saccharomyces cerevisiae. *FEMS microbiology reviews* **2007,** *31* (2), 109-33.
147. Muniandy, P. A.; Thapa, D.; Thazhathveetil, A. K.; Liu, S. T.; Seidman, M. M., Repair of laser-localized DNA interstrand cross-links in G1 phase mammalian cells. *The Journal of biological chemistry* **2009,** *284* (41), 27908-17.
148. Enoiu, M.; Jiricny, J.; Scharer, O. D., Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic acids research* **2012,** *40* (18), 8953-64.
149. Dardalhon, M.; Averbeck, D., Pulsed-field gel electrophoresis analysis of the repair of psoralen plus UVA induced DNA photoadducts in Saccharomyces cerevisiae. *Mutation research* **1995,** *336* (1), 49-60.
150. Magana-Schwencke, N.; Henriques, J. A.; Chanet, R.; Moustacchi, E., The fate of 8-methoxypsoralen photoinduced crosslinks in nuclear and mitochondrial yeast DNA: comparison of wild-type and repair-deficient strains. *Proceedings of the National Academy of Sciences of the United States of America* **1982,** *79* (6), 1722-6.
151. Saffran, W. A.; Ahmed, S.; Bellevue, S.; Pereira, G.; Patrick, T.; Sanchez, W.; Thomas, S.; Alberti, M.; Hearst, J. E., DNA repair defects channel interstrand DNA cross-links into alternate recombinational and error-prone repair pathways. *The Journal of biological chemistry* **2004,** *279* (35), 36462-9.
152. De Silva, I. U.; McHugh, P. J.; Clingen, P. H.; Hartley, J. A., Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. *Molecular and cellular biology* **2000,** *20* (21), 7980-90.
153. Niedernhofer, L. J.; Odijk, H.; Budzowska, M.; van Drunen, E.; Maas, A.; Theil, A. F.; de Wit, J.; Jaspers, N. G.; Beverloo, H. B.; Hoeijmakers, J. H.; Kanaar, R., The

structure-specific endonuclease Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks. *Molecular and cellular biology* **2004,** *24* (13), 5776-87.
154. Rothfuss, A.; Grompe, M., Repair kinetics of genomic interstrand DNA cross-links: evidence for DNA double-strand break-dependent activation of the Fanconi anemia/BRCA pathway. *Molecular and cellular biology* **2004,** *24* (1), 123-34.
155. Vare, D.; Groth, P.; Carlsson, R.; Johansson, F.; Erixon, K.; Jenssen, D., DNA interstrand crosslinks induce a potent replication block followed by formation and repair of double strand breaks in intact mammalian cells. *DNA repair* **2012,** *11* (12), 976-85.
156. McHugh, P. J.; Sones, W. R.; Hartley, J. A., Repair of intermediate structures produced at DNA interstrand cross-links in Saccharomyces cerevisiae. *Molecular and cellular biology* **2000,** *20* (10), 3425-33.
157. Zhang, J.; Walter, J. C., Mechanism and regulation of incisions during DNA interstrand cross-link repair. *DNA repair* **2014,** *19*, 135-42.
158. Ho, T. V.; Scharer, O. D., Translesion DNA synthesis polymerases in DNA interstrand crosslink repair. *Environmental and molecular mutagenesis* **2010,** *51* (6), 552-66.
159. Knipscheer, P. R. M. S. O. D. W. J. C., Replication-coupled DNA interstrand cross-link repair in Xenopus egg extracts. *Methods Mol Biol* **2012,** (920), 221-243.
160. Knipscheer, P.; Raschle, M.; Smogorzewska, A.; Enoiu, M.; Ho, T. V.; Scharer, O. D.; Elledge, S. J.; Walter, J. C., The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* **2009,** *326* (5960), 1698-701.
161. Long, D. T.; Raschle, M.; Joukov, V.; Walter, J. C., Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* **2011,** *333* (6038), 84-7.
162. Walden, H.; Deans, A. J., The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. *Annual review of biophysics* **2014,** *43*, 257-78.
163. Bogliolo, M.; Schuster, B.; Stoepker, C.; Derkunt, B.; Su, Y.; Raams, A.; Trujillo, J. P.; Minguillon, J.; Ramirez, M. J.; Pujol, R.; Casado, J. A.; Banos, R.; Rio, P.; Knies, K.; Zuniga, S.; Benitez, J.; Bueren, J. A.; Jaspers, N. G.; Scharer, O. D.; de Winter, J. P.; Schindler, D.; Surralles, J., Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *American journal of human genetics* **2013,** *92* (5), 800-6.
164. Kashiyama, K.; Nakazawa, Y.; Pilz, D. T.; Guo, C.; Shimada, M.; Sasaki, K.; Fawcett, H.; Wing, J. F.; Lewin, S. O.; Carr, L.; Li, T. S.; Yoshiura, K.; Utani, A.; Hirano, A.; Yamashita, S.; Greenblatt, D.; Nardo, T.; Stefanini, M.; McGibbon, D.; Sarkany, R.; Fassihi, H.; Takahashi, Y.; Nagayama, Y.; Mitsutake, N.; Lehmann, A. R.; Ogi, T., Malfunction of nuclease ERCC1-XPF results in diverse clinical manifestations and causes Cockayne syndrome, xeroderma pigmentosum, and Fanconi anemia. *American journal of human genetics* **2013,** *92* (5), 807-19.
165. Scharer, O. D., Nucleotide excision repair in eukaryotes. *Cold Spring Harbor perspectives in biology* **2013,** *5* (10), a012609.

166. Hoy, C. A.; Thompson, L. H.; Mooney, C. L.; Salazar, E. P., Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. *Cancer research* **1985,** *45* (4), 1737-43.
167. Wood, R. D., Mammalian nucleotide excision repair proteins and interstrand crosslink repair. *Environmental and molecular mutagenesis* **2010,** *51* (6), 520-6.
168. Mouw, K. W.; D'Andrea, A. D., Crosstalk between the nucleotide excision repair and Fanconi anemia/BRCA pathways. *DNA repair* **2014,** *19*, 130-4.
169. Crossan, G. P.; Patel, K. J., The Fanconi anaemia pathway orchestrates incisions at sites of crosslinked DNA. *The Journal of pathology* **2012,** *226* (2), 326-37.
170. Jakobs, P. M.; Fiddler-Odell, E.; Reifsteck, C.; Olson, S.; Moses, R. E.; Grompe, M., Complementation group assignments in Fanconi anemia fibroblast cell lines from North America. *Somat Cell Mol Genet* **1997,** *23* (1), 1-7.
171. Timmers, C.; Taniguchi, T.; Hejna, J.; Reifsteck, C.; Lucas, L.; Bruun, D.; Thayer, M.; Cox, B.; Olson, S.; D'Andrea, A. D.; Moses, R.; Grompe, M., Positional cloning of a novel Fanconi anemia gene, FANCD2. *Mol Cell* **2001,** *7* (2), 241-8.
172. Michaelson-Richie, E. D.; Ming, X.; Codreanu, S. G.; Loeber, R. L.; Liebler, D. C.; Campbell, C.; Tretyakova, N. Y., Mechlorethamine-induced DNA-protein cross-linking in human fibrosarcoma (HT1080) cells. *J Proteome Res* **2011,** *10* (6), 2785-96.
173. Kratz, K.; Schopf, B.; Kaden, S.; Sendoel, A.; Eberhard, R.; Lademann, C.; Cannavo, E.; Sartori, A. A.; Hengartner, M. O.; Jiricny, J., Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. *Cell* **2010,** *142* (1), 77-88.
174. MacKay, C.; Declais, A. C.; Lundin, C.; Agostinho, A.; Deans, A. J.; MacArtney, T. J.; Hofmann, K.; Gartner, A.; West, S. C.; Helleday, T.; Lilley, D. M.; Rouse, J., Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. *Cell* **2010,** *142* (1), 65-76.
175. Smogorzewska, A.; Desetty, R.; Saito, T. T.; Schlabach, M.; Lach, F. P.; Sowa, M. E.; Clark, A. B.; Kunkel, T. A.; Harper, J. W.; Colaiacovo, M. P.; Elledge, S. J., A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. *Mol Cell* **2010,** *39* (1), 36-47.
176. Klein Douwel, D.; Boonen, R. A.; Long, D. T.; Szypowska, A. A.; Raschle, M.; Walter, J. C.; Knipscheer, P., XPF-ERCC1 acts in Unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. *Mol Cell* **2014,** *54* (3), 460-71.
177. Pace, P.; Mosedale, G.; Hodskinson, M. R.; Rosado, I. V.; Sivasubramaniam, M.; Patel, K. J., Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. *Science* **2010,** *329* (5988), 219-23.
178. Adamo, A.; Collis, S. J.; Adelman, C. A.; Silva, N.; Horejsi, Z.; Ward, J. D.; Martinez-Perez, E.; Boulton, S. J.; La Volpe, A., Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. *Mol Cell* **2010,** *39* (1), 25-35.
179. Meng, Q.; Redetzke, D. L.; Hackfeld, L. C.; Hodge, R. P.; Walker, D. M.; Walker, V. E., Mutagenicity of stereochemical configurations of 1,2-epoxybutene and 1,2:3,4-

