The dynamic interplay between DNA damage and metabolism: The metabolic fate and transport of DNA lesions and novel DNA damage derived from intermediary metabolism

By

Watthanachai Jumpathong

B.S. in Chemistry (2008)
Kasetsart University, Thailand

SUBMITTED TO THE DEPARTMENT OF BIOLOGICAL ENGINEERING
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN BIOLOGICAL ENGINEERING
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2014

Author _____
Department of Biological Engineering
August 4, 2014

Certified by _____
Peter C. Dedon
Professor of Toxicology and Biological Engineering
Thesis Supervisor

Accepted by _____
Forest M. White
Professor of Biological Engineering
Chairman, Department Committee on Graduate Theses

Committee members who voted in favor of thesis:

Advisor: Peter C. Dedon
Title: Professor of Toxicology and Biological Engineering

Chair: John M. Essigmann
Title: Professor of Chemistry and Biological Engineering

Other member: Steven R. Tannenbaum
Title: Professor of Chemistry and Biological Engineering

Other member: John S. Wishnok
Title: Senior Research Scientist

The dynamic interplay between DNA damage and metabolism: The metabolic fate and transport of DNA lesions and novel DNA damage derived from intermediary metabolism

by

Watthanachai Jumpathong

Submitted to the Department of Biological Engineering on August 8, 2014, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biological Engineering

Abstract

The work presented in this thesis explores two novel and complementary facets of endogenous DNA damage: the development of biomarkers of inflammation based on metabolites of DNA damage products and the formation of DNA adducts by electrophilic products of intermediary metabolism. From the first perspective, endogenous DNA damage generated by reactive oxygen and nitrogen species from inflammation and oxidative stress has shown strong mechanistic links to the pathophysiology of cancer and other human diseases, with the damage products reflecting all types of damage chemistries including oxidation, deamination, halogenation, nitration and alkylation. However, the use of DNA damage products as biomarkers has been limited by poor understanding of the damage actually arising in tissues and a lack of appreciation of the fate of DNA damage products from the moment of formation at the site of damage to release from cells to final excretion from the body. The goal of the work presented in the first part of this thesis was to investigate the metabolic fates of the base propenal products arising from 4'-oxidation of 2'-deoxyribose in DNA, one of the most common products of DNA oxidation, and to define base propenal metabolites as potential biomarkers of oxidative stress. This project was approached with systematic metabolite profiling, starting with prediction of potential base propenal metabolites based on *a priori* knowledge of its chemical reactivity as an α,β-unsaturated aldehyde toward glutathione (GSH) in non-enzymatic reactions and in rat liver cell extracts. Of 15 potential candidates predicted and identified from these *in vitro* studies, analysis of urine samples from rats given intravenous doses (IV) of thymine propenal revealed three major metabolites: thymine propenoic acid and two mercapturic acid derivatives, which accounted for ~6% of the injected dose. An additional four metabolites, including conjugates with GSH, cysteinylglycine and cysteine, were observed in bile and accounted for ~22% of the dose. One of the major metabolites detected in urine and bile, a bis-mercapturic acid adduct of reduced thymine propenal was detected as a background excretory product in saline-treated rats and was significantly elevated after oxidative stress caused by treatment with bleomycin and CCl_4. Our observations suggest that metabolism and disposition of damaged biomolecules should be considered as crucial factors in the development of biomarkers relevant to inflammation and oxidative stress.

The second part of this thesis addresses the complementary hypothesis that electrophilic metabolites generated endogenously from intermediary metabolism can react with DNA to form adducts. This concept is illustrated here with glyoxylate from the glyoxylate metabolic cycle, whicvh plays a key role as an alternative to the TCA cycle in plants, bacteria, protists and fungi under changing conditions of environmental nutrients. The goal of this project was to characterize DNA adducts caused by glyoxylate in the mycobacterium *M. smegmatis*, with the studies motivated by the higher-than-expected mutation rate of mycobacteria during dormancy induced by nutrient deprivation and a shift to utilization of the glyoxylate cycle. Initially, *in vitro* reactions of 2'-deoxyguanosine (dG) with glyoxylate yielded N^2-carboxyhydroxymethyl dG (N^2-CHMdG) as the only adduct. However, the adduct proved to be unstable, so a reduction-based analytical method was developed to yield the stable amine derivative, N^2-carboxymethyl dG (N^2-CMdG). This stable adduct was used to develop an isotope-dilution chromatography-coupled tandem mass spectrometry method to quantify N^2-CHMdG as N^2-CMdG in calf thymus DNA treated with glyoxylate *in vitro*. This analytical method was then applied to quantify and compare the level N^2-CMdG in (1) wild-type *M. smegmatis* grown in rich medium (7H9) or in minimal M9 medium supplemented with acetate, the latter inducing a switch from the TCA cycle to the glyoxylate cycle; and (2) the isocitrate dehydrogenase (ICD)-deficient mutant of *M. smegmatis*. Mycobacteria grown in the acetate medium experienced a 2-fold increase in the adduct compared to those grown in 7H9. Similarly, the adduct increased 2-fold in the ICD mutant compared to wild-type *M. smegmatis* grown in 7H9. The results support the idea that shifts in intermediary metabolism can lead to DNA damage that may cause mutations associated with nutrient deprivation in mycobacteria, with implications for the genetic toxicology of other metabolism-derived electrophiles.

Thesis Supervisor: Peter C. Dedon
Title: Professor of Toxicology and Biological Engineering

Table of Contents

Abbreviations

AE-GSH	addition-elimination product generated by glutathione
AE-GSH^{red}	reduced addition-elimination product generated by glutathione;
AE-CG^{red}	reduced addition-elimination product generated by cysteinylglycine;
AE-Cys^{red}	reduced addition-elimination product generated by cysteine;
AE-NAc^{red}	reduced addition-elimination product generated by N-acetylcysteine;
bis-GSH	adduct conjugated with two molecules of glutathione;
bis-GSH^{red}	reduced adduct conjugated with two molecules of glutathione;
bis-CG^{red}	reduced adduct conjugated with two molecules of cysteinylglycine;
bis-Cys^{red}	reduced adduct conjugated with two molecules of cysteine;
bis-NAc^{red}	reduced adduct conjugated with two molecules of N-acetylcysteine;
$CO_3^{\bullet-}$	carbonate radical anion
Cys	cysteine
CG	cysteinylglycine
GSH	glutathione
HNE	4-hydroxy-2-nonenal
$HO^{\bullet}$	hydroxyl radical
H_2O_2	hydrogen peroxide
HOCl	hypochlorous acid
HPLC	high performance liquid chromatography
LC	liquid chromatography
LC-MS/MS	liquid chromatography-coupled tandem mass spectrometry
M_1dG	3-(2-Deoxy-beta-D-erythro-pentofuranosyl)pyrimido[1,2-α]purin-10(3H)-one
MDA	malondialdehyde
M-GSH	Michael-type adduct conjugated with GSH
M-GSH^{red}	reduced Michael-type adduct conjugated with glutathione
M-CG^{red}	reduced Michael-type adduct conjugated with cysteinylglycine
M-Cys^{red}	reduced Michael-type adduct conjugated with cysteine
M-NAc^{red}	reduced Michael-type adduct conjugated with N-acetylcysteine
MRM	multiple reaction monitoring
MS	mass spectrometry
N^2-CHMdG	N^2-carboxyhydroxymethyl 2'-deoxyguanosine
N^2-CMdG	N^2-carboxymethyl 2'-deoxyguanosine
NO	nitric oxide
$NO_2^{\bullet}$	nitrogen dioxide radical
NO_2^-	nitrite
N_2O_3	nitrous anhydride
$O_2^{\bullet-}$	superoxide anion radical
$ONOO^-$	peroxynitrite
$ONOOCO_2^-$	nitrosoperoxycarbonate
QQQ	triple quadrupole mass spectrometer
SOD	superoxide dismutase
Tp	thymine propenal
Tp^{red}	reduced thymine propenal
Tp^{ox}	oxidized thymine propenal

List of Figures

List of Scheme

List of Table

Acknowledgments

I first would like to thank Prof. Peter Dedon, my thesis advisor, for giving me a very good opportunity to do research in his laboratory. I am very appreciative of his support, guidance and kindness over the years, in particular during my first year as a graduate student at MIT. I would also like to thank Pete for his patience, encouragement and constructive scientific comments during my graduate research. I have learned a great deal from him.

Next, I am very grateful having Prof. John Essigmann as my thesis committee chair. John is a very great educator, providing me with intriguing knowledge tangential to my field of research. This immensely helps broaden my knowledge in sciences. I also want to thank Prof. Steven Tannenbaum and Dr. Pete Wishnok, my committee members, who have given me valuable suggestions in my research projects, which gives me different perspectives in my projects.

Thanks to all the past and present members of Dedon Lab. They make me feel myself at home and in turn assuage my home-sickness. I need to thank Olga Parkin and Christin Marzilli, the administrative staff I want to thank Dr. Michael Demott, Dr. Ramesh Babu, and Dr. Megan McBee for their advice, experiments and keeping the lab running smoothly. Paticularly, Dr. Megan McBee for her agility in animal experiment; I cannot start and finish it without her. I would like to thank Dr. Adides Williams for sharing with me a lot of experiences in the lab about nucleic acid chemistries—nucleic acid synthesis and phosphoramidite synthesis. Also, Adides always had funny stories to share with us, having me had enjoyable experiences in the lab among stressful atmosphere of sciences. I am also very thankful for Dr. Susovan Mohapatra, who has share his knowledge in biological sciences with me, a good colleague to do experiment with and his patience in explaining sciences for what I do not understand. I would also like to thank Yok Hian Chionh that we worked together for two months in the glyoxylate project. I would like to thank Dr. Wan Chan (Simon) for his mentorship. Simon was the very first lab member I worked with, and his tremendous helps in mass spectrometry and chromatography has accelerated my project to finish much faster. I also want to thank again Dr. Susova and Dr. Adides, Dr. Clement Chan, Dr. LianRong Wang, Dr. Joy Pang, Dr. KokSeong Lim, who share with me a lot of ups and downs. That means a lot to me.

I would like to thank my parents, Surasak and Daranee, for raising me very well and their supporting. Particularly, my mom has respected my decision for studying in sciences. Finally thank Buddha for his teaching that makes every day count in my life.

Chapter 1

Background & significance

1.1 Goals of this thesis project

The work presented in this thesis was motivated by two complementary ideas: that we must consider the biotransformational fate of DNA damage products in the development of biomarkers of inflammation and oxidative stress, and that cellular metabolism leads to the formation of reactive electrophiles capable of causing DNA damage. Using base propenals derived 4'-oxidation of DNA, the first goal aims to test the hypothesis that, after an oxidative insult occurs to DNA, the damage product is subjected to metabolism and the structure of the damage product is changed before it is excreted from the body and used as a biomarker of the stress. The second aim of this project focuses on characterizing a novel DNA adduct derived from the reaction of nucleophilic sites in DNA with electrophilic species generated from primary metabolism. The hypothesis is that electrophiles from metabolism, such as glyoxylate, can react with exocyclic amine of dG, for example, such that, when the metabolic flux is changed in the direction that the glyoxylate shunt is upregulated, the level of the DNA adduct is increased. This has implications for mutagenesis driven by endogenous metabolic processes. Both of these projects have revealed novel features of the biological chemistry of nucleic acids in both eukaryotic and prokaryotic cells.

1.2 Use of damage biomolecules as biomarkers—definition and examples

The phrases "biomarker" and "biological marker" were first coined in 1989 as a Medical Subject Heading (MeSH) term: "measurable and quantifiable biological parameters" (e.g., specific enzyme concentration, specific hormone concentration, specific gene phenotype distribution in a population, presence of biological substances). Subsequently, the term biomarker is used with various meanings. In 2011, The National Institutes of Health Biomarkers Definitions Working Group gave the definition of a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention." The International Programme on Chemical Safety defined

a biomarker as “any substance, structure, or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease.” WHO defined broader definition of a biomarker as “almost any measurement reflecting an interaction between a biological system and a potential hazard, which may be clinical, physical, or biological. The measured response may be functional and physiological, biochemical at the cellular level, or a molecular interaction.” Simplistically, biomarkers can be considered as indicators of the trait, state or rate of the disease [1].

Biomarkers are categorized based on the need or application. First, the biomarker of mechanism provides a link between cause and effect by explaining the mechanism underlying pathophysiolocal processes, such as reactive oxygen and nitrogen species (ROS, RNS) generated during inflammation [2]. Second, the biomarker of risk is used to indicate the correlation between the level of the markers and the risk of developing a disease. For example, the elevated plasma levels of cellular adhesion molecules (e.g., P- and E-selectin, etc.), cytokines (e.g., TNF-α) or C-reactive proteins are correlated with cardiovascular risk [3, 4]. The risk marker provides only a correlation with the risk of disease and not the mechanistic insights mentioned above for the marker of mechanism. Fourth, the biomarker of exposure can be used to assess the degree of exposure to chemical, biological, physical agents in air, food, water in the environment. One obvious example involves aflatoxins secreted by *Aspergillus flavus* and *Aspergillus parasiticus*. Aflatoxins are widespread contaminated in peanuts and cooling oils. When ingested, aflatoxin B1 is oxidized by cytochrome P450s to generate an oxirane species that can be nucleophilically attacked by N^7 of guanine to form a bulky DNA adduct [5]. The level of the N^7-guanine adducted with aflatoxin is quantitatively correlated with the exposure to aflatoxins [6]. The goal of my research is to define the fundamental chemical mechanisms underlying the biotransformation of DNA damage products as potential biomarkers of mechanism, risk and exposure, as well as to define the role of endogenous metabolites in the chemistry of DNA damage in cells.

1.3 Inflammation and oxidative stress as sources of cellular damage

During inflammation, innate immune cells, such as macrophages and neutrophils, generate a variety of ROS and RNA as a defensive mechanism (Figure 1.1A). These reactive species can react with virtually all types of biomolecules from nucleic acid, proteins, carbohydrates and lipids, with the chemistries of damage including oxidation, nitration, deamination, alkylation and halogenation [7]. Superoxide anion radical (O_2^-), generated from either the membrane-bound NADPH oxidase during inflammation [8] or the complex enzyme system in electron transport chain in mitochondria from oxidative stress [9], is converted to hydrogen peroxide (H_2O_2) by superoxide dismutase (SOD) [9, 10]. Subsequently, H_2O_2 can react with metals like ferrous ion (Fe^{2+}) or cuprous ion (Cu^+) to generate the reactive hydroxyl radicals ($OH^\bullet$) via Fenton chemistry. $OH^\bullet$ radical further oxidizes biomolecules as discussed below. In addition, $O_2^{\bullet-}$ can react with nitric oxide ($NO^\bullet$) produced by activated macrophages, resulting in the formation of peroxynitrite ($ONOO^-$) that is protonated to generate an unstable species peroxynitrous acid (ONOOH) [10]. The homolytic cleavage of ONOOH at the peroxide bond also gives rise to the reactive $OH^\bullet$ radical along with nitrogen dioxide radical ($^\bullet NO_2$) (Figure 1.1A). $ONOO^-$ can react with carbon dioxide to generate nitrosoperoxycarbonate ($ONOOCO_2^-$) that oxidatively damages nucleobases while $ONOO^-$ mostly performs oxidative damage to 2'-deoxyribose in DNA [11, 12]. Nitrosative deamination can also occur to biomolecules by nitrous anhydride (N_2O_3) - a nitrosating agent derived from oxidation of $NO^\bullet$ (Figure 1.1A) [13]. Moreover, halogenation mediated by hypochlorous acid (HOCl) damages biomolecules when myeloperoxidase in neutrophils catalyzes the reaction of H_2O_2 with Cl^- (Figure 1.1A) [14].

1.4 Nucleic acid damage products as biomarkers of inflammation and oxidative stress

ROS and RNS generated from innate immune cells during inflammation and oxidative stress can damage all types of biomolecules from carbohydrates, lipids, and proteins to nucleic acids (Figure 1.1) [7]. A lot of effort has been expended to develop DNA damage products as biomarkers

[15-19], which is reasonable given that DNA damage fulfills the three critical criteria for a biomarker. First, nucleic acid damage products reflect the full spectrum of damage chemistries at sites of inflammation: oxidation, nitration, nitrosative deamination, halogenation and indirect alkylation damage from inflammation-derived electrophiles [20-24]. Second, DNA damage causes mutations [25-27] that represent a mechanistic link between inflammation and cancer [28]. Third, the level of DNA damage increases according to the increase of the inflammation and oxidative stress, so there is a clear dose-response relationship. Given these features, it is important to review the variety of DNA lesions arising during inflammation and oxidative stress.

1.4.1 Direct DNA damage caused by ROS and RNS

a) Oxidation/nitration. The chemistries of DNA damage caused by oxidation and nitration are mainly mediated by $OH^{\bullet}$, $ONOO^-$ and $ONOOCO_2^-$ [29, 30]. Guanine is the main target for the oxidation due to its lowest oxidation potential compared to other nucleobases [31-33]. Unlike so-called one-electron oxidants, $OH^{\bullet}$ adds by nucleophilic attack at C4 (60%), C5 (15%) and C8 (25%) positions of guanine [34]. As to the C4-OH and C5-OH radical adduct, subsequent dehydration yields a guanine radical, eventually giving rise to oxazolone (Figure 1.1). The C8-OH adduct radical can be either oxidized to form 8-oxoguanine or reduced to generate formamidopyrimidine, 2,6-diamino-4-hydroxy-5-foramidopyrimidine (FapyG) (Figure 1.1) [34]. Owing to lower oxidation potential of 8-oxoG vs. G (+0.74 V vs. +1.29 V) [33], 8-oxoG is 1000-times more susceptible to further oxidation and yields various secondary lesions (Figure 1.1). When $ONOO^-$ combines with CO_2 to generate $ONOOCO_2^-$, the damage spectrum shifts from oxidizing 2'-deoxyribose to nitrating and oxidizing nucleobases of DNA [11, 12]. $ONOOCO_2^-$ reacts with dG to form 5-guanidino-4-nitroimidazole (NitroIm), 2,2-diamino-4-((2-deoxypentofuranosyl)amino)-5(2H)-oxazolone (oxazolone; OZ), and 8-nitrodG (Figure 1.1) [34]. It was shown by Tannenbaum and coworkers that 8-nitrodG can react with $ONOOCO_2^-$ to generate 8-oxodG [35].