diepoxybutane in human lymphblastoid cells. *Chemico-biological interactions* **2007,** *166* (1-3), 207-18.
180. Steen, A. M.; Meyer, K. G.; Recio, L., Analysis of hprt mutations occurring in human TK6 lymphoblastoid cells following exposure to 1,2,3,4-diepoxybutane. *Mutagenesis* **1997,** *12* (2), 61-7.
181. Vral, A.; Fenech, M.; Thierens, H., The micronucleus assay as a biological dosimeter of in vivo ionising radiation exposure. *Mutagenesis* **2011,** *26* (1), 11-7.
182. Khattab, M.; Walker, D. M.; Albertini, R. J.; Nicklas, J. A.; Lundblad, L. K. A.; Vacek, P. M.; Walker, V. E., Frequencies of micronucleated reticulocytes, a dosimeter of DNA double-strand breaks, in infants receiving computed tomography or cardiac catheterization. *Mutation research* **2017,** *820*, 8-18.
183. Lai, C.; Cao, H.; Hearst, J. E.; Corash, L.; Luo, H.; Wang, Y., Quantitative analysis of DNA interstrand cross-links and monoadducts formed in human cells induced by psoralens and UVA irradiation. *Analytical chemistry* **2008,** *80* (22), 8790-8.
184. Cao, H.; Hearst, J. E.; Corash, L.; Wang, Y., LC-MS/MS for the detection of DNA interstrand cross-links formed by 8-methoxypsoralen and UVA irradiation in human cells. *Analytical chemistry* **2008,** *80* (8), 2932-8.
185. Liu, S.; Wang, Y., A quantitative mass spectrometry-based approach for assessing the repair of 8-methoxypsoralen-induced DNA interstrand cross-links and monoadducts in mammalian cells. *Analytical chemistry* **2013,** *85* (14), 6732-9.
186. Bunting, S. F.; Callen, E.; Kozak, M. L.; Kim, J. M.; Wong, N.; Lopez-Contreras, A. J.; Ludwig, T.; Baer, R.; Faryabi, R. B.; Malhowski, A.; Chen, H. T.; Fernandez-Capetillo, O.; D'Andrea, A.; Nussenzweig, A., BRCA1 functions independently of homologous recombination in DNA interstrand crosslink repair. *Mol Cell* **2012,** *46* (2), 125-35.
187. Howard, S. M.; Yanez, D. A.; Stark, J. M., DNA damage response factors from diverse pathways, including DNA crosslink repair, mediate alternative end joining. *PLoS Genet* **2015,** *11* (1), e1004943.
188. Albertini, R. J.; Sram, R. J.; Vacek, P. M.; Lynch, J.; Nicklas, J. A.; van Sittert, N. J.; Boogaard, P. J.; Henderson, R. F.; Swenberg, J. A.; Tates, A. D.; Ward, J. B., Jr.; Wright, M.; Ammenheuser, M. M.; Binkova, B.; Blackwell, W.; de Zwart, F. A.; Krako, D.; Krone, J.; Megens, H.; Musilova, P.; Rajska, G.; Ranasinghe, A.; Rosenblatt, J. I.; Rossner, P.; Rubes, J.; Sullivan, L.; Upton, P.; Zwinderman, A. H., Biomarkers in Czech workers exposed to 1,3-butadiene: a transitional epidemiologic study. *Res Rep Health Eff Inst* **2003,** (116), 1-141; discussion 143-62.
189. Albertini, R. J.; Carson, M. L.; Kirman, C. R.; Gargas, M. L., 1,3-Butadiene: II. Genotoxicity profile. *Crit Rev Toxicol* **2010,** *40 Suppl 1*, 12-73.
190. Albertini, R. J.; Sram, R. J.; Vacek, P. M.; Lynch, J.; Rossner, P.; Nicklas, J. A.; McDonald, J. D.; Boysen, G.; Georgieva, N.; Swenberg, J. A., Molecular epidemiological studies in 1,3-butadiene exposed Czech workers: female-male comparisons. *Chemico-biological interactions* **2007,** *166* (1-3), 63-77.