In addition to nucleobases, 2'-deoxyribose can be oxidized by $OH^{\bullet}$ radical generated by γ-radiation and Fe(II)/EDTA, or by other oxidizing agents, resulting in strand breaks and abasic sites (Figure 1.2) [36, 37]. The oxidation takes place via homolytic cleavage through H-atom abstraction from any of the five carbons in 2'-deoxyribose [38]. The spectrum of oxidation products is unique for individual carbons according to the mechanism of the cleavage. For C1'-oxidation, the major product is 2'-deoxyribonolactone with the subsequent elimination to form 5-methylene-2(5H)-furanone (Figure 1.2) [39, 40]. Model oxidants for the study of 1'-oxidation were identified in the oxidizing agent bis(1,10-phenanthroline)copper (I) and the enediyne antibiotic family [41]. C1'-oxidation products are believed to be minor because H1' is embedded in the minor groove of B-DNA, which makes it inaccessible to solvent-borne oxidants [42]. As to C2'-oxidation, the product from the damage is an erythrose abasic site, as occurs in γ-irradiated DNA, with the erythrose abasic site subject to elimination to form 3'-phosphoglycolaldehyde (Figure 1.2) [43]. The C2' position is also not a major target for oxidation due to solvent inaccessibility [42]. Oxidation at C3' at the major groove of DNA yields 3'-phosphoglycolaldehyde and base propenoate (Figure 1.2) under aerobic conditions, while the 3'-radical gives rise to 2-methylene-3(2H)-furanone under anaerobic conditions [44]. When C4' oxidation occurs by bleomycin-Fe(II) (Figure 1.2), base propenal and 3'-phosphoglycolate are generated under high oxygen concentrations, but the 2-deoxypentos-4-ulose abasic site is formed at low oxygen conditions [45, 46]. However, if the damage is induced by γ-radiation, malondialdehyde, free bases and 3'-phosphoglycolaldehyde are generated [46]. The different product spectrum generated by bleomycin-Fe(II) and γ-radiation is explained by the mechanism that bleomycin remains bound to DNA close to the initial C4' radical after the homolytic cleavage occurs and it participates in subsequent reactions that lead to either base propenal or free base and malondialdehyde [46, 47]. C4' oxidation represents ~20% of oxidation events occurring at 2'-deoxyribose [48]. C5' oxidation in DNA occurs with Fe(II)/EDTA, γ-radiation, enediyne antibiotics, porphyrin-chelated metals, and perhydroxyl radical [49, 50]. The

product spectrum includes 5'-formylphosphate, 5'-(2-phosphoryl-1,4-dioxobutane) and nucleoside 5'-aldehyde (Figure 1.2) [51, 52]. When heated, 5'-(2-phosphoryl-1,4-dioxobutane) and nucleoside 5'-aldehyde undergo elimination to generate *trans*-2-butene-1,4-dial and furfural, respectively [51, 52]. C5' oxidation of 2'-deoxyribose is considered to be the major result of DNA oxidation, due to the highly solvent accessible hydrogen atoms [53].

An important feature of 2'-deoxyribose oxidation products is that they can react further with DNA and proteins to form adducts. The damage products generated from 2'-deoxyribose oxidation are highly electrophilic such as α,β-unsaturated species (i.e., base propenals, *trans*-2-butene-1,4-dial) and aldehyde-containing compounds (i.e., formyl phosphate, furfural, methylene furanone). These products can secondarily modify nucleophiles in cells such as proteins and nucleobases in DNA and RNA [52, 54-60].

b) Nitrosative deamination and halogenation. Deamination of nucleobases in DNA can occur by a variety of chemical mechanisms, including treatment with sodium nitrite in an acidic conditions to generate nitrous acid *in situ* for a diazotization reaction. *In vivo*, N_2O_3 is the major nitrosating agent that will eventually deaminate the primary exocyclic amines of the nucleobases to yield hydroxyl groups. Deamination by N_2O_3 converts adenine to inosine (hypoxanthine), cytosine to uracil, 5-methylcytosine to thymine, and guanine to both xanthine and oxanosine [13, 61]. Halogenation of DNA is a unique damage class associated with myeloperoxidase in neutrophils. The resulting generation of HOCl can cause lipid peroxidation, halogenation of DNA bases (e.g., 5-chlorocytosine) and cross-linking of DNA and protein [14, 62-64].

1.4.2 Indirect DNA damage by inflammation-derived electrophiles

Alkylation of DNA by Inflammation-induced electrophiles represents another important class of DNA lesions. Perhaps the most well studied example of this phenomenon is lipid peroxidation, which occurs with polyunsaturated fatty acids (PUFA) (e.g., arachidonic acid), and

gives rise to several types of α, β-unsaturated carbonyl species [65-67]. PUFA are vulnerable to oxidation by ROS or COXs to generate a host of reactive electrophiles, such as malondialdehyde (MDA), acrolein, crotonaldehyde, 4-hydroxy-2-nonenal (4-HNE) and 4-oxo-2-nonenal (4-ONE) [68-70]. These products react with DNA to form an even more complicated set of adducts. For example, MDA reacts with dG to generate a propano adduct pyrimido[1,2-α]purin-10(3H)one, M_1dG, while it can also react with dA to form N^6-(3-oxopropenyl)deoxyadenosine (M_1dA), albeit at levels 20% of M_1dG generation [71]. When dC reacts with MDA, N^4-(3-oxopropenyl)deoxycytidine (M_1dC) is formed in trace amounts. Unlike M_1dG, M_1dA and M_1dC are generated without further cyclization [72, 73].

4-HNE is generated at very high concentrations (10's of μM) in response to oxidative stresses [74] and it reacts with dG to form a propano adduct by a Michael addition reaction at N^2-position. When the double bond of 4-HNE is oxidized to form an oxirane ring, the epoxy 4-HNE can react with dG, dA and dC to yield unsubstituted etheno adducts. 4-ONE is a another major oxidation product of PUFA and reacts with dG and dC to form heptanone-substituted etheno adducts [75].

The generation of these DNA adducts from the adventitious electrophiles arising during inflammation illustrates an important concept that will be developed in the second goal of this thesis: to explore the formation of DNA adducts from shifts in normal intermediary metabolism caused by environmental changes. Both adventitious electrophiles and physiological electrophiles lead to the same endpoint – potentially toxic and mutagenic DNA damage.

1.5 Choice of sampling compartments for biomarker analysis

Having identified candidate biomarkers among the DNA damage products, the next step is to choose a sampling compartment for the biomarker assay. The major consideration here is the clinical utility of the sampling compartment. For example, while DNA damage products have been quantified in easily accessible urine, blood and feces samples, there has been little effort to

determine if and how the damage products are altered en route from the site of formation to the urine or feces. Most studies have employed sampling various tissues, based on necropsy in animal models or biopsy or surgical discards in human studies. This process obviously poses the challenge of invasiveness in terms of clinical utility, with very limited amounts of human tissues available for large scale studies. Moreover, the steady-state level of the DNA adducts is not constant throughout a tissue [2]. Therefore, we must consider other clinically accessible compartments for biomarker sampling, such as urine, blood and feces.

There has been significant attempt to develop nucleic acid biomarkers in urine due to the non-invasive nature of the sampling. 8-oxodG [76-78] and etheno adducts [79-81] have been characterized in urine by HPLC pre-purification with isotope dilution method in LC-MS/MS and GC-MS analysis, as well as post-labeling methods using ^{32}P. However, one of the major factors ignored in the above analysis is the metabolism of the adduct. Biotransformation of the adduct may play a pivotal role in converting the chemical structure of the adduct before its excretion into urine or feces, or its appearance in the general blood circulation. The metabolism of the DNA adduct is discussed below in the next section in this thesis.

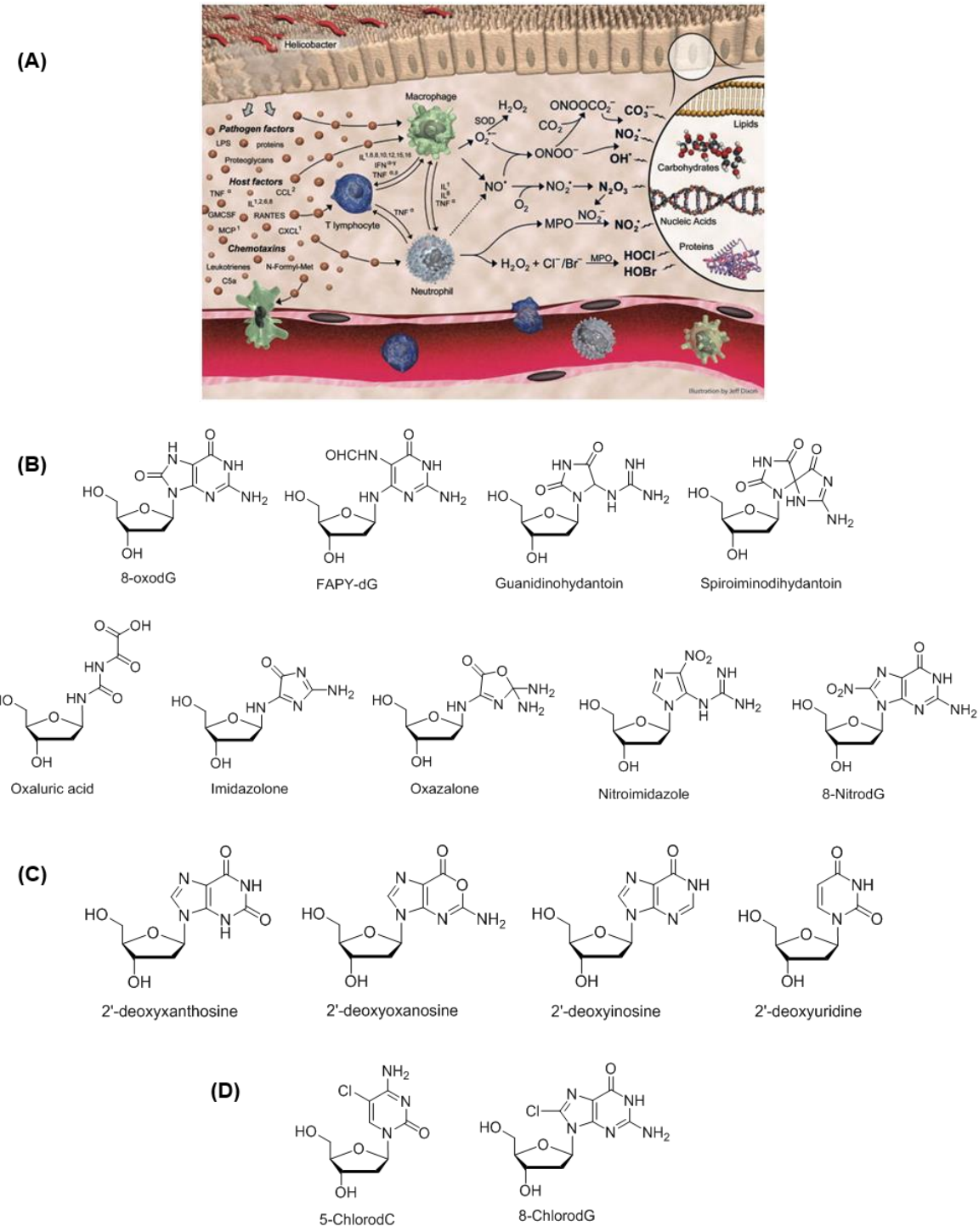


Figure 1.1. ROS and RNS during inflammation and oxidative stress causes DNA damage**.** (A) ROS and RNS generated by macrophages and neutrophils during inflammation. (B) – (D) Direct DNA damage products from oxidation, nitration, deamination and halogenation; (B) Oxidative damage products derived from the reaction of dG with ROS and RNS, $OH^{\bullet}$ and $ONOOCO_2$-; (C) Nitrosative deamination products from dA, dG and dC; (D) Halogenated products from dC and dG

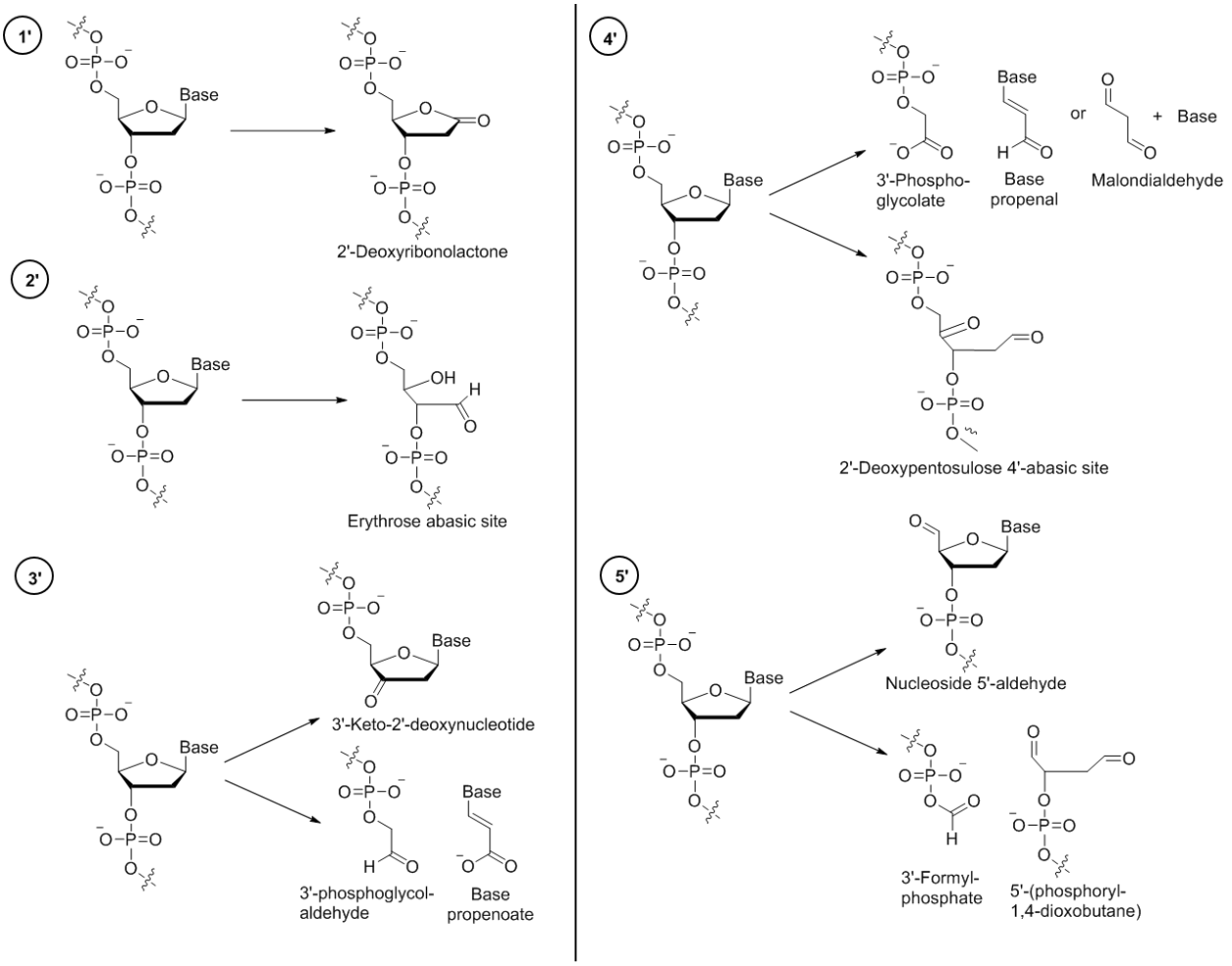


Figure 1.2. Damage products from 2'-deoxyribose oxidation at C1' to C5'.

1.6 Metabolic fates of DNA damage products

While there has been significant effort made to develop DNA damage products as biomarkers, there has been little effort made to understand the role that metabolism of the damage product may play altering candidate DNA damage products as biomarkers. Only a few studies have assessed the biotransformational fate of DNA damage products, including the *in vitro* metabolism of etheno adducts, and the *in vivo* metabolism of M_1dG and malondialdehyde as described here.

1.6.1 Metabolism of DNA nucleobases - etheno adducts and M_1dG

a) Etheno adducts. Etheno adducts consist of the alkylation products from the reaction of DNA with electrophilic species mostly derived from lipid peroxidation [65-67]. The structural spectrum of etheno adducts includes unsubstituted and substituted products (Figure 1.3A). As potential biomarkers, the quantity of unsubstituted etheno adducts excreted urine strongly correlates with human disease and exposure to oxidative stress [18, 82]. However, little attempt has been made to study the metabolic fate of etheno adducts. One study done by Marnett and co-workers, using rat liver cytosolic reactions, revealed that unsubstituted and substituted etheno adducts are subject to glycosylation and oxidation of the free bases at C2 [83] (Figure 1.3A and 1.3B). This raises concern that the previous analyses of etheno adducts in urine underestimated the actual amount of the adducts and supports the idea that metabolic pathways playing a crucial roles in altering the structure of the adduct should be considered when measuring the level of the adduct in urine.

b) The propano adduct, M_1dG. M_1dG is generated from the reaction of dG with either base propenals or malondialdehyde and is mutagenic by causing base pair substitutions ($M_1dG \rightarrow T$ and $M_1dG \rightarrow A$) and frameshift mutations in CpG rich regions [84]. The metabolic fate of M_1dG has been extensively studied by the Marnett group. The idea of M_1dG being subject to metabolism was originally derived from analysis of human urine, where the level of M_1dG excreted in the urine is extremely low (10–20 fmol/kg per 24 h) when correlated the amount of malondialdehyde generated from lipid peroxidation per day [85]. The metabolism of M_1dG was studied by exogenously administering [^{14}C] and [^{3}H]-labeled M1dG intravenously into Sprague-Dawley rats [86, 87]. M_1dG was rapidly distributed into tissues with the half-life ($t_{1/2}$) of 10 min [88]. The research group could detect a major oxidation product of M_1dG, 6-oxoM_1dG, catalyzed by xanthine oxidase (XO). The oxidation of M_1dG to 6-oxoM_1dG was confirmed by adding xanthine oxidase inhibitors, including menadione and allopurinol, into rat liver cytosol extracts [88]. 6-oxoM_1dG is

excreted in urine and bile with the recovery 25% and 70% relative to the injected dose of M_1dG, respectively [87]. This example serves as the motivation for the studies presented in this thesis, both for the concept of metabolism in biomarker development and for the focus on a "metabolic" product of 2-deoxyribose oxidation in DNA: base propenals.

1.6.2 Metabolism of 2-deoxyribose oxidation products

Even fewer studies have addressed the metabolic fate of DNA damage products involving the sugar moiety. For example, one study addressed the investigation of malondialdehyde metabolites. However, a more extensive study of an analog of a 2-deoxyribose oxidation product provides insights into the potential fates of these electrophiles. Peterson and coworkers have studied furan metabolism, which gives rise to *cis*-2-butene-1,4-dial - the geometrical isomer of *trans*-2-butene-1,4-dial derived from 5'-oxidation of 2'-deoxyribose. Therefore, the metabolic fates of malondialdehyde and *cis*-2-butene-1,4-dial are reviewed here.

a) cis-*2-Butene-1,4-dial.* β-Elimination of 5'-(2-phosphoryl-1,4-dioxobutane) gives rise to *trans*-2-butene-1,4-dial [52]. Studies pertaining to *trans*-2-butene-1,4-dial have been limited to the formation of bicyclic adducts with dG, dA and dC, but did not address the metabolism of this compound [57-60]. However, its geometrical isomer, *cis*-2-butene-1,4-dial, is an α,β-unsaturated dialdehyde derived from metabolism of furan by cytochrome P450s [89, 90]. *cis*-2-Butene-1,4-dial is believed to be the reactive intermediate that links with furan hepatotoxicity due to GSH depletion in freshly prepared hepatocytes. *In vitro*, *cis*-2-butene-1,4-dial reacts with model nucleophiles, such as N-acetyl lysine, N-acetyl cysteine and GSH [91]. As to the reaction of *cis*-2-butene-1,4-dial with N-acetyl lysine, the aldehyde condenses with the amino group of lysine to form a pyrrole-based adduct, as shown Figure 1.4A. The formation of these pyrrole-based GSH adducts derived from intramolecular Schiff base formation with the γ-amino group of glutamate suggests that *cis*-2-butene-1,4-dial is prone to crosslinking thiols with amino groups. The thiol crosslink with amino

groups also explains why GSH, but not N-acetylcysteine, can help prevent the protein adduct formation. Therefore, it was believed that GSH plays a vital role in detoxifying *cis*-2-butene-1,4-dial by protecting it from modifying cellular nucleophiles such as DNA and proteins. *In vivo*, *cis*-2-butene-1,4-dial was mainly metabolized to CO_2 by means of oxidation, with 26% of an 8 mg/kg injected dose in rats [92]. In addition, GSH-derived mercapturate metabolites were also identified in rat urine and bile [93-98]. Further oxidized mercapturate metabolites in the form of sulfoxide were also detected *in vivo*. It has recently been shown that polyamines, such as spermidine and putrescine, can also form the adduct with GSH-BDA [97]. GSH may not be able to completely detoxify BDA by forming mono- and bis-GSH adduct via S-conjugation addition and intramolecular Schiff base formation from γ-amino glutamate. It was discussed that when the GSH-BDA adduct forms, it may diffuse across the membrane and alkylate amine groups outside the site of its formation. Therefore, it has been thought that GSH-BDA probably contributes to the toxic properties of furan.

b) Malondialdehyde. Malondialdehyde is an aldehyde-containing damage product derived from DNA damage or lipid peroxidation. The pK_a value of the enolic OH group is 4.5 [74], so that, at physiological pH, malondialdehyde forms an enolate anion. The presence of the negative charge at the enolic OH makes the β-carbon of malondialdehyde not as electrophilic as typical α,β-unsaturated aldehydes owing to the delocalization of electrons from the enolate anion. Several pieces of evidence showed that malondialdehyde was more reactive at acidic pH. First, β-hydroxyacrolein, the enol form of malondialdehyde at acidic pH, forms 1:1 and 1:2 adducts with amino acids [99] and 3:2 adduct with cysteine [100] (Figure 1.5B). Second, when reacting with nucleophiles at physiological pH such as glycine, glutathione, cysteine and bovine serum albumin, malondialdehyde shows much slower kinetic reaction than does 4-hydroxynonenal [101]. The *in vitro* metabolism of malondialdehyde was studied by incubation with rat liver homogenates. An

aldehyde dehydrogenase in rat liver mitochondria was found to be responsible for oxidative metabolism of malondialdehyde to CO_2, with K_m = 0.5 mM, V_{max} = 9.3 mmol/min per mg protein for CO_2 formation [102]. In 1985, Marnett and co-workers studied the metabolism of malondialdehyde in Swiss mice using [^{14}C]-malondialdehyde injected into the animal intravenously [103]. Half of the malondialdehyde was metabolized to CO_2 as the major pathway. There are two pathways contributing to the conversion of malondialdehyde to CO_2 (Figure 1.5C). In the first pathway, one of the aldehydes in malondialdehyde is oxidized to malonic semialdehyde, which is then transformed to acetyl CoA and β-hydroxy propionate, and finally degraded to CO_2. This pathway was confirmed by the incorporation of [^{14}C]-labeled β-alanine due to its deamination to malonic semialdehyde and rapid metabolism to CO_2. The second and minor pathway contributing to the degradation of malondialdehyde to CO_2 is that malondialdehyde is oxidized to malonic semialdehyde and subsequently esterified by coenzyme A to generate malonic semialdehyde CoA. Then, malonic semialdehyde CoA is split to β-hydroxy propionyl CoA and malonyl CoA. The latter compound can be metabolized to acetyl CoA and CO_2. However, the major portion of malonyl CoA is subjected to incorporation with acetyl CoA to generate fatty acids in a biosynthesis pathway. This explains why this pathway would contribute less to the generation of CO_2.

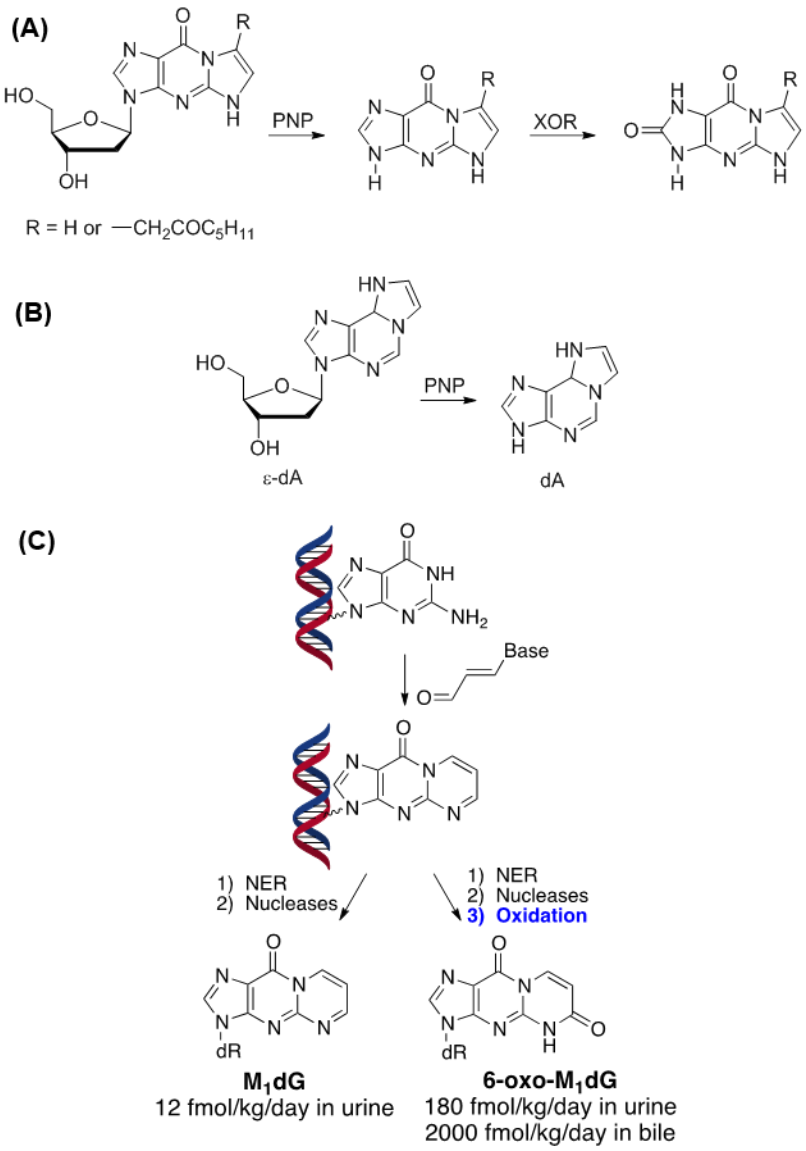


Figure 1.3. Metabolic conversion of etheno and propano adducts. (A) Metabolism of unsubstituted and substituted etheno dG in rat liver cytosol. (B) Metabolism of unsubstituted etheno dA in rat liver cytosol. (C) Metabolic fate of the propano adduct, M1dG.

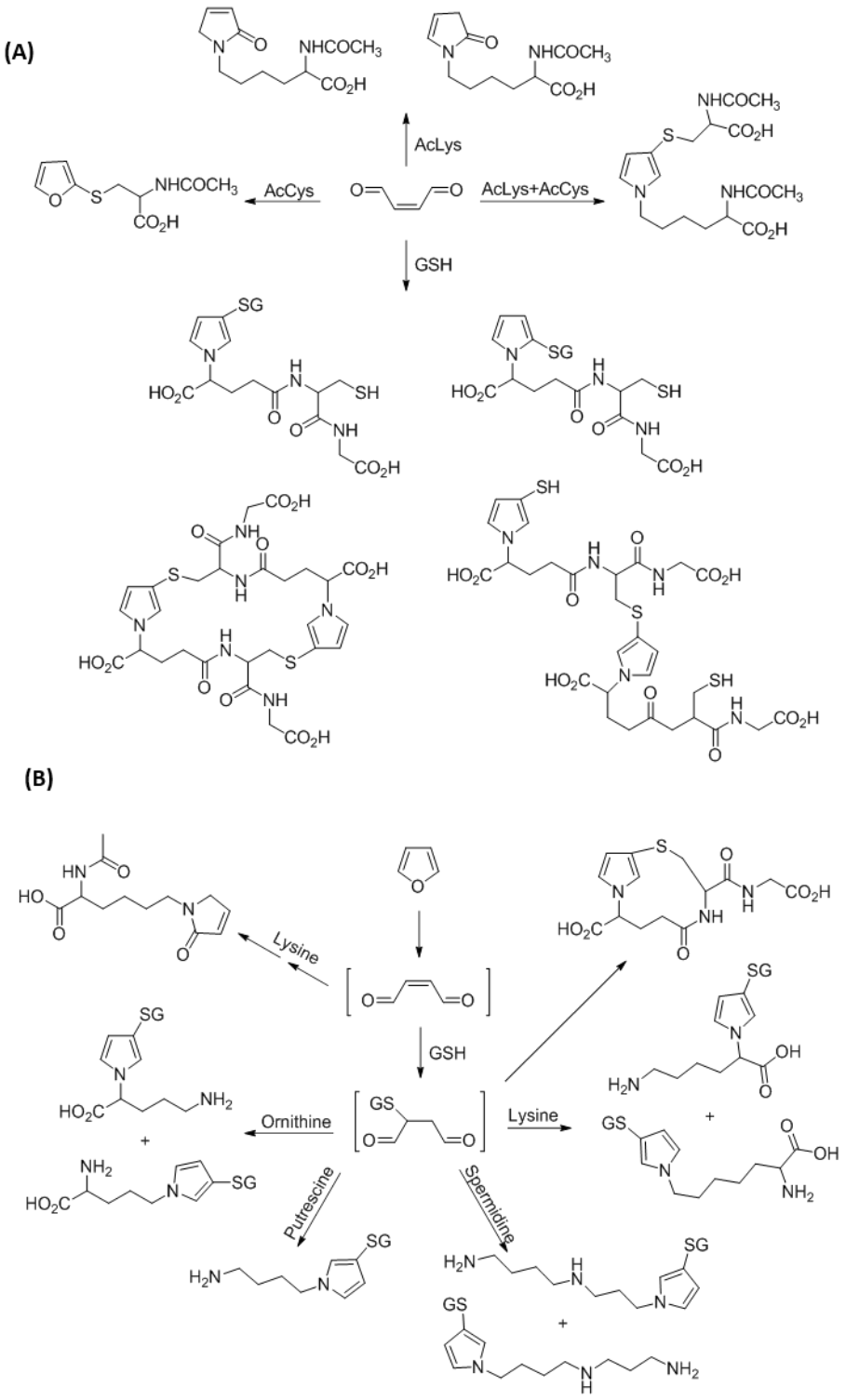


Figure 1.4. Metabolism of furan. (A) Chemical reaction of *cis*-2-butene-1,4-dial with model nucleophiles involving AcLys, AcCys and GSH. (B) *In vivo* metabolic fates of *cis*-2-butene-1,4-dial with several nucleophiles: GSH, lysine and polyamines.

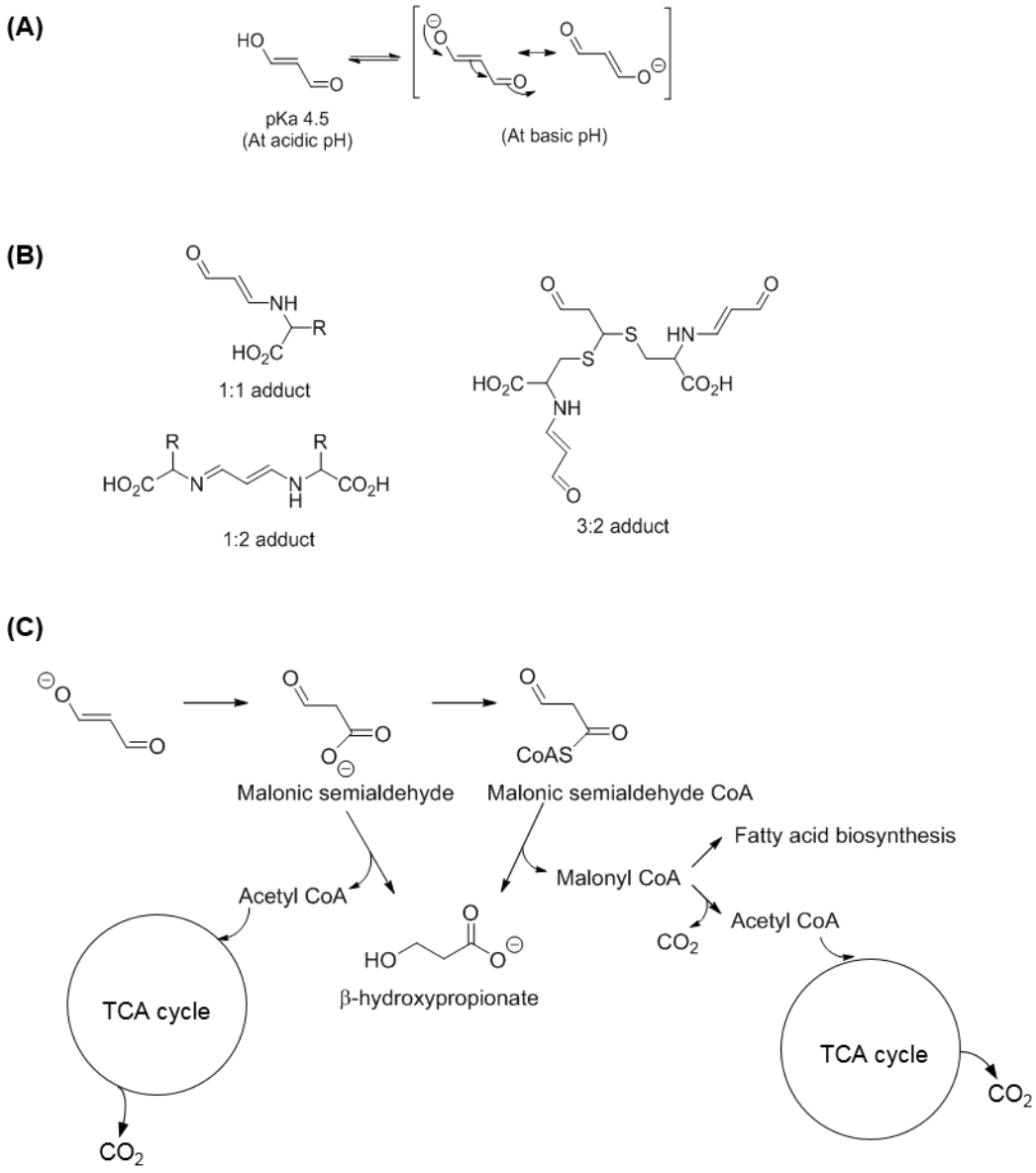


Figure 1.5. Physicochemical properties of MDA and its metabolism. (A) Chemical structure of MDA at acidic and basic pH. (B) MDA adduct formation with amino acids (1:1 and 1:2 adducts) and with cysteine (3:2 adduct). (C) Metabolic fates of MDA *in vivo*.

1.7 The problem driving this thesis: DNA damage metabolites as biomarkers

The goal of this part of the thesis is to explore the utility of DNA damage metabolites as biomarkers. DNA damage products reflect the full spectrum of pathological inflammation chemistries, including oxidation, nitration, deamination, halogenation and alkylation by inflammation-derived electrophiles. DNA damage products also provide mechanistic correlations between inflammation and the pathophysiology of diseases such as cancer, neurodegenerative disease and cardiovascular disease. In addition, there is a dose-response relationship between the level of DNA damage products and that of oxidative stress. However, sampling of DNA damage products as biomarkers typically involves tissue biopsies. This poses clinical challenges in terms of (1) the invasiveness of the tissue sampling, (2) the unreliable quantitation of the adduct owing to the inconsistency of the steady-state level of the adduct in cells across the tissue, and (3) the limited quantities of human tissues. Therefore, there has been significant interest in developing biomarkers from other clinically accessible compartments such as blood, urine and feces. In spite of this interest, the metabolism of the DNA adducts has not been considered as a crucial factor that contributes to the conversion of the chemical structure of the adduct before excretion from the body.

In this thesis, the metabolic fate of base propenals, a DNA damage product from 4'-oxidation of 2'-deoxyribose, is investigated using the reaction model with GSH - the major antioxidant in cells. The *in vitro* reaction of a base propenal and GSH is discussed in Chapter 2 followed by development of LC-MS/MS methods for *in vivo* targeted analyses in Chapter 3. The LC-MS/MS analytical platform is then applied to urinary and biliary analysis *in vivo* in Chapter 4. The results of the investigation of the metabolic fate of a base propenal described in Chapters 2 to 4 supports the concept of metabolic conversion of DNA damage adducts before excretion, which subsequently results in a new approach to biomarker development.

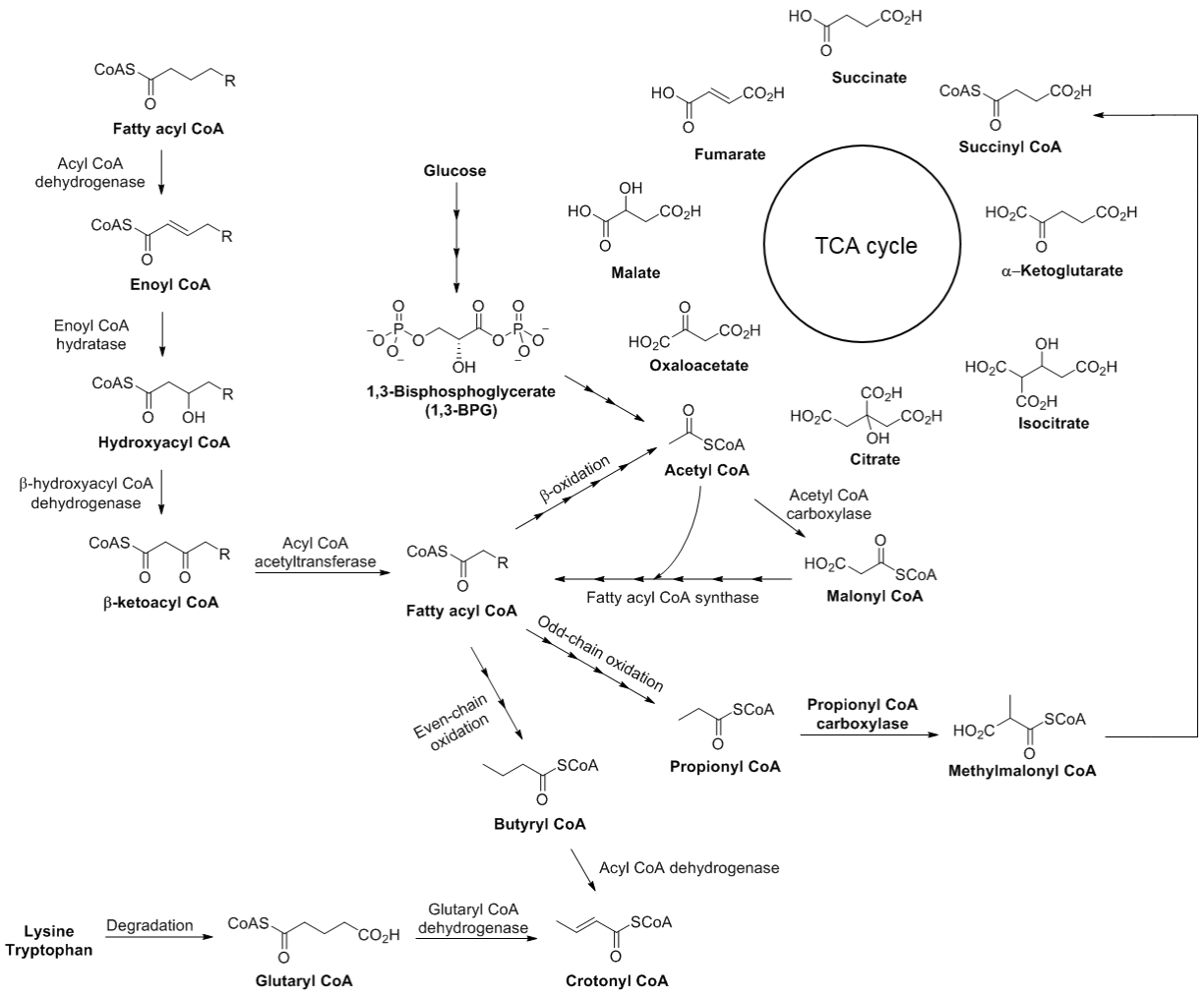


Figure 1.6. Various types of electrophilic species generated from primary metabolism; thioester, phosphate ester, α,β-unsaturated carbonyls.