191. Simmons, S. O.; Fan, C. Y.; Ramabhadran, R., Cellular stress response pathway system as a sentinel ensemble in toxicological screening. *Toxicol Sci* **2009,** *111* (2), 202-25.
192. Walker, D. M.; Patrick O'Neill, J.; Tyson, F. L.; Walker, V. E., The stress response resolution assay. I. Quantitative assessment of environmental agent/condition effects on cellular stress resolution outcomes in epithelium. *Environmental and molecular mutagenesis* **2013,** *54* (4), 268-80.
193. Walker, D. M.; Nicklas, J. A.; Walker, V. E., The stress response resolution assay. II. Quantitative assessment of environmental agent/condition effects on cellular stress resolution outcomes in epithelium. *Environmental and molecular mutagenesis* **2013,** *54* (4), 281-93.
194. Haiman, C. A.; Stram, D. O.; Wilkens, L. R.; Pike, M. C.; Kolonel, L. N.; Henderson, B. E.; Le Marchand, L., Ethnic and racial differences in the smoking-related risk of lung cancer. *The New England journal of medicine* **2006,** *354* (4), 333-42.
195. Basu, A. K., DNA Damage, Mutagenesis and Cancer. *Int J Mol Sci* **2018,** *19* (4).
196. Roethig, H. J.; Munjal, S.; Feng, S.; Liang, Q.; Sarkar, M.; Walk, R. A.; Mendes, P. E., Population estimates for biomarkers of exposure to cigarette smoke in adult U.S. cigarette smokers. *Nicotine Tob Res* **2009,** *11* (10), 1216-25.
197. Carmella, S. G.; Chen, M.; Han, S.; Briggs, A.; Jensen, J.; Hatsukami, D. K.; Hecht, S. S., Effects of smoking cessation on eight urinary tobacco carcinogen and toxicant biomarkers. *Chemical research in toxicology* **2009,** *22* (4), 734-41.
198. Boldry, E. J.; Patel, Y. M.; Kotapati, S.; Esades, A.; Park, S. L.; Tiirikainen, M.; Stram, D. O.; Le Marchand, L.; Tretyakova, N., Genetic Determinants of 1,3-Butadiene Metabolism and Detoxification in Three Populations of Smokers with Different Risks of Lung Cancer. *Cancer Epidemiol Biomarkers Prev* **2017,** *26* (7), 1034-1042.
199. Consortium, I. H., A haplotype map of the human genome. *Nature* **2005,** *437* (7063), 1299-320.
200. Citti, L.; Gervasi, P. G.; Turchi, G.; Bellucci, G.; Bianchini, R., The reaction of 3,4-epoxy-1-butene with deoxyguanosine and DNA in vitro: synthesis and characterization of the main adducts. *Carcinogenesis* **1984,** *5* (1), 47-52.
201. Cho, S. H.; Loecken, E. M.; Guengerich, F. P., Mutagenicity of a glutathione conjugate of butadiene diepoxide. *Chemical research in toxicology* **2010,** *23* (10), 1544-6.
202. McCarroll, S. A.; Hadnott, T. N.; Perry, G. H.; Sabeti, P. C.; Zody, M. C.; Barrett, J. C.; Dallaire, S.; Gabriel, S. B.; Lee, C.; Daly, M. J.; Altshuler, D. M., Common deletion polymorphisms in the human genome. *Nat Genet* **2006,** *38* (1), 86-92.
203. Vermes, I.; Haanen, C.; Steffens-Nakken, H.; Reutelingsperger, C., A novel assay for apoptosis. Flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled Annexin V. *J Immunol Methods* **1995,** *184* (1), 39-51.
204. Park, S. L.; Kotapati, S.; Wilkens, L. R.; Tiirikainen, M.; Murphy, S. E.; Tretyakova, N.; Le Marchand, L., 1,3-Butadiene exposure and metabolism among Japanese American, Native Hawaiian, and White smokers. In *Cancer Epidemiol*

*Biomarkers Prev*, (c)2014 American Association for Cancer Research.: United States, 2014; Vol. 23, pp 2240-9.
205. Kotapati, S.; Esades, A.; Matter, B.; Le, C.; Tretyakova, N., High throughput HPLC-ESI(-)-MS/MS methodology for mercapturic acid metabolites of 1,3-butadiene: Biomarkers of exposure and bioactivation. *Chemico-biological interactions* **2015,** *241*, 23-31.
206. Hinchman, C. A.; Ballatori, N., Glutathione conjugation and conversion to mercapturic acids can occur as an intrahepatic process. *J Toxicol Environ Health* **1994,** *41* (4), 387-409.
207. Zhang, P. P.; Wen, Y.; An, J.; Yu, Y. X.; Wu, M. H.; Zhang, X. Y., DNA damage induced by three major metabolites of 1,3-butadiene in human hepatocyte L02 cells. *Mutation research* **2012,** *747* (2), 240-5.
208. Hayes, J. D.; Flanagan, J. U.; Jowsey, I. R., Glutathione transferases. *Annu Rev Pharmacol Toxicol* **2005,** *45*, 51-88.
209. Wiencke, J. K.; Pemble, S.; Ketterer, B.; Kelsey, K. T., Gene deletion of glutathione S-transferase theta: correlation with induced genetic damage and potential role in endogenous mutagenesis. *Cancer Epidemiol Biomarkers Prev* **1995,** *4* (3), 253-9.
210. Fustinoni, S.; Soleo, L.; Warholm, M.; Begemann, P.; Rannug, A.; Neumann, H. G.; Swenberg, J. A.; Vimercati, L.; Colombi, A., Influence of metabolic genotypes on biomarkers of exposure to 1,3-butadiene in humans. *Cancer Epidemiol Biomarkers Prev* **2002,** *11* (10 Pt 1), 1082-90.
211. Granath, F. N.; Vaca, C. E.; Ehrenberg, L. G.; Tornqvist, M. A., Cancer risk estimation of genotoxic chemicals based on target dose and a multiplicative model. *Risk Anal* **1999,** *19* (2), 309-20.
212. Fred, C.; Tornqvist, M.; Granath, F., Evaluation of cancer tests of 1,3-butadiene using internal dose, genotoxic potency, and a multiplicative risk model. *Cancer research* **2008,** *68* (19), 8014-21.
213. Wiseman, M., The second World Cancer Research Fund/American Institute for Cancer Research expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. *The Proceedings of the Nutrition Society* **2008,** *67* (3), 253-6.
214. Carlsson, H.; Törnqvist, M., An Adductomic Approach to Identify Electrophiles In Vivo. *Basic & clinical pharmacology & toxicology* **2017,** *121 Suppl 3*, 44-54.
215. Törnqvist, M.; Svartengren, M.; Ericsson, C. H., Methylations in hemoglobin from monozygotic twins discordant for cigarette smoking: hereditary and tobacco-related factors. *Chemico-biological interactions* **1992,** *82* (1), 91-8.
216. Shimada, S.; Tanaka, Y.; Ohmura, C.; Tamura, Y.; Shimizu, T.; Uchino, H.; Watada, H.; Hirose, T.; Nakaniwa, T.; Miwa, S.; Kawamori, R., N-(carboxymethyl)valine residues in hemoglobin (CMV-Hb) reflect accumulation of oxidative stress in diabetic patients. *Diabetes research and clinical practice* **2005,** *69* (3), 272-8.
217. Uchimura, T.; Nakano, K.; Hashiguchi, T.; Iwamoto, H.; Miura, K.; Yoshimura, Y.; Hanyu, N.; Hirata, K.; Imakuma, M.; Motomiya, Y.; Maruyama, I., Elevation of N-