1.8 DNA lesions caused by electrophilic metabolites from intermediary metabolism

The second half of this thesis addresses a direct complement to the concept of the metabolism of DNA lesions: the formation of DNA lesions as a result of metabolism. Primary metabolism is indispensible in terms of providing cellular energy through oxidation of biomolecules including carbohydrates, lipids and proteins. However, many of the intermediates generated during the oxidative metabolism are electrophilic, such as acetyl CoA from glycolysis and β-oxidation, crotonyl CoA from β-oxidation, and succinyl CoA from the TCA cycle, among others. Several studies have shown that these electrophilic metabolites can react with nucleophilic side chains of lysine and cysteine in proteins [104-109]. This part of the thesis reviews the chemical biology of those protein modifications as the foundation for our goal of characterizing DNA adducts derived from the reaction with electrophilic metabolites from the primary metabolism.

1.8.1 Acylation of lysine by bisphosphoglycerate (1,3-BPG)

1,3-BPG is an intermediate in glycolysis as a product of oxidation of glyceraldehyde 3-phosphate by glyceraldehyde 3-phosphate dehydrogenase (GAPDH). The physiological acylation of lysine by 1,3-BPG has been studied by Cravatt and coworkers [104]. 1,3-BPG is unstable and isomerizes to 2,3-BPG, which reacts non-covalently with hemoglobin and regulates the protein function in terms of oxygen binding affinity known as “Bohr effect” [110]. On the other hand, 1,3-BPG reacts convalently with lysine residues in glycolytic enzymes [104] due to its reactive phosphate ester group caused by the anhydride bond linking between carbonyl and phosphoryl group [111]. 1,3-BPG-modfied lysine (pgK) was detected in commercial GAPDH preparations, which suggests that pgK modifies proteins constitutively [104]. Cravatt and co-workers used a phosphoproteomic platform to identify and quantify pgK-modified proteins. They observed that 1, 3-BPG covalently modifies lysines in several proteins such as metabolic enzymes in glycolysis and TCA cycle, nuclear proteins, transporters, etc.

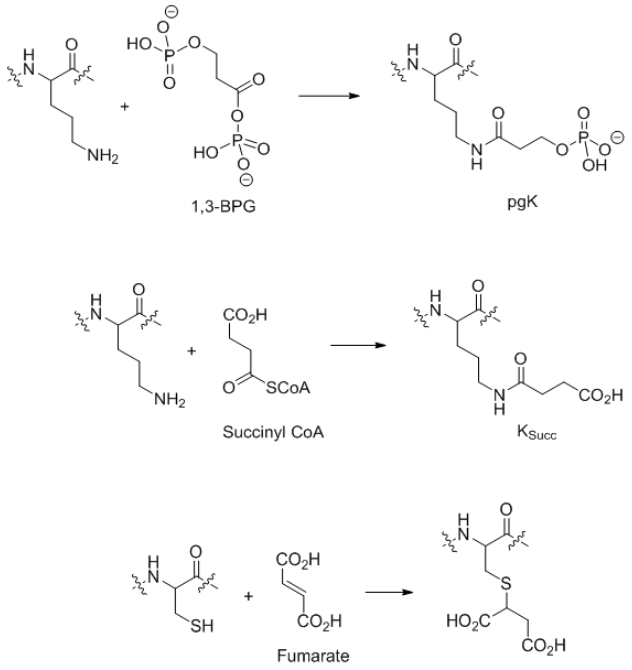


Figure 1.7. Nucleophilic addition of electrophiles, 1,3-BPG, succinyl CoA and fumarate to generate protein modifications: phosphoglyceryllysine, succinyllysine and cysteine succination.

The pgK modification was detected in catalytic sites and regulatory sites of those enzymes, suggesting that a biological function of this novel type of covalent modification. For example, two key regulatory enzymes in glycolysis, glyceralehyde 3-phosphate dehydrogenase and pyruvate kinase 2, are subject to pgK modification when the concentration of glucose is increased, resulting in higher concentration of 1, 3-BPG. Furthermore, enolase 1 is also modified by 1,3-BPG and is presumably responsible for modifying enzymes catalysis of three-carbon metabolites in glycolytic pathway. It is thought that 1,3-BPG down-regulates those enzymes because, after modification by 1,3-BPG, for example, GAPDH exhibits higher K_M for its substrate (3-phosphoglyceraldehyde). The structure of the modification has the negative charge of a phosphate, thus it is thought that 1,3-BPG modification of lysine affects the structure of enzymes and regulates their function. This impairment of the enzyme might be another regulatory mechanism to shift the metabolic pathway to an alternative shunt such as the pentose phosphate pathway. However, there is no conclusive evidence showing that there is an enzyme that removes phosphoglycerate from lysine to close the loop in terms of biological regulation.

1.8.2 Acylation of lysine by succinyl CoA

Lysine succinylation (K_{succ}) was first observed in the modification of lysine residues in homoserine succinyl transferase [112]. However, this phenomenon was recently observed as a post-translational modification in two studies done by Zhao and co-workers, which underscores the importance of this modification that appears in prokaryote and eukaryote cells and mammalian tissues [105, 113, 114]. It was concluded that succinyl CoA is the donor for lysine succinylation in enzyme-independent reactions [106, 107]. This was based on two observations, that a decrease in lysine succinylation occurs when α-ketoglutarate dehydrogenase is knocked out, leading to the depletion of succinyl CoA [106], and that when succinyl CoA ligase is knocked out, lysine succinylation is more abundant due to the accumulation of succinyl CoA in the pool. K_{Succ} was globally detected in cells – histone and non-histone proteins - reflecting the importance of this post-translational modification [115], but the biological role of Ksucc in histones has not been studied. K_{succ} was found in numerous types of enzymes, including isocitrate dehydrogenase II, serine hydroxymethyl transferase, glyceraldehyde 3-phosphate dehydrogenase, GLUD1, malate dehydrogenase II, citrate synthase, carbamoyl phosphate synthetase, HMGCoA synthase II, thiosulfate sulfurtransferase, aspartate aminotransferase [106]. K_{succ} was shown to regulate the carbon flow in energy metabolism in particular glycolysis, TCA cycle, beta-oxidation of fatty acids and ketone bodies [116]. The level of K_{Succ} is regulated by a deacylase SIRT5. which was shown to suppress the function of pyruvate dehydrogenase complex (PDC) and succinate dehydrogenase (PDC) [116]. In the same study, β-oxidation and ketone bodies was decreased when the SIRT5 gene was deleted. Also, the function of K_{Succ} pertaining to lipid metabolism was confirmed by the X-ray crystal structures of HMGCoA synthase II in that lysine succinylation was found in the vicinity of active site (K83, K310, K350, K354 and K358) [116] along with mutation of K83E, K306E and K310E leading to the loss of function of HMGCoA synthase II.

1.8.3 Cysteine succination

Succination involves the 1,4-nucleophilic addition of fumarate to the thiol in cysteine, which differs from lysine succinylation involving the nucleophilic acyl substitution occurring as 1,2-addition at the carbonyl carbon of thioesters. The purpose of this "succination" versus "succinylation" nomenclature is to differentiate these two mechanisms. Cysteine succination was accidentally discovered in the analysis of S-(carboxymethyl)cysteine, the product generated from the reaction between glyoxal or glycolaldehyde and cysteine [117]. This adduct was then quantified with N-ethyl maleimide, which generates S-(2-succinyl)cysteine after acid hydrolysis. However, without N-ethyl maleimide treatment, S-(2-succinyl)cysteine was detected. Elevated level of cysteine succination were found in two pathological conditions: fumarate hydratase (FH) deficiency and diabetes [118-120]. In FH-deficient cancer cells, fumarate accumulates due to loss of function of FH responsible for conversion of fumarate to malate. The increase in fumarate concentration strongly correlates with the development of tumors [121]. One model to explain the elevated concentration of fumarate causing hereditary leiomyomatosis and renal cancer (HLRCC) is the dysregulation of KEAP1-NRF2 pathway [122]. When the mitochondrial fumarate concentration is increased by the loss of function of FH, fumarate leaks into the cytosol and is subjected to the attack by regulatory cysteine residues (cys151 and cys 288) in the KEAP1 protein. This succination modification to KEAP1 serves as a signal for ubiquitylation by the proteasome, which in turn stabilizes NRF2 transcription factor and leads to the upregulation of several genes regarding glutathione S-transferases (GSTs), aldo-keto reductases, pentose phosphate pathway (PPP), glutamine metabolism, and the PI3K-Akt pathway. PPP provides ribose and NADPH, which are critical for DNA replication and biosynthesis of macromolecules, respectively [110]. It was shown in cancer cells that glutamine is metabolized via the TCA cycle and used for cancer cell growth [123]. The PI3K/Akt pathway is one of the major contributors to cancer development because it promotes cancer cell growth, survival and metabolism.

1.8.4 Can glyoxylate modify cellular nucleophiles?

Glyoxylate is an aldehyde-containing two-carbon compound that is endogenously derived from the glyoxylate shunt, where isocitrate lyases (ICLs) catalyze the formation of glyoxylate through retro-aldo reaction with the mechanism of C-C bond cleavage of isocitrate to generate glyoxylate and succinate [110]. The glyoxylate cycle plays a very important role in plant seed germination and in positioning mycobacteria for the dormant state. The aldehyde of glyoxylate serves as an electrophilic functional group, which should be capable of modifying nucleoophilic sites in nucleic acids and proteins in the same manner as formaldehyde and acetaldehyde [124, 125]. In particular, in mycobacteria that rely heavily on the glyoxylate shunt during starvation in granulomas, the potential for a reaction between dG and glyoxylate was expected to be high, which served as the driving hypothesis for this portion of the thesis project. Here we explore the role of glyoxylate in mycobacterial physiology and tuberculosis.

1.9 Tuberculosis

Tuberculosis is the disease caused by the complex of genetically related mycobacteria including *M. tuberculosis, M africanum, M. bovis*, *M. microti, M. canettii, M. caprae, M. pinnipedii* and *M. mungi* [126]. The major source of human infection in tuberculosis is M. tuberculosis, which infects one-third population of the world. In 2013, WHO reported 10 million new TB cases and 3 million deaths. The reason that *M. tuberculosis* is one of the most successful human pathogens is its rigid, thick cell wall that regulates entry of drugs and other small molecules [127]. The cell envelope is composed of a polypeptide layer, a peptidoglycan layer, and free lipids. In addition, there is also a complex structure of polyketides, including mycolic acids, which are further modified by addition of functional molecules to form glycolipids, sulpholipids, phenolic lipids, etc. In addition, lipid metabolism in mycobacteria also plays a crucial role in the infection and persistence in hosts [128] as reflected by the fact that lipid metabolism genes account for 30% of the genes in the

mycobacterial genome, with 250 genes contributing to fatty acid metabolism and 39 genes encoding polyketide biosynthesis enzymes to generate the waxy coat in the cell wall [129].

1.10 Mycobacterial infection

Tuberculosis infection initially involves the respiratory tract and spreads through aerosols [130, 131]. When the bacilli reach the lung, they are phagocytized by alveolar macrophages, which are then activated and induce an inflammatory response by secreting cytokines in order to recruit lymphocytes to the site of infection. Instead of being destroyed by the host immune cells, mycobacteria can survive in phagosome under servere conditions of nutritional deprivation, hypoxia and low pH in the highly organized cellular structure of the immune-mediated granuloma. As a hallmark of tuberculosis infection, granulomas prevent mycobacteria from distributing to other tissues [132]. Under these stress conditions, mycobacteria reprogram their metabolism and shift to the non-replicative state of dormancy [133, 134], which represents a metabolically quiescent state with no colony forming ability when cultured on agar plates. Dormancy differs from the clinical state of "latency" that means an asymptomatic state arising from either a host immune-mediated control of small number of bacilli or a non-replicating state established by the bacilli [134]. The clinically latent mycobacteria are not perilous as long as they are suppressed by host immune system. However, mycobacteria can reverse their metabolism to the active state, with the reactivated or resuscitated mycobacteria breaking out of the granuloma [135]. The effect of resuscitation is often fatal. The impairment of the immune surveillance of the host occurs when the host encounters malnutrition, old age or HIV infection [135]. In this project, the formation of glyoxylate-derived DNA adducts are being studied as a result of the shift to the glyoxylate cycle in dormancy and as a possible source of the higher-than-expected mutation rate observed in dormant mycobacteria.

1.11 A shift of the carbon metabolic flux in dormant mycobacteria

Due to the hypoxic conditions and nutrient starvation inside the granuloma, *M. tuberculosis* reprograms its metabolism during the transition to dormancy [134]. Studies of the metabolic shifts in mycobacteria during dormancy have been mostly obtained from differential gene expression data in a variety of model of infection systems, including *M. tuberculosis*-infected mouse bone marrow macrophages, human macrophages and animals [136-140]. The data [141] show the upregulation of genes involved in the glyoxylate cycle, gluconeogenesis and fatty acid metabolism, including isocitrate lyase (aceA1), citrate synthase (gltA), PEP carboxykinase (pckA), pyruvate phosphate dikinase (ppdK), fructose-bisphosphate aldolase (fba), acyl CoA dehydrogenases (fadE28) and enoyl CoA hydratase (echA20), as well as the genes for uptake of glycerol-3-phosphate (ugp and glpD) and cholesterol (mce4 operon). This strongly suggests that the bacilli utilize host lipids as the major carbon source, with β-oxidation to convert even-chain fatty acids to acetyl CoA and conserve the energy using the glyoxylate shunt [142] by generating glyoxylate instead of losing it as CO_2. Then, glyoxylate is condensed with a molecule of acetyl CoA to form oxaloacetate that can be further converted to carbohydrate via gluconeogenesis. In this way, the bacilli can both conserve the energy by not losing CO_2 and convert lipids to carbohydrates during the dormant state [110]. Moreover, *M. tuberculosis* uses the methylcitrate cycle to detoxify propionyl CoA, arising from the metabolism of odd-chain fatty acids and cholesterol, to pyruvate - an important intermediate that can be recycled to glucose or a high energy compound like acetyl CoA [143]. Furthermore, a study by McKinney et al. indicates that the isocitrate lyase genes, responsible for metabolizing isocitrate lyase to glyoxylate, are required for *M. tuberculosis* virulence and persistence in a pulmonary-infection mouse model [128]. This implies that the glyoxylate shunt may not only play a role in mycobacterial survival during the stress, but also acts as a potential virulence factor for mycobacterial infection.

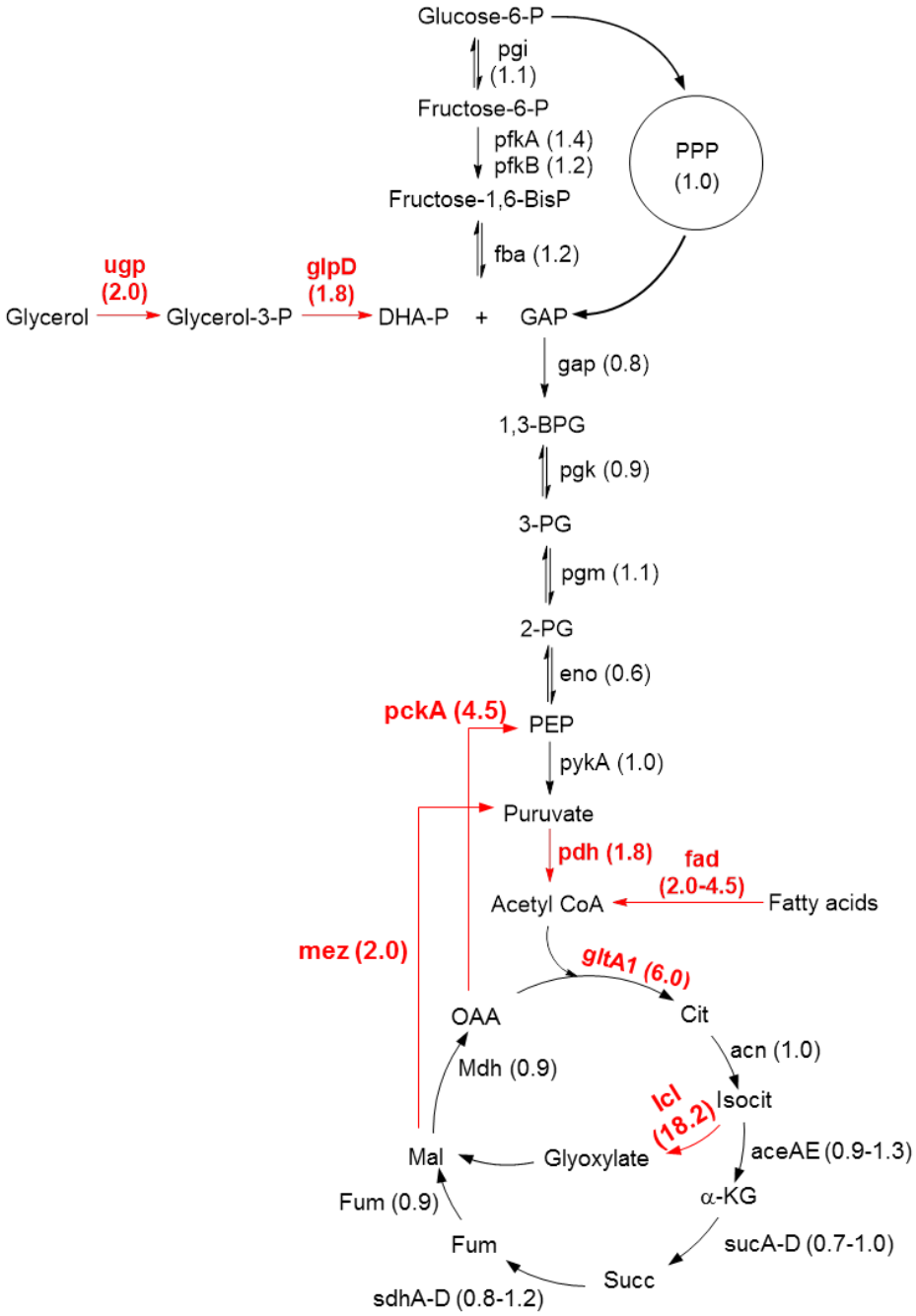


Figure 1.8. Model for *M. tuberculosis* metabolism based on differential gene expression data obtained from bacteria grown in bone-marrow-derived macrophages and 7H9 culture medium [136]. The number in the bracket represents the fold change of mRNA level in each gene of *M. tuberculosis* in bone-marrow-derived macrophages over in 7H9 culture medium.

1.12 The problem driving this portion of the thesis: DNA lesions caused by shifts in intermediary metabolism

Intermediary metabolism gives rise to several types of electrophilic species ranging from thioesters, phosphate esters, α, β-unsturated carbonyl compounds, and aldehyde-containing compounds. For example, it has recently been shown that acetyl CoA and succinyl CoA can act as electrophiles and undergo nucleophilic acyl substitution with the N_{ε}-amino group of lysine and the resulting modification exerts biological functions in terms of regulation of enzymes in central metabolic pathways [104-106]. Acetyl CoA and succinyl CoA were also shown to modify lysines in an enzyme-independent manner, facilitated by the basic pH in the mitochondrial matrix [107]. 1,3-Bisphosphoglycerate (1,3-BPG), a phosphate ester that is more electrophilic than thioesters due to the anhydride bond [111], can also nonenzymtically modify lysines in proteins and impose biological negative feedback to glycolytic enzymes via covalent modification of regulatory lysines [104]. Fumarate, an α, β-unsaturated dicarboxylic acid, can attack with regulatory cysteine residues in KEAP1 (Kelch-like ECH-associated protein 1) through conjugate addition and inactivates the function of KEAP1, which in turn stabilizes NRF2 (nuclear factor 2) transcription factor responsible for transcription of a lot of genes that promote cancer cell proliferation such as glutathione S-transferases (GSTs), pentose phosphate pathway (PPP) and PI3K/Akt signaling pathway [108, 109].

Glyoxylate is an aldehydic species derived from glyoxylate shunt found in plants, fungi, bacteria, archaea, protists, and nematodes. For example, during plant seedling, the organism utilizes the glyoxylate pathway to convert stored lipids to carbohydrates, a primary source for plant growth and development prior to plant growth and development [144-146]. Glyoxylate pathway also allows mycobacteria to survive in fatty acid or acetate-based nutrient and hypoxic conditions, since it conserves two carbon in the form of C2-glyoxylate instead of losing it as two molecules of CO_2 per turn that happens in the normoxic condition in the TCA cycle [110, 147, 148]. Two unique enzymes in the glyoxylate shunt are isocitrate lyase (ICL) and malate synthase (MS). ICL converts isocitrate

to glyoxylate and succinate by retro-aldol cleavage mechanism [149]. glyoxylate is further nucleophilically attacked by acetyl CoA through the aldol addition reaction to generate malate, with the catalysis by malate synthase [110]. Levels of ICL mRNA in *M. tuberculosis* grown in bone marrow-derived macrophages (BMM) were upregulated by 18-fold compared to cultured cells grown in standard Middlebrook 7H9 medium [141]. In addition, loss of the ICL1 and ICL2 genes reduces the persistence and virulence of *M. tuberculosis* during infection of mouse lungs [128]. Recently, the mutation rate of *M. tuberculosis* during dormant infections in the *Cynomolgus macaque* was found to be similar to that in acute infections of *M. tuberculosis* [150]. With the high upregulation of ICL transcripts in the glyoxylate shunt, the vital role of ICL for mycobacterium virulence, and the intrinsic electrophilicity of glyoxylate due to the aldehyde functional group, I therefore propose that glyoxylate can nucleophilically attack the amino group of 2'-deoxyguanosine and bring about a novel DNA lesion, N^2-CHMdG, in *M. smegmatis*.

References

1. Nick Fox, J.H.G., *Biomarkers and Surrogates.* NeuroRX, 2004. **1**(2): p. 181.
2. Dedon, P.C., et al., *Challenges in developing DNA and RNA biomarkers of inflammation.* Biomark Med, 2007. **1**(2): p. 293-312.
3. Blake, G.J. and P.M. Ridker, *Inflammatory bio-markers and cardiovascular risk prediction.* J Intern Med, 2002. **252**(4): p. 283-94.
4. Ridker, P.M., et al., *C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women.* N Engl J Med, 2000. **342**(12): p. 836-43.
5. Miller, E.C., *Some Current Perspectives on Chemical Carcinogenesis in Humans and Experimental-Animals - Presidential-Address.* Cancer Research, 1978. **38**(6): p. 1479-1496.
6. Das, A., et al., *Quantification of Aflatoxin B1 Produced by Aspergillus flavus MTCC 2798 in Rice Straw under Solid-State Fermentation.* Journal of Pure and Applied Microbiology, 2012. **6**(4): p. 1663-1669.
7. Lonkar, P. and P.C. Dedon, *Reactive species and DNA damage in chronic inflammation: reconciling chemical mechanisms and biological fates.* International Journal of Cancer, 2011. **128**(9): p. 1999-2009.
8. Marnett, L.J., *Inflammation and Cancer: Chemical Approaches to Mechanisms, Imaging, and Treatment.* Journal of Organic Chemistry, 2012. **77**(12): p. 5224-5238.
9. Turrens, J.F., *Mitochondrial formation of reactive oxygen species.* J Physiol, 2003. **552**(Pt 2): p. 335-44.
10. Marnett, L.J., J.N. Riggins, and J.D. West, *Endogenous generation of reactive oxidants and electrophiles and their reactions with DNA and protein.* J Clin Invest, 2003. **111**(5): p. 583-93.
11. Kennedy, L.J., et al., *Quantitation of 8-oxoguanine and strand breaks produced by four oxidizing agents.* Chem Res Toxicol, 1997. **10**(4): p. 386-92.
12. Burney, S., et al., *DNA damage in deoxynucleosides and oligonucleotides treated with peroxynitrite.* Chem Res Toxicol, 1999. **12**(6): p. 513-20.
13. Lucas, L.T., D. Gatehouse, and D.E.G. Shuker, *Modification of nucleotides and 2 '-deoxyguanosine by nitrosated indoles results in deamination, depurination and the formation of a novel product oxanine via a transnitrosation mechanism.* British Journal of Cancer, 1999. **81**(4): p. 578-578.
14. Sugiyama, S., et al., *Hypochlorous acid, a macrophage product, induces endothelial apoptosis and tissue factor expression - Involvement of myeloperoxidase-mediated oxidant in plaque erosion and thrombogenesis.* Arteriosclerosis Thrombosis and Vascular Biology, 2004. **24**(7): p. 1309-1314.
15. Cooke, M.S., et al., *Measurement and meaning of oxidatively modified DNA lesions in urine.* Cancer Epidemiology Biomarkers & Prevention, 2008. **17**(1): p. 3-14.
16. Nair, U., H. Bartsch, and J. Nair, *Lipid peroxidation-induced DNA damage in cancer-prone inflammatory diseases: A review of published adduct types and levels in humans.* Free Radical Biology and Medicine, 2007. **43**(8): p. 1109-1120.
17. Nair, L., et al., *High urinary excretion of lipid peroxidation-derived DNA damage in patients with cancer-prone liver diseases.* Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis, 2010. **683**(1-2): p. 23-28.
18. Dechakhamphu, S., et al., *Lipid Peroxidation and Etheno DNA Adducts in White Blood Cells of Liver Fluke-Infected Patients: Protection by Plasma alpha-Tocopherol and Praziquantel.* Cancer Epidemiology Biomarkers & Prevention, 2010. **19**(1): p. 310-318.
19. Chen, H.J.C. and C.F. Kao, *Effect of gender and cigarette smoking on urinary excretion of etheno DNA adducts in humans measured by isotope dilution gas chromatography/mass spectrometry.* Toxicology Letters, 2007. **169**(1): p. 72-81.

20. De Bont, R. and N. van Larebeke, *Endogenous DNA damage in humans: a review of quantitative data.* Mutagenesis, 2004. **19**(3): p. 169-85.
21. Epe, B., *Role of endogenous oxidative DNA damage in carcinogenesis: what can we learn from repair-deficient mice?* Biol Chem, 2002. **383**(3-4): p. 467-75.
22. Dedon, P.C. and S.R. Tannenbaum, *Reactive nitrogen species in the chemical biology of inflammation.* Arch Biochem Biophys, 2004. **423**(1): p. 12-22.
23. Marnett, L.J., *Oxyradicals and DNA damage.* Carcinogenesis, 2000. **21**(3): p. 361-70.
24. Wang, Y.S., *Bulky DNA lesions induced by reactive oxygen species.* Chemical Research in Toxicology, 2008. **21**(2): p. 276-281.
25. Barnes, D.E. and T. Lindahl, *Repair and genetic consequences of endogenous DNA base damage in mammalian cells.* Annual Review of Genetics, 2004. **38**: p. 445-476.
26. Delaney, J.C. and J.M. Essigmann, *Biological properties of single chemical-DNA adducts: A twenty year perspective.* Chemical Research in Toxicology, 2008. **21**(1): p. 232-252.
27. You, Y.H. and G.P. Pfeifer, *Similarities in sunlight-induced mutational spectra of CpG-methylated transgenes and the p53 gene in skin cancer point to an important role of 5-methylcytosine residues in solar UV mutagenesis.* Journal of Molecular Biology, 2001. **305**(3): p. 389-399.
28. Loeb, L.A. and C.C. Harris, *Advances in chemical carcinogenesis: a historical review and prospective.* Cancer Res, 2008. **68**(17): p. 6863-72.
29. Burney, S., et al., *The chemistry of DNA damage from nitric oxide and peroxynitrite.* Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis, 1999. **424**(1-2): p. 37-49.
30. Shafirovich, V., et al., *Nitrogen dioxide as an oxidizing agent of 8-oxo-7,8-dihydro-2 '-deoxyguanosine but not of 2 '-deoxyguanosine.* Chemical Research in Toxicology, 2001. **14**(2): p. 233-241.
31. Niles, J.C., J.S. Wishnok, and S.R. Tannenbaum, *Peroxynitrite-induced oxidation and nitration products of guanine and 8-oxoguanine: structures and mechanisms of product formation.* Nitric Oxide, 2006. **14**(2): p. 109-21.
32. Cadet, J., T. Douki, and J.L. Ravanat, *One-electron oxidation of DNA and inflammation processes.* Nat Chem Biol, 2006. **2**(7): p. 348-9.
33. Steenken, S. and S.V. Jovanovic, *How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution.* Journal of the American Chemical Society, 1997. **119**(3): p. 617-618.
34. Cadet, J., et al., *Hydroxyl radicals and DNA base damage.* Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis, 1999. **424**(1-2): p. 9-21.
35. Lee, J.A., et al., *Peroxynitrite reacts with 8-nitropurines to yield 8-oxopurines.* Chemical Research in Toxicology, 2002. **15**(1): p. 7-14.
36. Povirk, L.F. and I.H. Goldberg, *Endonuclease-Resistant Apyrimidinic Sites Formed by Neocarzinostatin at Cytosine Residues in DNA - Evidence for a Possible Role in Mutagenesis.* Proceedings of the National Academy of Sciences of the United States of America, 1985. **82**(10): p. 3182-3186.
37. Weinfeld, M., et al., *Response of base excision repair enzymes to complex DNA lesions.* Radiation Research, 2001. **156**(5): p. 584-589.
38. Dedon, P.C., *The chemical toxicology of 2-deoxyribose oxidation in DNA.* Chem Res Toxicol, 2008. **21**(1): p. 206-19.
39. Yamamoto, K. and S. Kawanishi, *Hydroxyl Free-Radical Is Not the Main Active Species in Site-Specific DNA Damage Induced by Copper(Ii) Ion and Hydrogen-Peroxide.* Journal of Biological Chemistry, 1989. **264**(26): p. 15435-15440.

40. Goyne, T.E. and D.S. Sigman, *Nuclease Activity of 1,10-Phenanthroline Copper-Ion - Chemistry of Deoxyribose Oxidation.* Journal of the American Chemical Society, 1987. **109**(9): p. 2846-2848.
41. Johnson, G.R.A. and N.B. Nazhat, *Kinetics and Mechanism of the Reaction of the Bis(1,10-Phenanthroline)Copper(I) Ion with Hydrogen-Peroxide in Aqueous-Solution.* Journal of the American Chemical Society, 1987. **109**(7): p. 1990-1994.
42. Miaskiewicz, K. and R. Osman, *Theoretical-Study on the Deoxyribose Radicals Formed by Hydrogen Abstraction.* Journal of the American Chemical Society, 1994. **116**(1): p. 232-238.
43. Dizdaroglu, M., et al., *Radiation-Chemistry of Carbohydrates .10. Gamma-Radiolysis of Crystalline D-Glucose and D-Fructose.* Zeitschrift Fur Naturforschung Section B-a Journal of Chemical Sciences, 1977. **32**(2): p. 213-224.
44. Sitlani, A., et al., *DNA Photocleavage by Phenanthrenequinone Diimine Complexes of Rhodium(Iii) - Shape-Selective Recognition and Reaction.* Journal of the American Chemical Society, 1992. **114**(7): p. 2303-2312.
45. Wu, J.C., J. Stubbe, and J.W. Kozarich, *Mechanism of Bleomycin - Evidence for 4'-Ketone Formation in Poly(Da-Du) Associated Exclusively with Free Base Release.* Biochemistry, 1985. **24**(26): p. 7569-7573.
46. Rashid, R., et al., *Bleomycin versus OH-radical-induced malonaldehydic-product formation in DNA.* International Journal of Radiation Biology, 1999. **75**(1): p. 101-109.
47. Henner, W.D., et al., *Gamma-Ray Induced Deoxyribonucleic-Acid Strand Breaks - 3' Glycolate Termini.* Journal of Biological Chemistry, 1983. **258**(2): p. 711-713.
48. Kappen, L.S., et al., *Neocarzinostatin-Induced Hydrogen-Atom Abstraction from C-4' and C-5' of the T-Residue at a D(Gt) Step in Oligonucleotides - Shuttling between Deoxyribose Attack Sites Based on Isotope Selection Effects.* Biochemistry, 1991. **30**(8): p. 2034-2042.
49. Isildar, M., et al., *Gamma-Radiolysis of DNA in Oxygenated Aqueous-Solutions - Alterations at the Sugar Moiety.* International Journal of Radiation Biology, 1981. **40**(4): p. 347-354.
50. Kappen, L.S., I.H. Goldberg, and J.M. Liesch, *Identification of Thymidine-5'-Aldehyde at DNA Strand Breaks Induced by Neocarzinostatin Chromophore.* Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences, 1982. **79**(3): p. 744-748.
51. Pratviel, G., et al., *Furfural as a Marker of DNA Cleavage by Hydroxylation at the 5' Carbon of Deoxyribose.* Angewandte Chemie-International Edition in English, 1991. **30**(6): p. 702-704.
52. Chen, B.Z., et al., *5 '-(2-phosphoryl-1,4-dioxobutane) as a product of 5 '-oxidation of deoxyribose in DNA: Elimination as trans-1,4-dioxo-2-butene and approaches to analysis.* Chemical Research in Toxicology, 2004. **17**(11): p. 1406-1413.
53. Pogozelski, W.K. and T.D. Tullius, *Oxidative strand scission of nucleic acids: Routes initiated by hydrogen abstraction from the sugar moiety.* Chemical Reviews, 1998. **98**(3): p. 1089-1107.
54. Zhou, X.F., K. Taghizadeh, and P.C. Dedon, *Chemical and biological evidence for base propenals as the major source of the endogenous M(1)dG adduct in cellular DNA.* Journal of Biological Chemistry, 2005. **280**(27): p. 25377-25382.
55. Dedon, P.C., et al., *Indirect mutagenesis by oxidative DNA damage: formation of the pyrimidopurinone adduct of deoxyguanosine by base propenal.* Proc Natl Acad Sci U S A, 1998. **95**(19): p. 11113-6.
56. Plastaras, J.P., et al., *Reactivity and mutagenicity of endogenous DNA oxopropenylating agents: base propenals, malondialdehyde, and N(epsilon)-oxopropenyllysine.* Chem Res Toxicol, 2000. **13**(12): p. 1235-42.
57. Bohnert, T., L. Gingipalli, and P.C. Dedon, *Reaction of 2 '-deoxyribonucleosides with cis- and trans-1,4-dioxo-2-butene.* Biochemical and Biophysical Research Communications, 2004. **323**(3): p. 838-844.

58. Chen, B.Z., et al., *Formation of 1,4-dioxo-2-butene-derived adducts of 2 '-deoxyadenosine and 2 '-deoxycytidine in oxidized DNA.* Chemical Research in Toxicology, 2006. **19**(8): p. 982-985.
59. Byrns, M.C., D.P. Predecki, and L.A. Peterson, *Characterization of nucleoside adducts of cis-2-butene-1,4-dial, a reactive metabolite of furan.* Chemical Research in Toxicology, 2002. **15**(3): p. 373-379.
60. Byrns, M.C., et al., *Detection of DNA adducts derived from the reactive metabolite of furan, cis-2-butene-1,4-dial.* Chemical Research in Toxicology, 2006. **19**(3): p. 414-420.
61. Lucas, L.T., D. Gatehouse, and D.E. Shuker, *Efficient nitroso group transfer from N-nitrosoindoles to nucleotides and 2'-deoxyguanosine at physiological pH. A new pathway for N-nitrosocompounds to exert genotoxicity.* J Biol Chem, 1999. **274**(26): p. 18319-26.
62. Spickett, C.M. and G. Dever, *Studies of phospholipid oxidation by electrospray mass spectrometry: From analysis in cells to biological effects.* Biofactors, 2005. **24**(1-4): p. 17-31.
63. Jiang, Q., B.C. Blount, and B.N. Ames, *5-chlorouracil, a marker of DNA damage from hypochlorous acid during inflammation - A gas chromatography-mass spectrometry assay.* Journal of Biological Chemistry, 2003. **278**(35): p. 32834-32840.
64. Kulcharyk, P.A. and J.W. Heinecke, *Hypochlorous acid produced by the myeloperoxidase system of human phagocytes induces covalent cross-links between DNA and protein.* Biochemistry, 2001. **40**(12): p. 3648-3656.
65. Bartsch, H. and J. Nair, *Oxidative stress and lipid peroxidation-derived DNA-lesions in inflammation driven carcinogenesis.* Cancer Detection and Prevention, 2004. **28**(6): p. 385-391.
66. Marnett, L.J., *Oxy radicals, lipid peroxidation and DNA damage.* Toxicology, 2002. **181**: p. 219-222.
67. Montine, T.J., et al., *Lipid peroxidation in aging brain and Alzheimer's disease.* Free Radical Biology and Medicine, 2002. **33**(5): p. 620-626.
68. Porter, N.A., S.E. Caldwell, and K.A. Mills, *Mechanisms of Free-Radical Oxidation of Unsaturated Lipids.* Lipids, 1995. **30**(4): p. 277-290.
69. Kaneko, T., et al., *Lethal Effects of a Linoleic-Acid Hydroperoxide and Its Autoxidation Products, Unsaturated Aliphatic-Aldehydes, on Human-Diploid Fibroblasts.* Chemico-Biological Interactions, 1987. **63**(2): p. 127-137.
70. Esterbauer, H., H. Zollner, and R.J. Schaur, *Hydroxyalkenals - Cyto-Toxic Products of Lipid-Peroxidation.* Isi Atlas of Science-Biochemistry, 1988. **1**(4): p. 311-317.
71. Marnett, L.J., *Lipid peroxidation-DNA damage by malondialdehyde.* Mutat Res, 1999. **424**(1-2): p. 83-95.
72. Stone, K., M.B. Ksebati, and L.J. Marnett, *Investigation of the adducts formed by reaction of malondialdehyde with adenosine.* Chem Res Toxicol, 1990. **3**(1): p. 33-8.
73. Stone, K., A. Uzieblo, and L.J. Marnett, *Studies of the reaction of malondialdehyde with cytosine nucleosides.* Chem Res Toxicol, 1990. **3**(5): p. 467-72.
74. Esterbauer, H., R.J. Schaur, and H. Zollner, *Chemistry and biochemistry of 4-hydroxynonenal, malonaldehyde and related aldehydes.* Free Radic Biol Med, 1991. **11**(1): p. 81-128.
75. Rindgen, D., et al., *Covalent modifications to 2'-deoxyguanosine by 4-oxo-2-nonenal, a novel product of lipid peroxidation.* Chem Res Toxicol, 1999. **12**(12): p. 1195-204.
76. Olinski, R., et al., *Urinary measurement of 8-OxodG, 8-OxoGua, and 5HMUra: a noninvasive assessment of oxidative damage to DNA.* Antioxid Redox Signal, 2006. **8**(5-6): p. 1011-9.
77. Svoboda, P., et al., *Urinary 8-hydroxyguanine may be a better marker of oxidative stress than 8-hydroxydeoxyguanosine in relation to the life spans of various species.* Antioxid Redox Signal, 2006. **8**(5-6): p. 985-92.
78. Weimann, A., D. Belling, and H.E. Poulsen, *Quantification of 8-oxo-guanine and guanine as the nucleobase, nucleoside and deoxynucleoside forms in human urine by high-performance liquid*