(carboxymethyl)valine residue in hemoglobin of diabetic patients. Its role in the development of diabetic nephropathy. *Diabetes care* **2001,** *24* (5), 891-6.
218. Poulsen, M. W.; Hedegaard, R. V.; Andersen, J. M.; de Courten, B.; Bugel, S.; Nielsen, J.; Skibsted, L. H.; Dragsted, L. O., Advanced glycation endproducts in food and their effects on health. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* **2013,** *60*, 10-37.
219. Peter, M. G., Chemical Modifications of Biopolymers by Quinones and Quinone Methides. *Angew. Chem. Int. Ed. Engl.* **1989,** *28* (5), 555-570.
220. Enoch, S. J.; Ellison, C. M.; Schultz, T. W.; Cronin, M. T., A review of the electrophilic reaction chemistry involved in covalent protein binding relevant to toxicity. *Crit Rev Toxicol* **2011,** *41* (9), 783-802.
221. Huang, C.; Liu, Y.; Rokita, S. E., Targeting duplex DNA with the reversible reactivity of quinone methides. *Signal transduction and targeted therapy* **2016,** *1*.
222. Rokita, S. E.; Yang, J.; Pande, P.; Greenberg, W. A., Quinone Methide Alkylation of Deoxycytidine. *The Journal of organic chemistry* **1997,** *62* (9), 3010-3012.
223. Weinert, E. E.; Frankenfield, K. N.; Rokita, S. E., Time-dependent evolution of adducts formed between deoxynucleosides and a model quinone methide. *Chemical research in toxicology* **2005,** *18* (9), 1364-70.
224. Rydberg, P. Patent: Method for Analyzing N-terminal Protein Adducts Using Isothiocyanate Reagents. International Publication Number: WO 2005/101020 AI. 2005.
225. *Quinone Methides*. John Wiley & Sons, Inc.: 2009.
226. Bolton, J. L.; Comeau, E.; Vukomanovic, V., The influence of 4-alkyl substituents on the formation and reactivity of 2-methoxy-quinone methides: evidence that extended pi-conjugation dramatically stabilizes the quinone methide formed from eugenol. *Chemico-biological interactions* **1995,** *95* (3), 279-90.
227. Thompson, D. C.; Thompson, J. A.; Sugumaran, M.; Moldeus, P., Biological and toxicological consequences of quinone methide formation. *Chemico-biological interactions* **1993,** *86* (2), 129-62.
228. Bolton, J. L.; Valerio, L. G., Jr.; Thompson, J. A., The enzymatic formation and chemical reactivity of quinone methides correlate with alkylphenol-induced toxicity in rat hepatocytes. *Chemical research in toxicology* **1992,** *5* (6), 816-22.
229. Toteva, M. M.; Moran, M.; Amyes, T. L.; Richard, J. P., Substituent effects on carbocation stability: the pK(R) for p-quinone methide. *Journal of the American Chemical Society* **2003,** *125* (29), 8814-9.
230. McCrane, M. P.; Hutchinson, M. A.; Ad, O.; Rokita, S. E., Oxidative quenching of quinone methide adducts reveals transient products of reversible alkylation in duplex DNA. *Chemical research in toxicology* **2014,** *27* (7), 1282-93.
231. McCrane, M. P.; Weinert, E. E.; Lin, Y.; Mazzola, E. P.; Lam, Y. F.; Scholl, P. F.; Rokita, S. E., Trapping a labile adduct formed between an ortho-quinone methide and 2'-deoxycytidine. *Organic letters* **2011,** *13* (5), 1186-9.
232. Weinert, E. E.; Dondi, R.; Colloredo-Melz, S.; Frankenfield, K. N.; Mitchell, C. H.; Freccero, M.; Rokita, S. E., Substituents on quinone methides strongly modulate formation

and stability of their nucleophilic adducts. *Journal of the American Chemical Society* **2006,** *128* (36), 11940-7.
233. Bolton, J. L.; Turnipseed, S. B.; Thompson, J. A., Influence of quinone methide reactivity on the alkylation of thiol and amino groups in proteins: studies utilizing amino acid and peptide models. *Chemico-biological interactions* **1997,** *107* (3), 185-200.
234. Stich, T. A.; Myers, W. K.; Britt, R. D., Paramagnetic intermediates generated by radical S-adenosylmethionine (SAM) enzymes. *Accounts of chemical research* **2014,** *47* (8), 2235-43.
235. Graf, E., *Chinese Drugs of Plant Origin. Chemistry, Pharmacology, and Use in Traditional and Modern Medicine*. Springer-Verlag Berlin: 1992.
236. Zhang, J.; Mohamad, H.; Wong, J. H.; Bilal, M.; Lloyd, A. J.; Yusoff, A. A. M.; Osman, H.; Wong, K. T.; Idris, Z.; Abdulluh, J. M., The Effect of 4-hydroxybenzaldehyde on the γ-aminobutyric Acid Type A Receptor. *Maylays J Med Sci* **2017,** *24* (2), 94-99.
237. Manley, S. L.; Chapman, D. J., Metabolism of l-Tyrosine to 4-Hydroxybenzaldehyde and 3-Bromo-4-Hydroxybenzaldehyde by Chloroplast-containing Fractions of Odonthalia floccosa (Esp.) Falk. *Plant Physiol* **1979,** *64* (6), 1032-8.
238. Jakas, A.; Horvat, S., Study of degradation pathways of Amadori compounds obtained by glycation of opioid pentapeptide and related smaller fragments: stability, reactions, and spectroscopic properties. *Biopolymers* **2003,** *69* (4), 421-31.
239. Podstolski, A.; Havkin-Frenkel, D.; Malinowski, J.; Blount, J. W.; Kourteva, G.; Dixon, R. A., Unusual 4-hydroxybenzaldehyde synthase activity from tissue cultures of the vanilla orchid Vanilla planifolia. *Phytochemistry* **2002,** *61* (6), 611-20.
240. Sinha, A. K.; Sharma, U. K.; Sharma, N., A comprehensive review on vanilla flavor: extraction, isolation and quantification of vanillin and others constituents. *International journal of food sciences and nutrition* **2008,** *59* (4), 299-326.
241. Pérez Silva, A.; Gunata, Z.; Lepoutre, J.; Odoux, E., New insight on the genesis and fate of odor-active compounds in vanilla beans (Vanilla planifolia G. Jackson) during traditional curing. *Food Res Int* **2011,** *44* (9), 2930-2937.
242. Janes, D.; Kreft, S., Salicylaldehyde is a characteristic aroma component of buckwheat groats. *Food chemistry* **2008,** *109* (2), 293-8.
243. Sugitani, N.; Sivley, R. M.; Perry, K. E.; Capra, J. A.; Chazin, W. J., XPA: A key scaffold for human nucleotide excision repair. *DNA repair* **2016,** *44*, 123-135.
244. Kang, H.; Tolbert, T. J.; Schoneich, C., Photoinduced Tyrosine Side Chain Fragmentation in IgG4-Fc: Mechanisms and Solvent Isotope Effects. *Mol Pharm* **2019,** *16* (1), 258-272.

# APPENDICES

## A. NMR Spectra

**Figure A1.** $^{1}H$ NMR spectrum of synthesized EB-GSH.

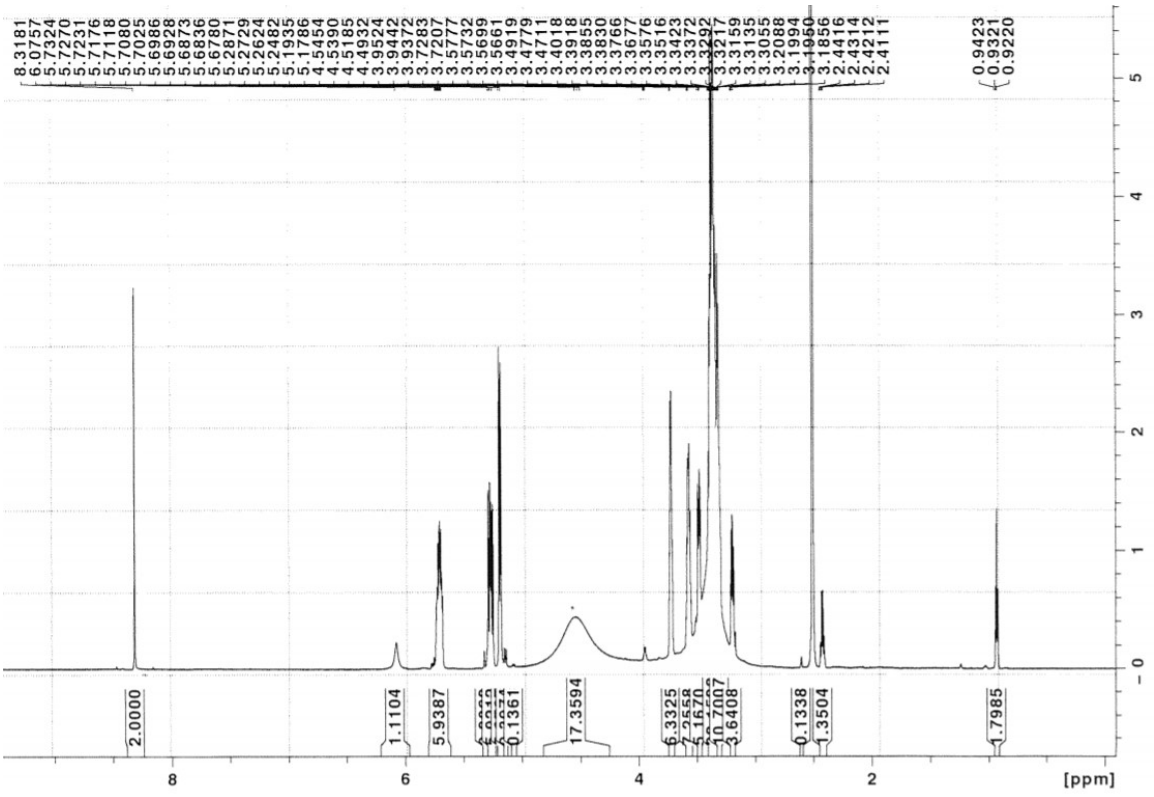
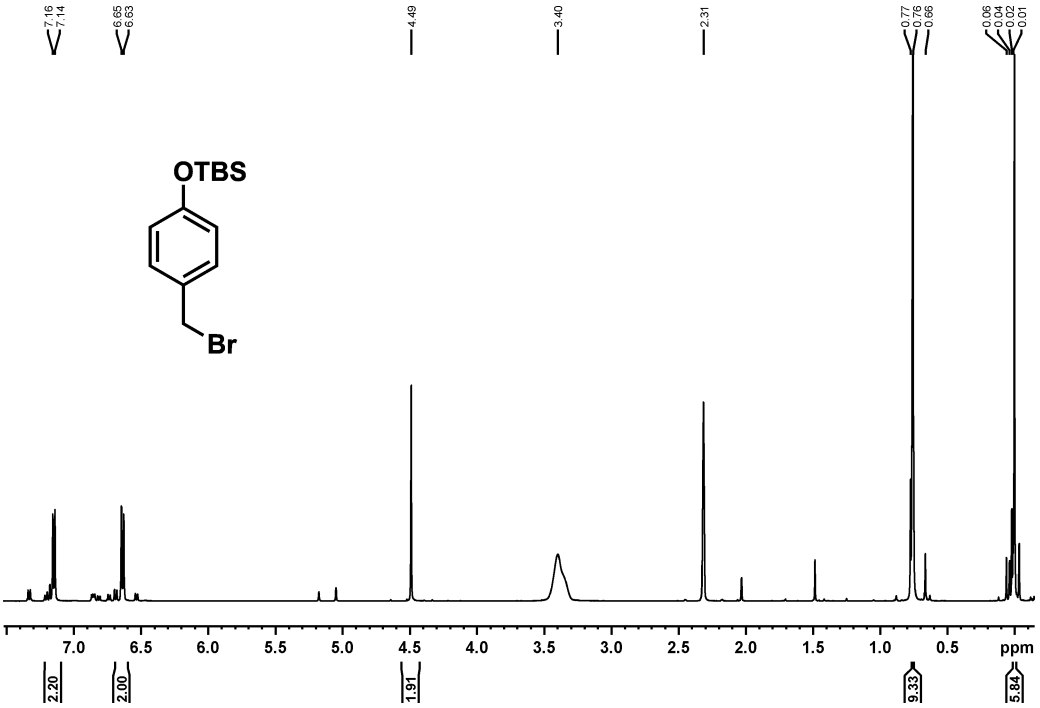


**Figure A2.** NMR spectrum of 4-QM precursor (compound 4.3).

a) $^{1}$H NMR spectrum



b) $^{135}$DEPT NMR spectrum

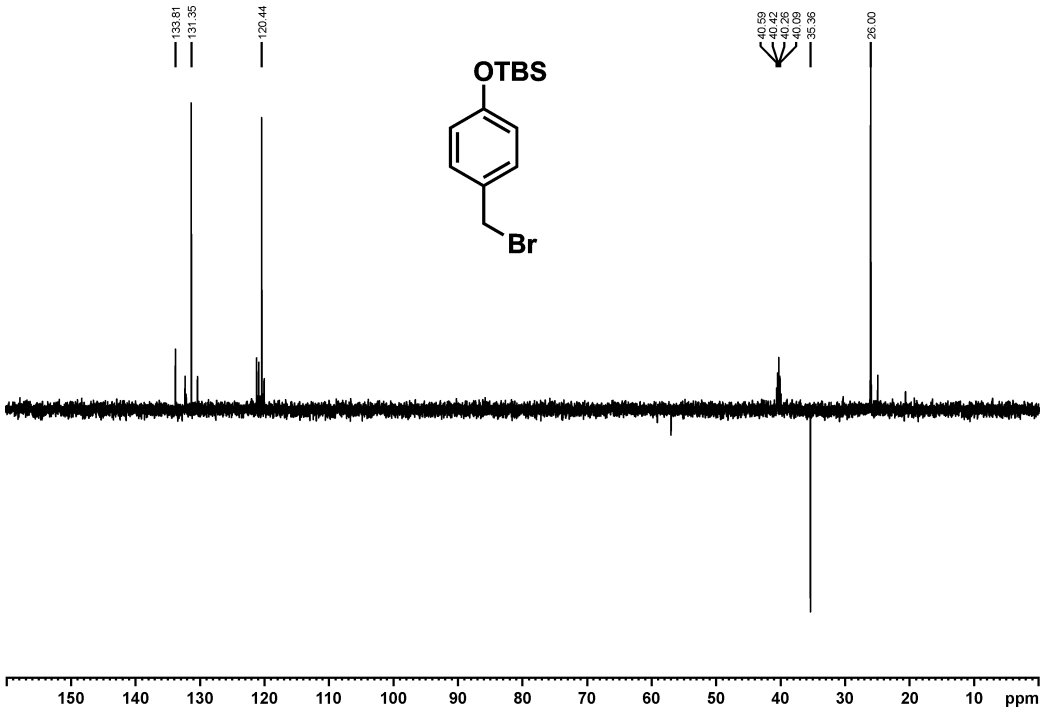
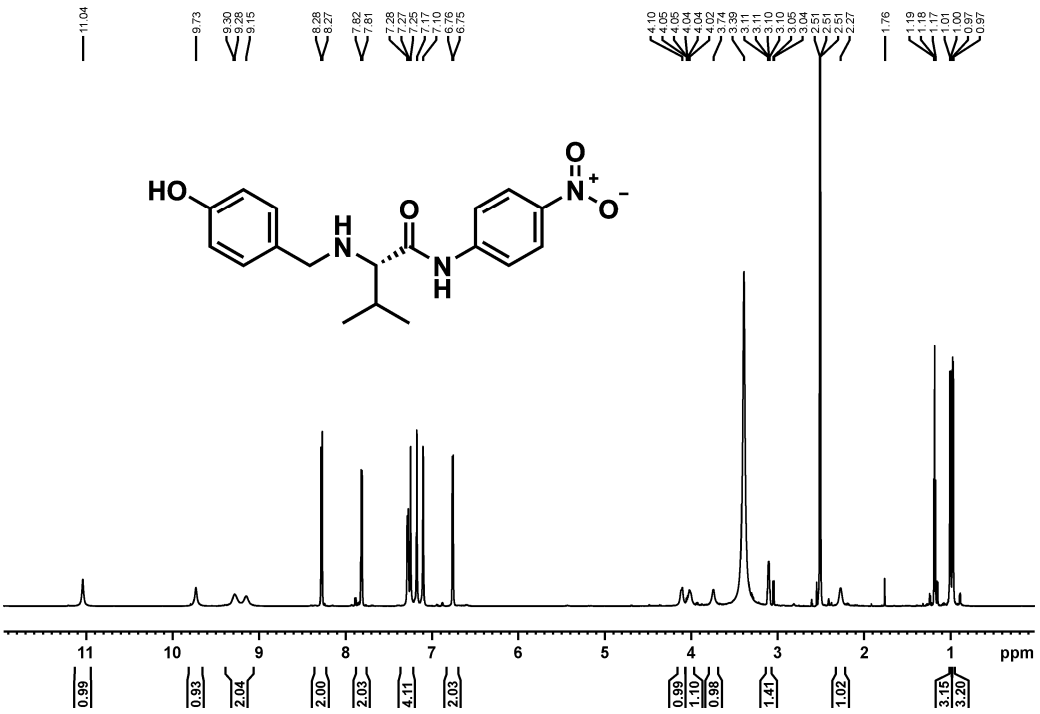