chromatography-electrospray tandem mass spectrometry. Nucleic Acids Res, 2002. **30**(2): p. E7.
79. Chen, H.J. and C.M. Chang, *Quantification of urinary excretion of 1,N6-ethenoadenine, a potential biomarker of lipid peroxidation, in humans by stable isotope dilution liquid chromatography-electrospray ionization-tandem mass spectrometry: comparison with gas chromatography-mass spectrometry.* Chem Res Toxicol, 2004. **17**(7): p. 963-71.
80. Hillestrom, P.R., et al., *Quantification of 1,N6-etheno-2'-deoxyadenosine in human urine by column-switching LC/APCI-MS/MS.* Free Radic Biol Med, 2004. **36**(11): p. 1383-92.
81. Chen, H.J., et al., *Analysis of 3,N(4)-ethenocytosine in DNA and in human urine by isotope dilution gas chromatography/negative ion chemical ionization/mass spectrometry.* Chem Res Toxicol, 2001. **14**(12): p. 1612-9.
82. Dechakhamphu, S., et al., *High excretion of etheno adducts in liver fluke-infected patients: protection by praziquantel against DNA damage.* Cancer Epidemiol Biomarkers Prev, 2008. **17**(7): p. 1658-64.
83. Knutson, C.G., et al., *Oxidation and glycolytic cleavage of etheno and propano DNA base adducts.* Biochemistry, 2009. **48**(4): p. 800-9.
84. VanderVeen, L.A., et al., *Induction of frameshift and base pair substitution mutations by the major DNA adduct of the endogenous carcinogen malondialdehyde.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14247-52.
85. Hoberg, A.M., et al., *Measurement of the malondialdehyde-2 '-deoxyguanosine adduct in human urine by immunoextraction and liquid chromatography/atmospheric pressure chemical ionization tandem mass spectrometry.* Journal of Mass Spectrometry, 2004. **39**(1): p. 38-42.
86. Knutson, C.G., et al., *Metabolism and elimination of the endogenous DNA adduct, 3-(2-deoxy-beta-D-erythropentofuranosyl)-pyrimido[1,2-alpha]purine-10(3H)-one, in the rat.* J Biol Chem, 2007. **282**(50): p. 36257-64.
87. Knutson, C.G., et al., *Monitoring in vivo metabolism and elimination of the endogenous DNA adduct, M1dG {3-(2-deoxy-beta-D-erythro-pentofuranosyl)pyrimido[1,2-alpha]purin-10(3H)-one}, by accelerator mass spectrometry.* Chem Res Toxicol, 2008. **21**(6): p. 1290-4.
88. Otteneder, M.B., et al., *In vivo oxidative metabolism of a major peroxidation-derived DNA adduct, M1dG.* Proc Natl Acad Sci U S A, 2006. **103**(17): p. 6665-9.
89. Guengerich, F.P., *Cytochrome P450 oxidations in the generation of reactive electrophiles: epoxidation and related reactions.* Arch Biochem Biophys, 2003. **409**(1): p. 59-71.
90. Ravindranath, V., L.T. Burka, and M.R. Boyd, *Reactive metabolites from the bioactivation of toxic methylfurans.* Science, 1984. **224**(4651): p. 884-6.
91. Chen, L.J., S.S. Hecht, and L.A. Peterson, *Characterization of amino acid and glutathione adducts of cis-2-butene-1,4-dial, a reactive metabolite of furan.* Chem Res Toxicol, 1997. **10**(8): p. 866-74.
92. Burka, L.T., K.D. Washburn, and R.D. Irwin, *Disposition of [14C]furan in the male F344 rat.* J Toxicol Environ Health, 1991. **34**(2): p. 245-57.
93. Peterson, L.A., et al., *Identification of a cis-2-butene-1,4-dial-derived glutathione conjugate in the urine of furan-treated rats.* Chem Res Toxicol, 2006. **19**(9): p. 1138-41.
94. Kellert, M., et al., *Biomarkers of furan exposure by metabolic profiling of rat urine with liquid chromatography-tandem mass spectrometry and principal component analysis.* Chem Res Toxicol, 2008. **21**(3): p. 761-8.
95. Lu, D., et al., *Degraded protein adducts of cis-2-butene-1,4-dial are urinary and hepatocyte metabolites of furan.* Chem Res Toxicol, 2009. **22**(6): p. 997-1007.
96. Lu, D. and L.A. Peterson, *Identification of furan metabolites derived from cysteine-cis-2-butene-1,4-dial-lysine cross-links.* Chem Res Toxicol, 2010. **23**(1): p. 142-51.

97. Peterson, L.A., et al., *Polyamines are traps for reactive intermediates in furan metabolism.* Chem Res Toxicol, 2011. **24**(11): p. 1924-36.
98. Hamberger, C., et al., *Hepatobiliary toxicity of furan: identification of furan metabolites in bile of male f344/n rats.* Drug Metab Dispos, 2010. **38**(10): p. 1698-706.
99. Crawford, D.L., T.C. Yu, and Sinnhube.Ro, *Reaction of Malonaldehyde with Glycine.* Journal of Agricultural and Food Chemistry, 1966. **14**(2): p. 182-&.
100. Buttkus, H., *Reaction of Cysteine and Methionine with Malonaldehyde.* Journal of the American Oil Chemists Society, 1969. **46**(2): p. 88-&.
101. McBrien, D.C.H. and T.F. Slater, *Free radicals, lipid peroxidation and cancer*. NFCR cancer symposia. 1982, London ; New York: Academic Press. xii, 447 p., 3 p. of plates.
102. Siu, G.M. and H.H. Draper, *Metabolism of malonaldehyde in vivo and in vitro.* Lipids, 1982. **17**(5): p. 349-55.
103. Marnett, L.J., et al., *Distribution and oxidation of malondialdehyde in mice.* Prostaglandins, 1985. **30**(2): p. 241-54.
104. Moellering, R.E. and B.F. Cravatt, *Functional Lysine Modification by an Intrinsically Reactive Primary Glycolytic Metabolite.* Science, 2013. **341**(6145): p. 549-553.
105. Park, J., et al., *SIRT5-mediated lysine desuccinylation impacts diverse metabolic pathways.* Mol Cell, 2013. **50**(6): p. 919-30.
106. Weinert, B.T., et al., *Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation.* Cell Rep, 2013. **4**(4): p. 842-51.
107. Wagner, G.R. and R.M. Payne, *Widespread and enzyme-independent Nepsilon-acetylation and Nepsilon-succinylation of proteins in the chemical conditions of the mitochondrial matrix.* J Biol Chem, 2013. **288**(40): p. 29036-45.
108. Bardella, C., et al., *Aberrant succination of proteins in fumarate hydratase-deficient mice and HLRCC patients is a robust biomarker of mutation status.* J Pathol, 2011. **225**(1): p. 4-11.
109. Yang, Y., et al., *A novel fumarate hydratase-deficient HLRCC kidney cancer cell line, UOK268: a model of the Warburg effect in cancer.* Cancer Genet, 2012. **205**(7-8): p. 377-90.
110. Berg, J.M., J.L. Tymoczko, and L. Stryer, *Biochemistry*. 7th ed. 2012, New York: W.H. Freeman. xxxii, 1054, 43, 41, 48 p.
111. McMurry, J. and T.P. Begley, *The organic chemistry of biological pathways*. 2005, Englewood, Colo.: Roberts and Co. Publishers. xxi, 490 p.
112. Rosen, R., et al., *Probing the active site of homoserine trans-succinylase.* FEBS Lett, 2004. **577**(3): p. 386-92.
113. Zhang, Z., et al., *Identification of lysine succinylation as a new post-translational modification.* Nat Chem Biol, 2011. **7**(1): p. 58-63.
114. Colak, G., et al., *Identification of lysine succinylation substrates and the succinylation regulatory enzyme CobB in Escherichia coli.* Mol Cell Proteomics, 2013. **12**(12): p. 3509-20.
115. Xie, Z., et al., *Lysine succinylation and lysine malonylation in histones.* Mol Cell Proteomics, 2012. **11**(5): p. 100-7.
116. Rardin, M.J., et al., *SIRT5 regulates the mitochondrial lysine succinylome and metabolic networks.* Cell Metab, 2013. **18**(6): p. 920-33.
117. Alderson, N.L., et al., *S-(2-Succinyl)cysteine: a novel chemical modification of tissue proteins by a Krebs cycle intermediate.* Arch Biochem Biophys, 2006. **450**(1): p. 1-8.
118. Pollard, P.J., et al., *Accumulation of Krebs cycle intermediates and over-expression of HIF1alpha in tumours which result from germline FH and SDH mutations.* Hum Mol Genet, 2005. **14**(15): p. 2231-9.
119. Isaacs, J.S., et al., *HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: novel role of fumarate in regulation of HIF stability.* Cancer Cell, 2005. **8**(2): p. 143-53.

120. Frizzell, N., M. Lima, and J.W. Baynes, *Succination of proteins in diabetes.* Free Radic Res, 2011. **45**(1): p. 101-9.
121. Tomlinson, I.P., et al., *Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer.* Nat Genet, 2002. **30**(4): p. 406-10.
122. Adam, J., et al., *Renal cyst formation in Fh1-deficient mice is independent of the Hif/Phd pathway: roles for fumarate in KEAP1 succination and Nrf2 signaling.* Cancer Cell, 2011. **20**(4): p. 524-37.
123. Burgess, D.J., *Metabolism: Glutamine connections.* Nat Rev Cancer, 2013. **13**(5): p. 293.
124. Moeller, B.C., et al., *Biomarkers of exposure and effect in human lymphoblastoid TK6 cells following [13C2]-acetaldehyde exposure.* Toxicol Sci, 2013. **133**(1): p. 1-12.
125. Lu, K., et al., *Molecular dosimetry of N2-hydroxymethyl-dG DNA adducts in rats exposed to formaldehyde.* Chem Res Toxicol, 2011. **24**(2): p. 159-61.
126. Vasconcellos, S.E., et al., *Distinct genotypic profiles of the two major clades of Mycobacterium africanum.* BMC Infect Dis, 2010. **10**: p. 80.
127. Brennan, P.J., *Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis.* Tuberculosis (Edinb), 2003. **83**(1-3): p. 91-7.
128. McKinney, J.D., et al., *Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase.* Nature, 2000. **406**(6797): p. 735-8.
129. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.* Nature, 1998. **393**(6685): p. 537-44.
130. Kaufmann, S.H., *How can immunology contribute to the control of tuberculosis?* Nat Rev Immunol, 2001. **1**(1): p. 20-30.
131. Russell, D.G., *Mycobacterium tuberculosis: here today, and here tomorrow.* Nat Rev Mol Cell Biol, 2001. **2**(8): p. 569-77.
132. Pieters, J., *Mycobacterium tuberculosis and the macrophage: maintaining a balance.* Cell Host Microbe, 2008. **3**(6): p. 399-407.
133. Flynn, J.L. and J. Chan, *Tuberculosis: Latency and reactivation.* Infection and Immunity, 2001. **69**(7): p. 4195-4201.
134. Zhang, Y., *Persistent and dormant tubercle bacilli and latent tuberculosis.* Front Biosci, 2004. **9**: p. 1136-56.
135. Ramakrishnan, L., *Revisiting the role of the granuloma in tuberculosis.* Nat Rev Immunol, 2012. **12**(5): p. 352-66.
136. Schnappinger, D., et al., *Transcriptional adaptation of Mycobacterium tuberculosis within macrophages: Insights into the phagosomal environment.* Journal of Experimental Medicine, 2003. **198**(5): p. 693-704.
137. Tailleux, L., et al., *Probing Host Pathogen Cross-Talk by Transcriptional Profiling of Both Mycobacterium tuberculosis and Infected Human Dendritic Cells and Macrophages.* Plos One, 2008. **3**(1).
138. Cappelli, G., et al., *Profiling of Mycobacterium tuberculosis gene expression during human macrophage infection: Upregulation of the alternative sigma factor G, a group of transcriptional regulators, and proteins with unknown function.* Research in Microbiology, 2006. **157**(5): p. 445-455.
139. Talaat, A.M., et al., *The temporal expression profile of Mycobacterium tuberculosis infection in mice.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(13): p. 4602-4607.
140. Talaat, A.M., et al., *Mycobacterial bacilli are metabolically active during chronic tuberculosis in murine lungs: Insights from genome-wide transcriptional profiling.* Journal of Bacteriology, 2007. **189**(11): p. 4265-4274.

141. Schnappinger, D., et al., *Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment.* J Exp Med, 2003. **198**(5): p. 693-704.
142. Eisenreich, W., et al., *Carbon metabolism of intracellular bacterial pathogens and possible links to virulence.* Nature Reviews Microbiology, 2010. **8**(6): p. 401-412.
143. Upton, A.M. and J.D. McKinney, *Role of the methylcitrate cycle in propionate metabolism and detoxification in Mycobacterium smegmatis.* Microbiology, 2007. **153**(Pt 12): p. 3973-82.
144. Kornberg, H.L. and H. Beevers, *The Glyoxylate Cycle as a Stage in the Conversion of Fat to Carbohydrate in Castor Beans.* Biochimica Et Biophysica Acta, 1957. **26**(3): p. 531-537.
145. Kornberg, H.L. and H. Beevers, *A mechanism of conversion of fat to carbohydrate in castor beans.* Nature, 1957. **180**(4575): p. 35-6.
146. Eastmond, P.J., et al., *Postgerminative growth and lipid catabolism in oilseeds lacking the glyoxylate cycle.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5669-74.
147. Murthy, P.S., M. Sirsi, and T. Ramakrishnan, *Effect of age on the enzymes of tricarboxylic acid and related cycles in Mycobacterium tuberculosis H37Rv.* Am Rev Respir Dis, 1973. **108**(3): p. 689-90.
148. Wayne, L.G. and K.Y. Lin, *Glyoxylate metabolism and adaptation of Mycobacterium tuberculosis to survival under anaerobic conditions.* Infect Immun, 1982. **37**(3): p. 1042-9.
149. Quartararo, C.E., et al., *Solvent isotope-induced equilibrium perturbation for isocitrate lyase.* Biochemistry, 2013. **52**(51): p. 9286-93.
150. Ford, C.B., et al., *Use of whole genome sequencing to estimate the mutation rate of Mycobacterium tuberculosis during latent infection.* Nat Genet, 2011. **43**(5): p. 482-6.

Chapter 2

In vitro metabolite profiling to develop candidate biomarkers of thymine propenal metabolism *in vivo*

Abstract

Toward the goal of developing DNA damage products as biomarkers, this chapter addresses the *in silico* and *in vitro* analysis of the biotransformational reactions of base propenals to identify metabolites as candidate biomarkers for analysis in rats in Chapter 3. Base propenals are the α,β-unsaturated aldehydic products arising from 4'-oxidation of the sugar in DNA and they are among the most common products of DNA oxidation. The well known reactivity of α,β-unsaturated carbonyls provides an opportunity for predictive modeling of base propenal metabolism and *in vitro* analysis of the models, starting from basic reactions with cellular nucleophiles such as glutathione (GSH). Here we use thymine propenal as a model to conduct *in vitro* reactions with GSH and other cellular nucleophiles, as well as reactions in liver cell extracts, to obtain a metabolite spectrum predictive of *in vivo* biotransformation. A critical feature of this work involves the development of analytical tools to characterize and quantify the metabolites in vitro and to apply to the *in vivo* analysis in Chapter 3. We report that the direct reaction of thymine propenal with GSH ($t_{1/2}$ ~105 s) generates a Michael adduct (M-GSH), an addition-elimination product (AE-GSH), and a bis-GSH adduct (bis-GSH). M-GSH and bis-GSH are unstable and undergo reversal to thymine propenal and GSH, or transformation to AE-GSH, respectively. In studies with rat liver extracts, reactions with the microsomal fraction yielded a product spectrum similar to direct reaction, while reactions in the cytosolic fraction were very rapid (Tp consumption within 15 s) and resulted in aldehydic reduction of the 3 GSH adducts (AE-GSH, bis-GSH, M-GSH) and oxidation and reduction of parent Tp to, thymine propenol and thymine propenoic acid. These stable metabolites represent potential candidates for in vivo metabolism, so chemical standards were synthesized and sensitive LC-MS/MS methods developed to further characterize the metabolites *in vivo* in Chapter 3.

Introduction

The goal of this Chapter of the thesis is to create a set of metabolic products of thymine propenal (Tp) to serve as candidate biomarkers for Tp metabolism in *in vivo* in Chapter 3. As products of 4'-oxidation of DNA, base propenals possess an α,β-unsaturated carbonyl structure with well characterized reactivity [1], especially for nucleophilic reactions with GSH and other cellular nucleophiles. For example, the lipid peroxidation product 4-hydroxynonenal spontaneously reacts with GSH to form Michael adducts that are subsequently metabolized to mercapturic acid derivatives before excretion into urine [2, 3]. Another example would be *cis*-2-buten-1,4-dial, which reacts with GSH to form cyclic GSH adducts as mentioned earlier in Chapter 1. Base propenals have also been shown to undergo both enzyme-independent and -dependent reactions with GSH [4]. Tp thus provides an opportunity for predictive modeling of DNA damage product metabolism and *in vitro* analysis of the models to create a metabolite spectrum predictive of *in vivo* biotransformation.

The 2'-deoxyribose moiety of DNA is as vulnerable to oxidative damage as the nucleobases, with a similar 1000-fold protective effect of the cellular environment against oxidation [5-7]. Oxidation of each carbon in 2'-deoxyribose brings about not only DNA strand breaks but also generates highly electrophilic products possessing α, β-unsaturated carbonyl groups that are susceptible to attack by nucleophiles in DNA nucleobases, proteins, and low-molecular-weight thiols. As the most abundant cellular thiol, GSH is ubiquitous throughout cells, in mitochondria, cytosol, and nuclei [8-11], with very high concentrations ranging from 1 to 10 mM [12]. This suggests that endogenous electrophiles have a high likelihood of reacting directly with GSH to form conjugates. Moreover, glutathione S-transferases (GSTs) play a central role in protecting cells against oxidative stress in partnership with glutathione peroxidase (GSHPxs) [13], with GSTs catalyzing GSH conjugation with cellular electrophiles [14, 15]. 2'-Deoxyribose oxidation products are likely to be susceptible to GST-mediated conjugation with GSH, because α, β-unsaturated

carbonyls satisfy the two criteria for GST substrates: they contain an electrophilic site and they readily react with GSH [16]. For these reasons, base propenals were chosen in this study, with their α, β-unsaturated carbonyl structure illustrating a common feature of other 2'-deoxyribose oxidation products and other cellular oxidation products (Scheme 2.1). In this Chapter, the direct reactions and enzymatic reactions of thymine propenal with GSH and other cellular nucleophiles are described in terms of the formation and structural characterization of the metabolites, along with other in vitro metabolic reactions arising in liver cell extracts. A major feature of this work involves the chemical synthesis of metabolite candidates as tools to develop analytical methods and as standards to quantify metabolites discovered *in vivo* in Chapter 3.