**Figure A3.** NMR spectrum of 4-OHBn-ValpNA (compound 4.9).

a) $^{1}$H NMR spectrum



b) HSQC NMR spectrum

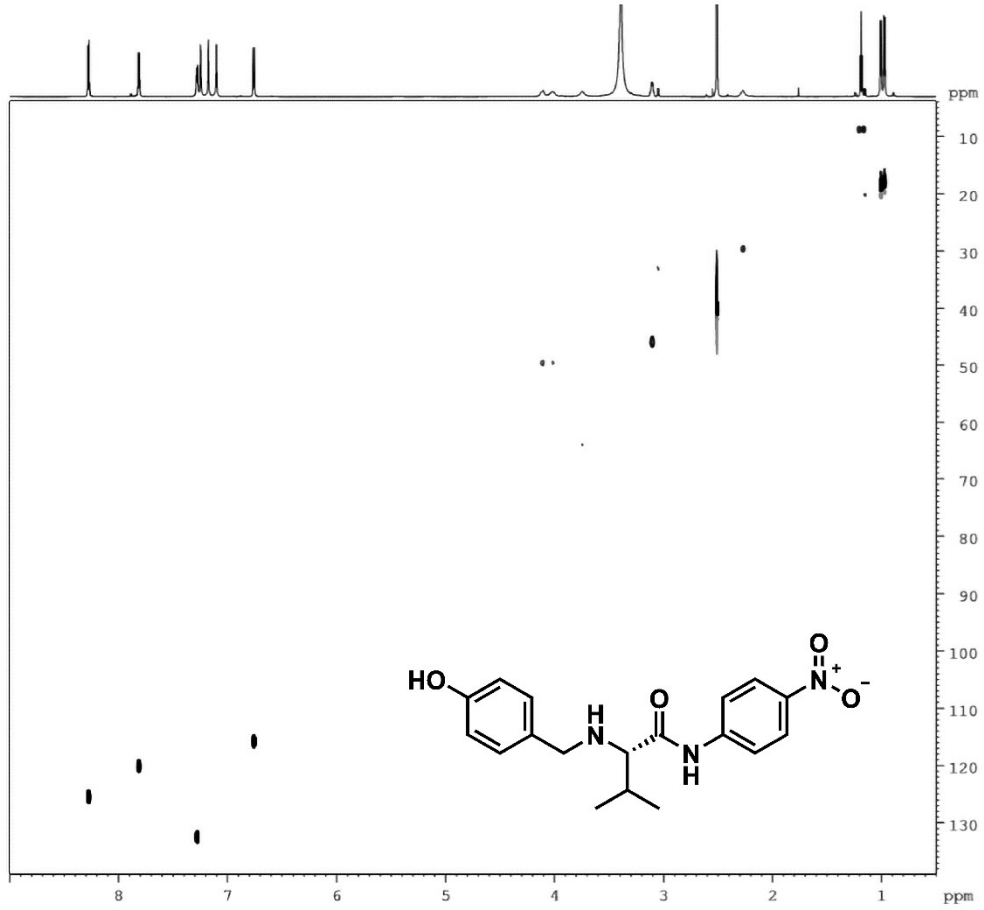
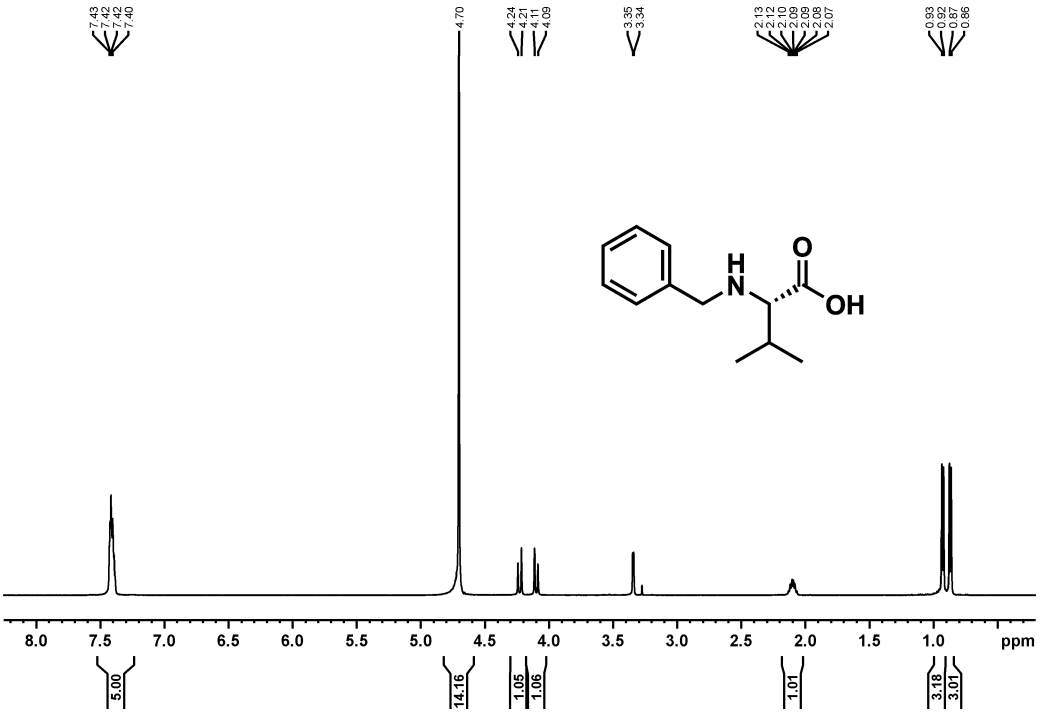


**Figure A4.** NMR spectrum of N-benzylvaline (Bn-Val) (compound 4.11) synthesized from benzaldehyde.

a) $^{1}$H NMR spectrum



b) HSQC NMR spectrum

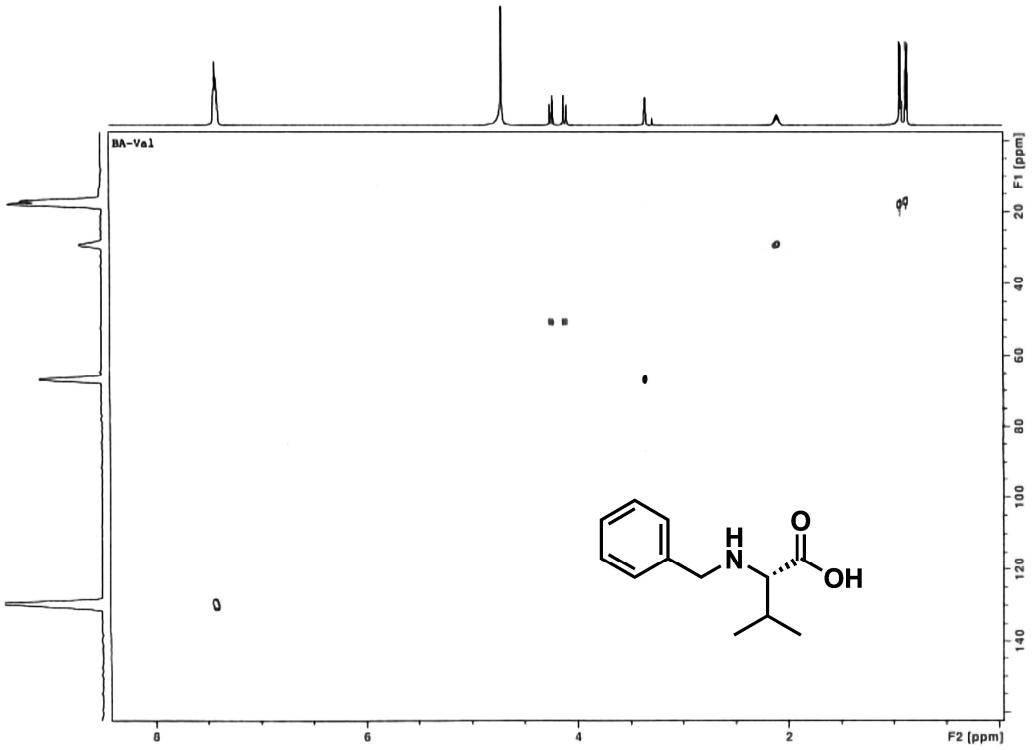
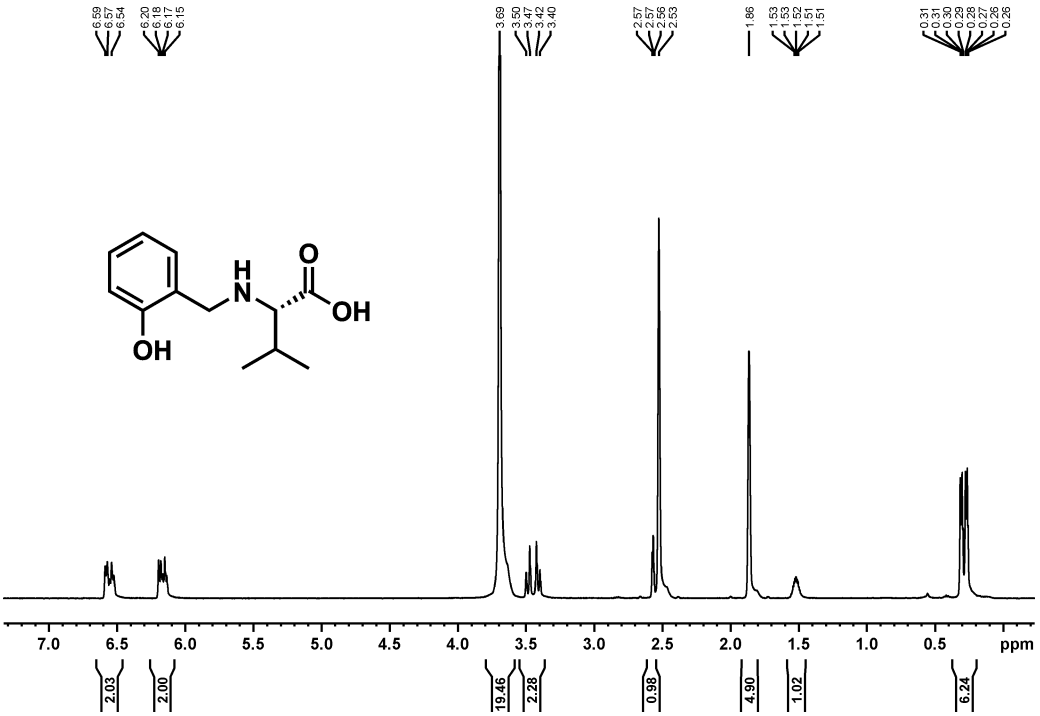


**Figure A5.** NMR spectrum of N-(2-hydroxybenzyl)valine (2-OHBn-Val) (compound 4.12) synthesized from 2-hydroxybenzaldehyde.

a) $^{1}$H NMR spectrum



b) HSQC NMR spectrum

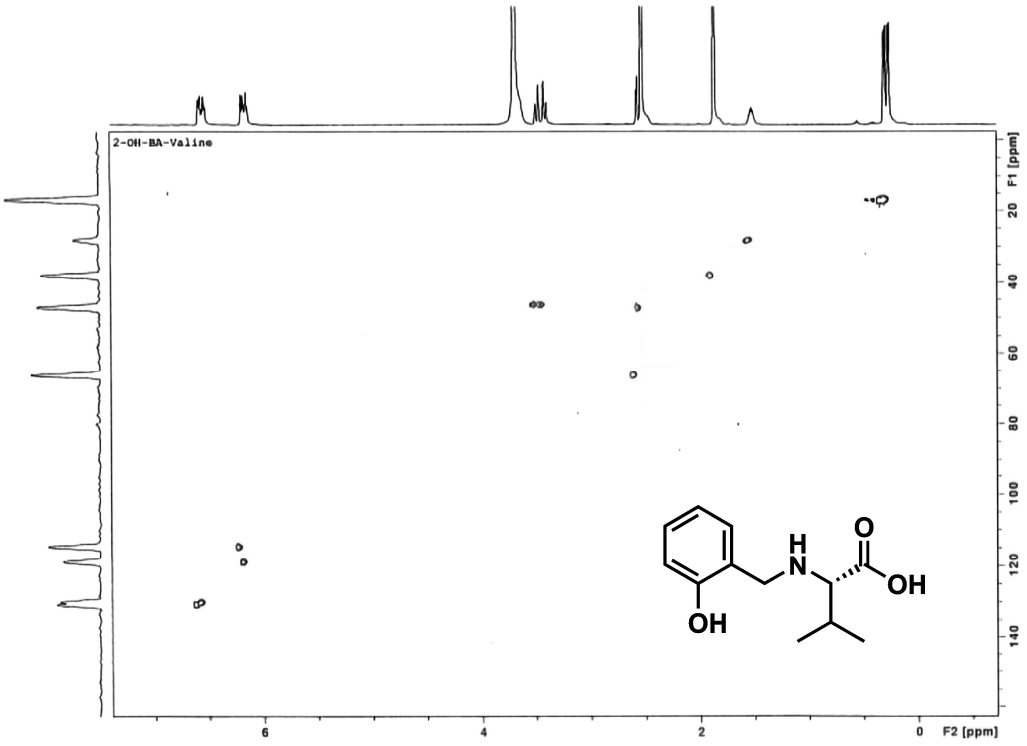
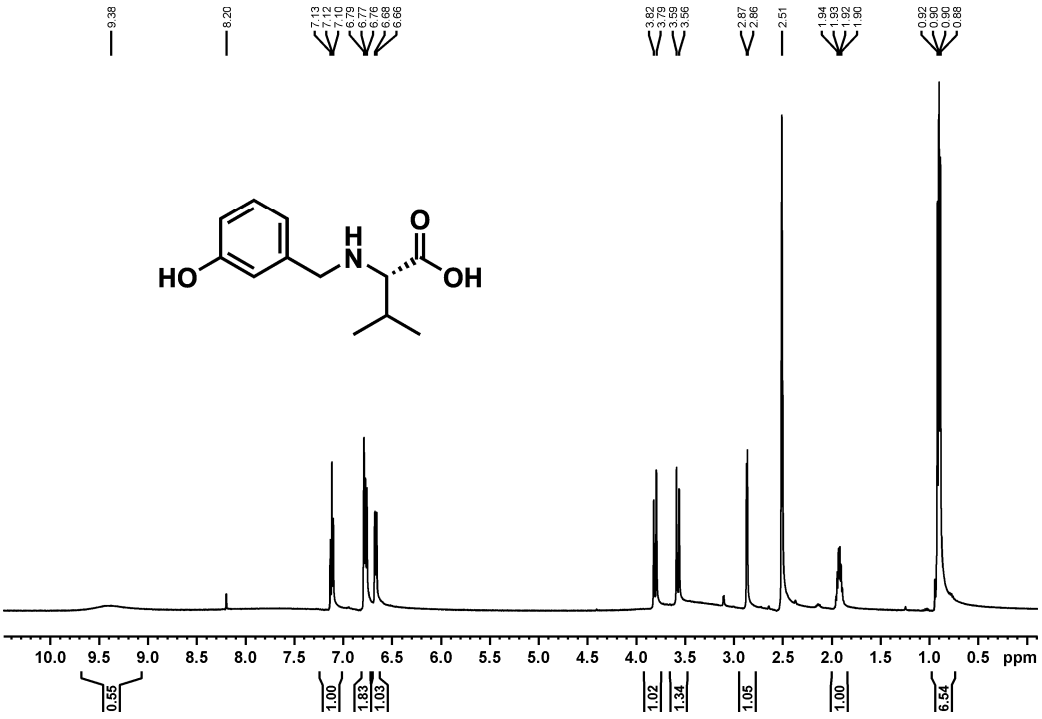