Materials & methods

Materials. The following chemicals were commercially obtained and used without further purification: thymine; 3, 3-diethoxy-1-propyne; glutathione (GSH); potassium monohydrogen phosphate (K_2HPO_4); potassium dihydrogen phosphate (KH_2PO_4); N, O-bis-(trifluoromethylsilyl)trifluoroacetamide (BSTFA); liver microsome from female Sprague-Dawley rats; liver cytosol from human; β-nicotinamide adenine dinucleotide phosphate, reduced, sodium salt (NADPH); glutathione S-transferases from equine liver, were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO). Deionized water was further purified with a Milli-Q system (Millipore Corporation, Bedford, MA) and was used in all experiments.

Synthesis of AE-GSH 2a, bis-GSH 3a, AE-NAc 2e and bis-NAc 3e. Deprotection of 3, 3-diethoxy-1-propynal to propargyl aldehyde is modified from [17]. Briefly, 3, 3-diethoxy-1-propynal 128 or 1.64 g N-acetylcysteine is added to 1 mL of formic acid in a 10 mL round bottom flask. The reaction is stirred at 4 °C for 30 min. Then, 1.84 g GSH (3 mmol) in 2 mL of water is added to the reaction and allowed to stir for 10 min at ambient temperature. Then, 120 mg NaOH in 2 mL of water is added dropwise to the solution. The reaction is then allowed to stir for 4 hr to obtain **2a**, **3a, 2e and 3e**. Purification to obtain **2a, 3a, 2e ad 3e** was performed by HPLC with water/acetonitrile gradient in Table 2.2.

Synthesis of M-GSH^{red} 1b and M-NAc^{red} 1e. GSH or N-acetylcysteine (3 mmol) is added to the solution of 90 mg (0.5 mmol) of (a) in 4 mL of 1:1 acetonitrile/water in a 10 mL round bottom flask at 4 °C. The solution is rigorously stirred for 15 h to obtain 1a or 1e (not isolable). At 4 °C, 380 mg of $NaBH_4$ (10 mmol) is gradually added, and the reaction mixture is allowed to stir for 1 h to obtain 1b or 1e. The solution is then neutralized using 20% v/v formic acid until no gas develops. 1b and 1e is purified from the reaction using HPLC with water/acetonitrile gradient. Structural characterization of the compounds was achieved as described in Table 2.2.

Biochemical preparation of M-CG^{red} 1c. ~1 μmol of 1b was treated with γ-glutamyltransferase in 0.5 mL tris-HCl buffer pH 10 for 2 h. The protein was filter through 10 kD exclusion filter and the filtrate was purified using HPLC with water/acetonitrile gradient. The structure of 1c was confirmed by HR-MS and MS/MS as shown in Table 2.2.

Biochemical preparation of M-Cys^{red} 1d. ~1 μmol of 1e was treated with acylase I in 0.5 mL phosphate buffer pH 7.0 for 4 h. The protein was filter through 10 kD exclusion filter and the filtrate was purified using HPLC with water/acetonitrile gradient. The structure of 1d was confirmed by HR-MS and MS/MS as shown in Table 2.2.

Synthesis of AE-GSH^{red} 2b**, bis-GSH^{red}** 3b**, AE-NAc^{red}** 2e **and bis-NAc^{red}** 3e. 3, 3-diethoxy-1-propynal 128 mg (1 mmol) is added to 10 mL of formic acid in 50 mL round bottom flask. The reaction is stirred at 4°C for 30 min. Then, 9.21 g GSH (30 mmol) or 1.64 g N-acetyl cysteine in 2 mL of water is added to the reaction. Then, 480 mg NaOH in 10 mL of water is added dropwise to the solution. The reaction is then allowed to stir for 3 hr to obtain the aldehyde intermediate. Then, $NaBH_4$ is added to the reaction mixture until no gas was developed at 4°C, and stir for 1 hr at room temperature. The reaction is then quenched by 20% v/v formic acid until no gas develops. The crude is then purified using HPLC with water/acetonitrile gradient.

Biochemical preparation of AE-CG^{red} 2c and bis-CG^{red} 3c. ~1 μmol of 2b and 3b was treated with γ-glutamyltransferase in 0.5 mL tris-HCl buffer pH 10 for 2 h. The protein was filter through 10 kD exclusion filter and the filtrate was purified using HPLC with water/acetonitrile gradient. The structure of 2c and 3c was confirmed by HR-MS and MS/MS as shown in Table 2.2.

Biochemical preparation of AE-Cys^{red} 2d and bis-Cys^{red} 3d. ~1 μmol of 2e and 3e was treated with acylase I in 0.5 mL phosphate buffer pH 7.0 for 4 h. The protein was filter through 10 kD exclusion filter and the filtrate was purified using HPLC with water/acetonitrile gradient. The structure of 2d and 3d was confirmed by HR-MS and MS/MS as shown in Table 2.2.

3-Thymine propenol (Tp^{red} 4). 18 mg of Tp (0.01 mmol) in 5 mL of 1:1 acetonitrile water in 25 mL round bottom flask was reduced with 38 mg $NaBH_4$ (1 mmol). The reaction was allowed to stir for 1 h and quenched with 20% v/v formic acid until no gas developed. The crude product was purified by HPLC with water/acetonitrile gradient to obtain 4. Structural characterization of the compound was achieved as described in Table 2.2.

3-Thymine propenoic acid (Tp^{ox} 5). The first step of synthesis of 5 was modified from published work [18]. Briefly, thymine 128 mg (1 mmol) was added with 3 mL BSTFA and the suspension was heated at 120 °C for 1 h or until the mixture became homogenous, which was the sign of complete derivatization. The solution was then allowed to cool to ambient temperature. Then, 336 mg (4 mmol) methyl propiolate is added and the reaction mixture was heated again at 50 ºC for 16 h. Then, the reaction was cooled to ambient temperature and water was added to obtain the white precipitate of 3-thyminyl methyl acrylate. The suspension was filtered and the white precipitate collected, which was then washed three times with cyclohexane. Then, 2 mL of 1 M aqueous NaOH was added to dissolve the precipitate, followed by slow addition of 0.5 mL of 1 M formic acid to adjust pH to 4, with further purification by HPLC with water/acetonitrile gradient to obtain **5**. Structural characterization of the compound was achieved as described in Table 2.2.

3-Thymine propenal (Tp 6). The synthesis of thymine propenal in this study was modified from published methods [17, 18]. To prepare propargyl aldehyde, 3, 3-diethoxy-1-propyanl was deprotected by formic acid. The reaction is stirred for 1 h and then phosphorus pentoxide (P_2O_5) was added to eliminate formic acid to carbon monoxide and water. Then the crude propargyl aldehyde was added to silylated thymine generated from the reaction of thymine with BSTFA. The reaction was stirred for 4 h. Water was then added to the reaction and thymine propenal was precipitated as a white solid. The suspension was filtered to obtain the white precipitate, which was dissolved in DMSO. The crude of thymine propenal in DMSO was further purified by flash column chromatography using 230-400 mesh ASTM, particle 0.040-0.063 mm, Silica gel 60. The mobile

phase used to elute the system was: 10% to 35% EtOAc/hexane. Structural characterization of the compound was achieved as described in Table 2.2.

Instrumentation. UV-vis absorbance measurements were made on an HP8452 diode-array spectrometer (Agilent Technologies). HPLC-UV analyses were carried out on Agilent 1290 series HPLC system with binary pumps and degassers. Full scan and collision-induced dissociation (CID) mass spectra of *in vitro* reaction products were obtained with an Agilent Technologies 6430 ion trap LC/MS coupled to Agilent 1200 series HPLC. High resolution mass spectral data for the synthetic standards were obtained from Agilent Technologies 6510 QTOF coupled to HPLC Agilent 1200 series HPLC. Quantitative data were obtained from multiple reaction monitoring mode (MRM) on an Agilent Technologies 6430 QQQ coupled to Agilent 1290 series HPLC.

HPLC methods. Four HPLC methods were applied in the experiment:

HPLC method 1. An HPLC column was equilibrated with 100% solvent A (0.1% formic acid in H_2O) and 0%B (acetonitrile) at a flow rate of 0.200 mL/min at 30°C. The solvent was programmed as follows: a linear gradient from the starting solvent to 5% B in 30 min; a linear gradient increasing from 5%B to 46%B for 12 min; increasing to 100% B in 0.1min, holding for 10 min; decreasing to 0% B in 0.1 min; and reequilibrating at initial conditions for 15 min.

HPLC method 2. ZORBAX Eclipse RRHD, XDB-C18 (2.1 × 100 mm), 1.8 μm (Agilent Technologies) was equilibrated with 95.5% solvent A (0.1% formic acid in H_2O) and 0.5%B (acetonitrile) at a flow rate of 0.250 mL/min at 20°C. The solvent was programmed as follows: a linear gradient from the starting solvent to 13% B in 13 min; increasing to 100% B in 0.1min, holding for 10 min; decreasing to 0.5% B in 0.1 min; and reequilibrating at initial conditions for 5 min.

HPLC method 3. ZORBAX Eclipse, XDB-C8 (4.6 × 250 mm), 5 μm (Agilent Technologies) was equilibrated with 95.5% solvent A (0.1% formic acid in H_2O) and 0.5%B (acetonitrile) at a flow rate

of 0.300 mL/min at 20°C. The solvent was programmed as follows: a linear gradient from the starting solvent to 7% B in 12 min; increasing to 100% B in 0.1min, holding for 10 min; decreasing to 0.5% B in 0.1 min; and reequilibrating at initial conditions for 15 min.

HPLC method 4. Dionex Acclaimed Polar Advantage, C18 (2.1 × 150 mm), 3 μm (Thermo Scientifics) was equilibrated with 100 % solvent A (0.1% formic acid in H_2O) and 0%B (acetonitrile) at a flow rate of 0.250 mL/min at 20°C. The solvent was programmed as follows: an isocratic elution of 0%B for 6 min; a linear gradient to 3% B in 13 min; a linear gradient from 3%B to 20%B in 10 min; increasing to 100% B in 0.1min, holding for 10 min; decreasing to 0% B in 0.1 min; and reequilibrating at initial conditions for 15 min.

Direct reaction of thymine propenal with GSH. 0.1 mM thymine propenal was reacted with 1 mM GSH at 37 °C in potassium phosphate buffer pH 7.4 for 1 h. The reaction mixture was then analyzed by LC-MS with HPLC method 1.

GST-catalyzed reaction of thymine propenal with GSH. 0.1 mM thymine propenal was reacted with 1 mM GSH with or without 10 U of equine GSTs at 37 °C in Clelex-treated phosphate buffer, pH 7.4 at these time points: 10, 15, 30, 45 and 60 s. For each time point, the reaction mixture was then quenched by 0.05% trifluoroacetic acid (final concentration) at 4 °C. GSTs was removed using 10 kD filter. The filtrant then analyzed by HPLC-UV Agilent 1100 series using HPLC method 4.

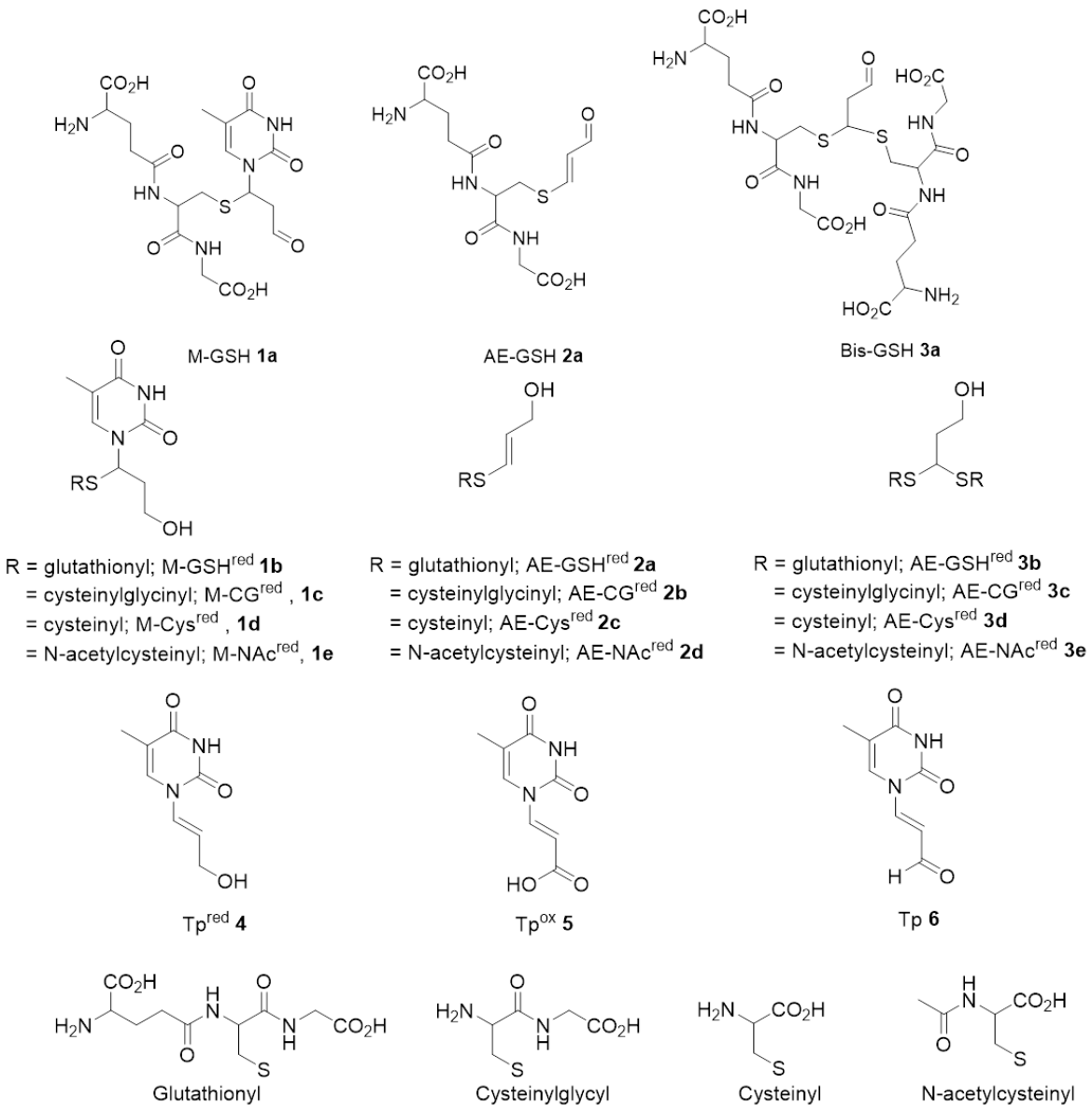
Reaction with rat liver microsomes. 0.1 mM thymine propenal was added to the solution of 1 mM GSH, 1 mM NADPH, 20 mg/mL microsomal fraction at 37 °C and 50 mM clelex-treated phosphate buffer, pH 7.4 for 1 h. Then, the reaction mixture was filtered through 10 kD filter to remove enzymes prior to LC-MS analysis using HPLC method 1. The control experiment was conducted similarly but NADPH was not added.

Reaction in rat liver cytosol. 0.1 mM thymine propenal was added to the solution of 1 mM GSH, 1 mM NADPH, 20 mg/mL cytosolic fraction at 37 °C and 50 mM clelex-treated phosphate buffer, pH 7.4 for 1 h. The reaction mixture was then analyzed by LC-MS with HPLC method B. The

control experiment was conducted similarly but NADPH was not added. Then, the reaction mixture was filtered through 10 kDa filter prior to injection.

LC-MS analysis of *in vitro* products. LC-MS analyses of the direct reactions, GST-catalyzed reactions, and the reactions in liver extracts were performed with a Dionex Acclaimed Polar Advantage, C18 (2.1 × 150 mm, Thermo Scientific) column at a flow rate of 0.2 mL/min. The column was eluted with HPLC method 1. An Agilent 1200 series HPLC was coupled to LC/MSD ion trap mass spectrometer operating in positive ion mode. Nitrogen was the nebulizing and drying gas (20.0 psi, 10.0 L/min) and had a temperature set at 325 °C. The analysis was performed with the mass spectrometer set to scanning mode with a scan rage (m/z) from 160 to 1000. The metabolites were then re-analyzed with the mass spectrometer set in manual MS^2 mode to obtain fragmentation patterns for each of the compounds.

Stability of the products. M-GSH 1a, AE-GSH 2a, and bis-GSH 3a were purified from the reaction in different fraction tube already sit on ice bath from LC-MS (ion trap) using HPLC method 1. Then, each compound was incubated in potassium phosphate buffer pH 7.4, 37 °C at these time points: 0, 15, 30, 60 and 120 min for M-GSH 1a; 0, 2, 4, 6, 8 and 12 h for AE-GSH 2a and 1, 3, 5 and 9 min for bis-GSH 3a. To quench to reaction for each time point, the reaction tubes were added with 0.05% trifluoroacetic acid to quench the reaction.



Scheme 2.1. Abbreviations for the chemical structures for the metabolites.

Results

The goal of performing *in vitro* metabolic reactions with Tp, including direct GSH reaction and enzymatic reactions, was to scan through the possible metabolites and to characterize the their structures. LC-MS/MS methods were developed accordingly based on the metabolites detected in the *in vitro* studies in order to apply the methods as an analytical platform to analyze the metabolites *in vivo*.

Direct reactions. We begin our experiment by directly reacting 0.1 mM thymine propenal with 1 mM GSH in phosphate buffer pH 7.4 at 37 °C. The reaction is very rapid with the half life of thymine propenal of ~105 s (Figure 2.1). Three reaction products are observed: M-GSH 1a, AE-GSH 2a and bis-GSH 3a, as shown in the chromatogram in Figures 2.2A and 2.2B. The two peaks observed for 1a stem from two enantiomers, which are not stable and transform into two sets of compounds, including AE-GSH 2a and thymine propenal 6 via β-elimination through E1cB mechanism, with a half-life ($t_{1/2}$) ~47 min (Figure 2.2C and Figure 2.3). The level of AE-GSH 2a, which is indirectly derived from quantitation of thymine based on 1:1 stoichiometry (Figure 2.2C), and that of thymine propenal 6 are slightly different, with the level of thymine propenal ~1.5 times higher than AE-GSH. This is probably caused by similar pKa values of GSH (pKa ~ 9.2, [19]) and thymine (pKa ~ 9.9, [20]) that make GSH and thymine act similarly as leaving groups. The instability of Michael-type adducts was also observed in a 1,4-butylthiol adduct in an experiment performed in the 1980s [21]. For the purpose of structural characterization of M-GSH 1a, $NaBH_4$ was used to reduce the aldehyde to an alcohol to make it isolable for characterization by ^{1}H-NMR (Table 2.2). By contrast, AE-GSH 2a is stable (Figure 2.2D), and the *trans* isomer of 2a is confirmed by ^{1}H-NMR with a coupling constant of $J_{H_\alpha H_\beta}$ = 15 Hz (Table 2.2) and a spectrum that is in agreement with previous work [22]. 3a is unstable and rapidly transforms to 2a via β-elimination (Figure 2.3) with a half life of ~8 min for the formation of 2a. This mechanism of transformation is similar to

that of 1a. Thus, the reaction of Tp with GSH catalyzed by equine GSTs was conducted to test if 1a is an enzymatic product. To do so, the relative quantitation of 1a and 2a was conducted by HPLC coupled to a UV/vis detector at λ_{max} = 260 nm and 290 nm for characterization of 1a and 2a, respectively. We found that the level of 1a catalyzed by GSTs ($1a^{GST}$) is approximately two-times higher (Figure 2.4A), while 2a formation is not different regardless of the presence of GSTs (Figure 2.4B). This result suggests that Tp is a substrate of GSTs, which is in agreement with the previous work [4], and that 1a is the enzymatic product from the reaction. However, 1a can also form in the nonenzymatic reaction as suggested by our results.