**Figure A6.** NMR spectrum of N-(3-hydroxybenzyl)valine (3-OHBn-Val) (compound 4.13) synthesized from 3-hydroxybenzaldehyde.

a) $^{1}$H NMR spectrum



b) HSQC NMR spectrum

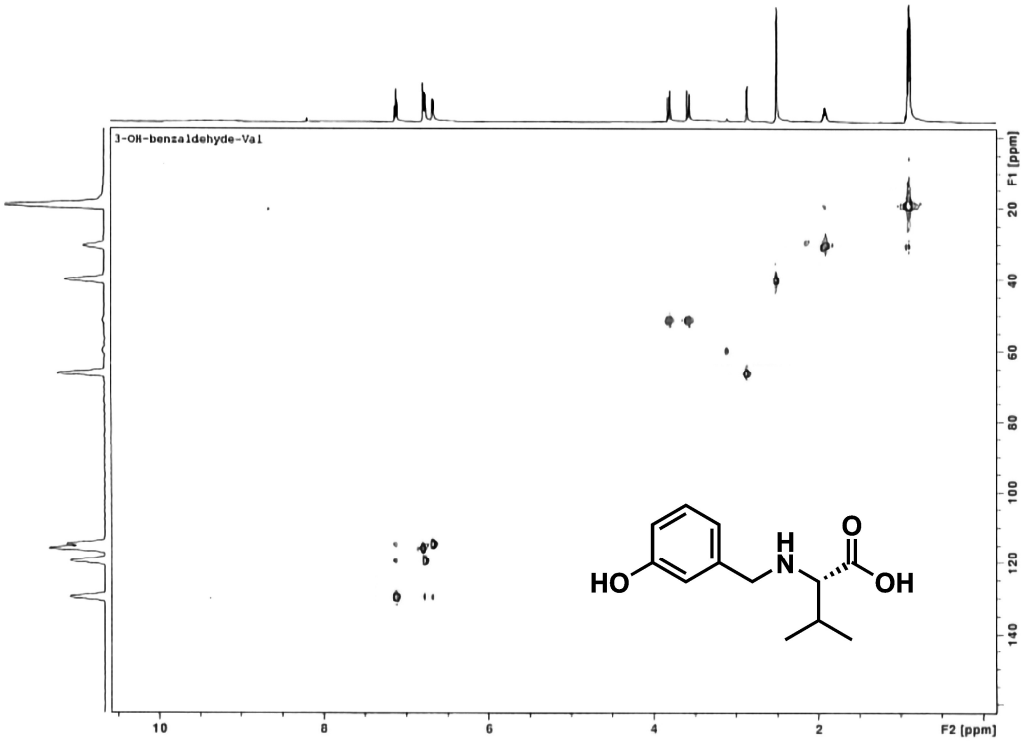
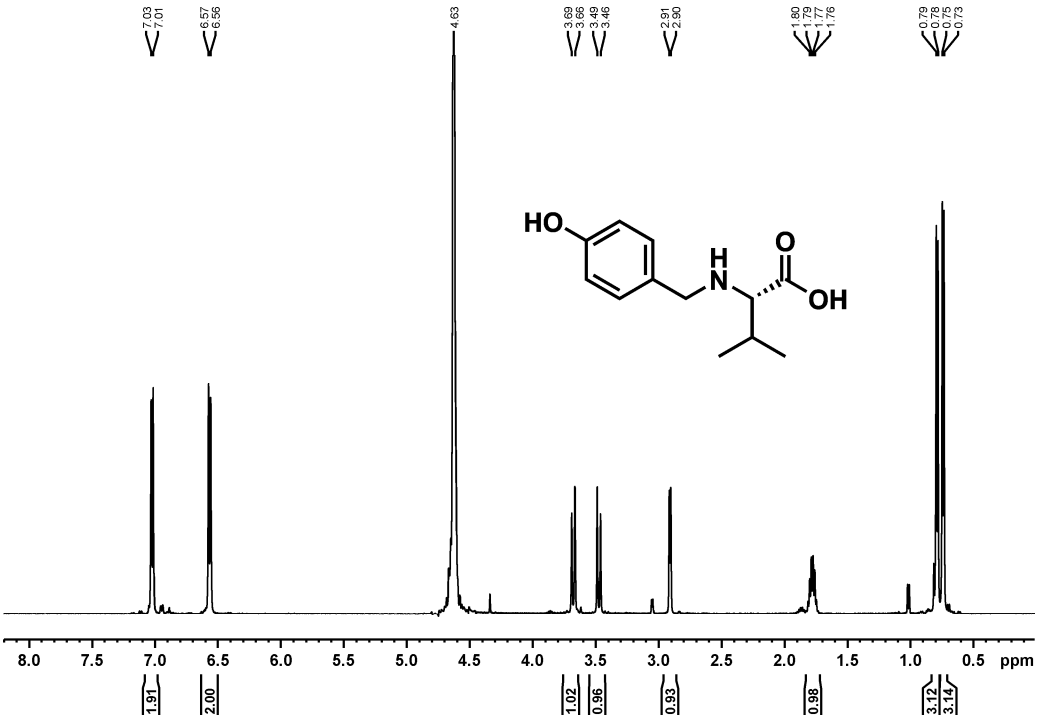


**Figure A7.** NMR spectrum of N-(4-hydroxybenzyl)valine (4-OHBn-Val) (compound 4.14) synthesized from 4-hydroxybenzaldehyde.

a) $^{1}$H NMR spectrum



b) HSQC NMR spectrum