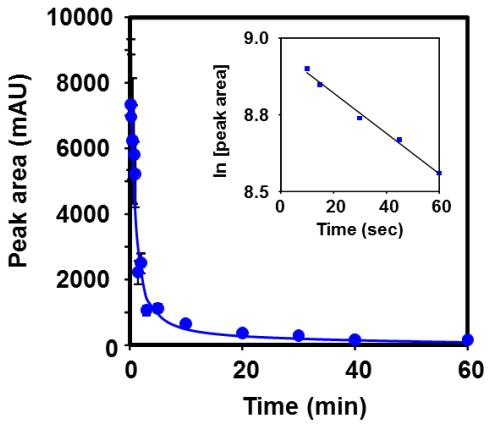


Figure 2.1. Relative quantitation of thymine propenal 6 loss in a reaction with GSH. The half life ($t_{1/2}$) of 6 in this reaction is ~105 s is obtained from the initial condition (loss <10%) of 6 during a reaction running from 0 to 60 s.

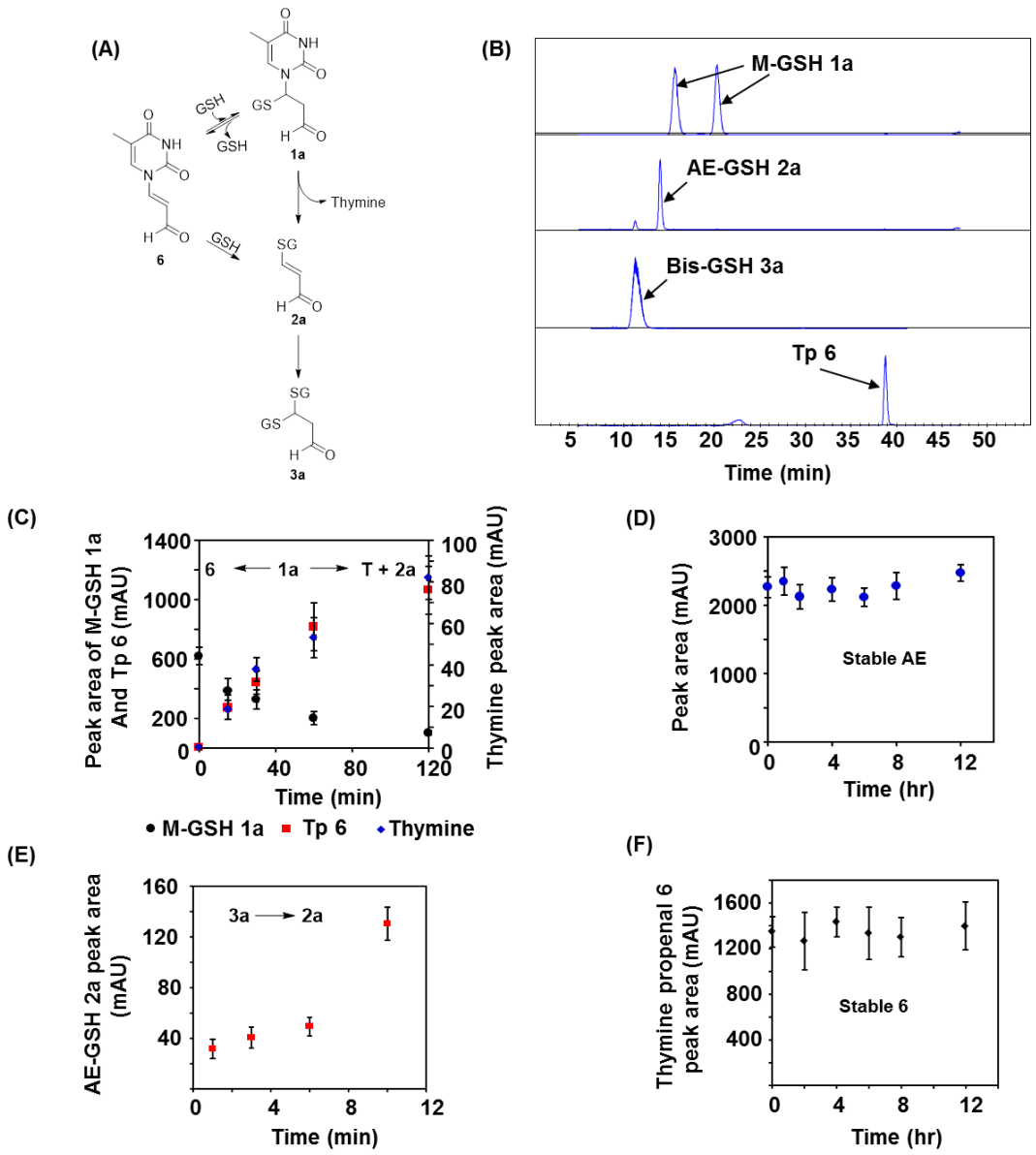


Figure 2.2. Metabolites generated from direct reaction of 0.1 mM thymine propenal 6 with 1 mM GSH. (A) Chemical structures of three GSH-conjugated metabolites detected from the reaction including: M-GSH 1a, AE-GSH 2a and bis-GSH 3a. (B) Extracted ion chromatogram with base-line separation of the three 1a to 3a as well as thymine propenal 6. (C) Stability of M-GSH 1a, which is converted to AE-GSH 2a and thymine propenal 6. (D) AE-GSH 2a is stable. (E) Formation of AE-GSH 2a from bis-GSH 3a. (F) Thymine propenal 6 is stable.

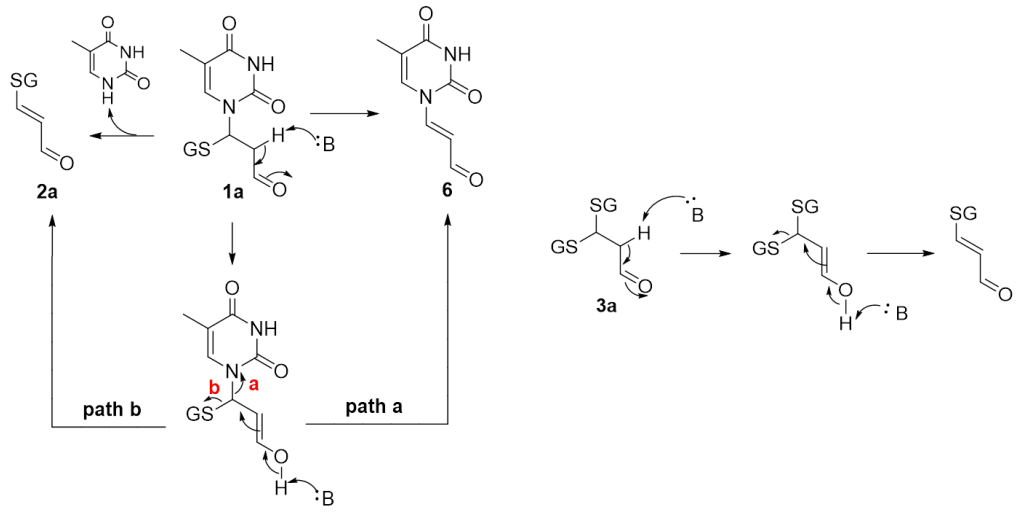


Figure 2.3. β-Elimination through E1cB mechanism of **1a** and **3a**

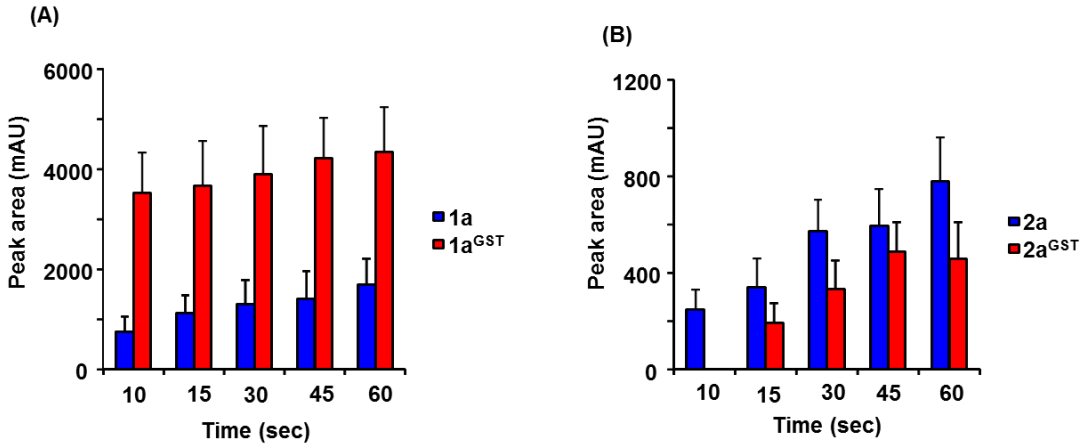


Figure 2.4. Relative quantitation of M-GSH 1a and AE-GSH 2a in the direct reaction compared with that of $1a^{GST}$ and $2a^{GST}$ from the GST-catalyzed reaction by LC-UV, suggesting that 1a is the enzymatic product. (A) Peak area (mAU) of 1a from direct reaction and $1a^{GST}$ from GST-catalyzed reaction; the UV signal is measured at λ_{max} = 260 nm (B) Peak area (mAU) of 2a from direct reaction and $2a^{GST}$ from GST-catalyzed reaction; the UV signal is measured at λ_{max} = 290 nm.

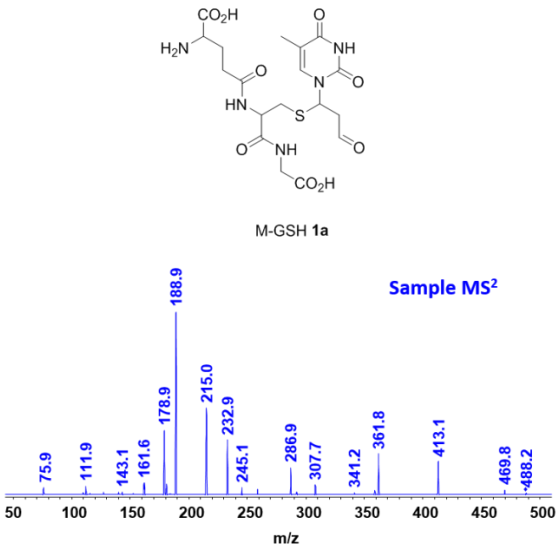


Figure 2.5. MS^2 of M-GSH 1a.

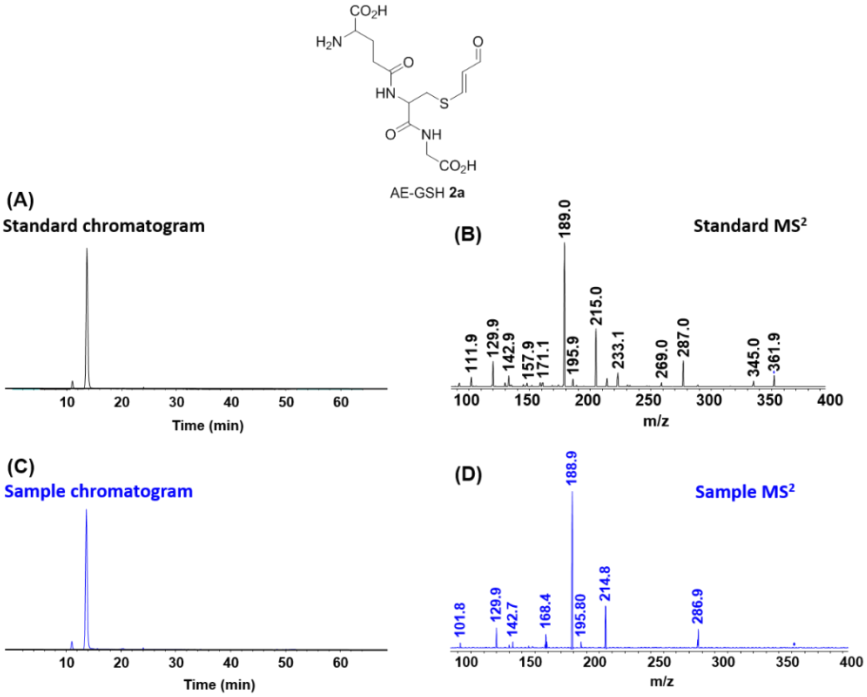


Figure 2.6. Comparison of retention time and MS^2 of **2a** from standard ((A) and (B), black highlighted text) to those of **2a** from the direct reaction ((C) and (D), blue highlighted text)

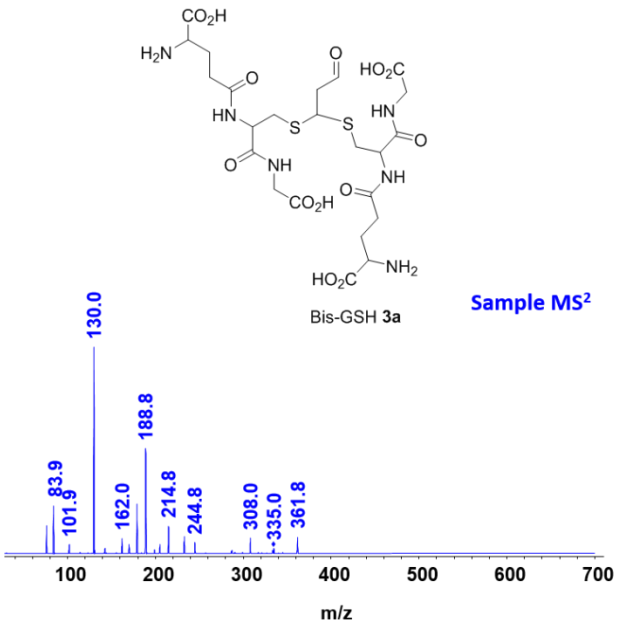


Figure 2.7. MS^2 of **3a**

Reactions with liver extracts. The instability of 1a and 3a implies that the reaction product from direct reaction should not be used as potential candidates for analysis *in vivo*. We next moved to perform the reaction of thymine propenal and GSH in rat liver extracts including microsome and cytosol. Microsomal fraction exhibits little metabolic activity to thymine propenal, since the product spectrum is similar to that of the direct reaction (Figure 2.9B). Conversely, the metabolism in the cytosolic fraction is so rapid that we could not detect the residual Tp as well as the metabolites from direct reaction even after immediate mixing (Figure 2.8). Therefore, Tp is extensively metabolized in liver cytosol and the reaction generates different types of metabolites (Figure 2.8) ranging from oxidation/reduction products to different GSH conjugations, including $M\text{-}GSH^{red}$ 1b, $AE\text{-}GSH^{red}$ 2b, $bis\text{-}GSH^{red}$ 3b, Tp^{red} 4 and Tp^{ox} 5. The structures of these products were confirmed by their high resolution mass spectrometry, CID fragmentation patterns and ^{1}H-NMR (Table 2.2). Although the amount of 3b is not enough for ^{1}H-NMR characterization, we confirmed the structure by comparing the retention time of $NaBH_4$-reduced 3a (Figure 2.12) in addition to its high resolution mass and MS^2. The reduction of the aldehyde to an alcohol in compounds 1a, 2a and 3a to

yield compounds 1b, 2b and 3b, respectively, is most likely derived from the aldose reductase superfamily of enzymes [23], which in turn converts the unstable 1a and 3a to the stable products. The result suggests the existence of defensive mechanisms to neutralize reactive products and preventing them from further reacting with other nucleophilic biomolecules such as DNA and proteins. Since all of the metabolites detected in the liver extract reactions are stable, the result from the these studies helps us predict that metabolites 1b to 3b, and 4 and 5 should arise *in vivo*.

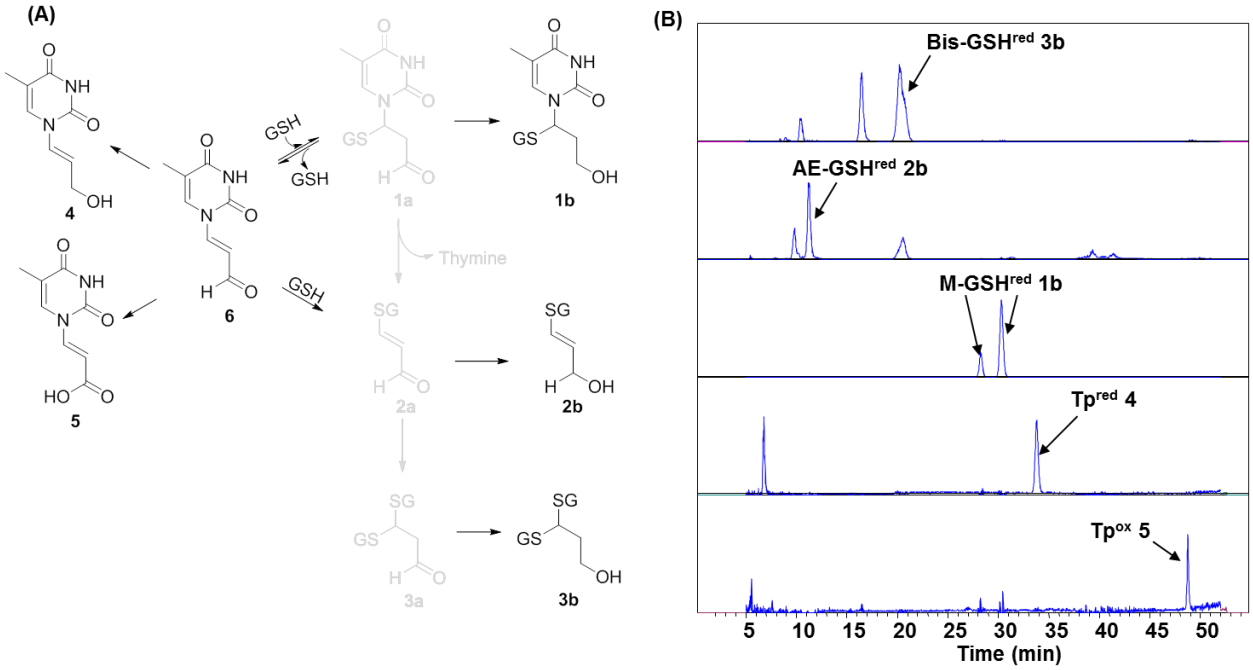


Figure 2.8. The reaction of 0.1 mM thymine propenal with 1 mM GSH in liver cytosol fraction. (A) Metabolites generated from the reaction including 1b – 3b and 4 and 5; 1a – 3a as well as starting material 6, highlighted in grey, are not detected in the reaction. (B) EIC with base-line separation of the products from the reaction; the structures for each product are confirmed with synthesized standards as described below.

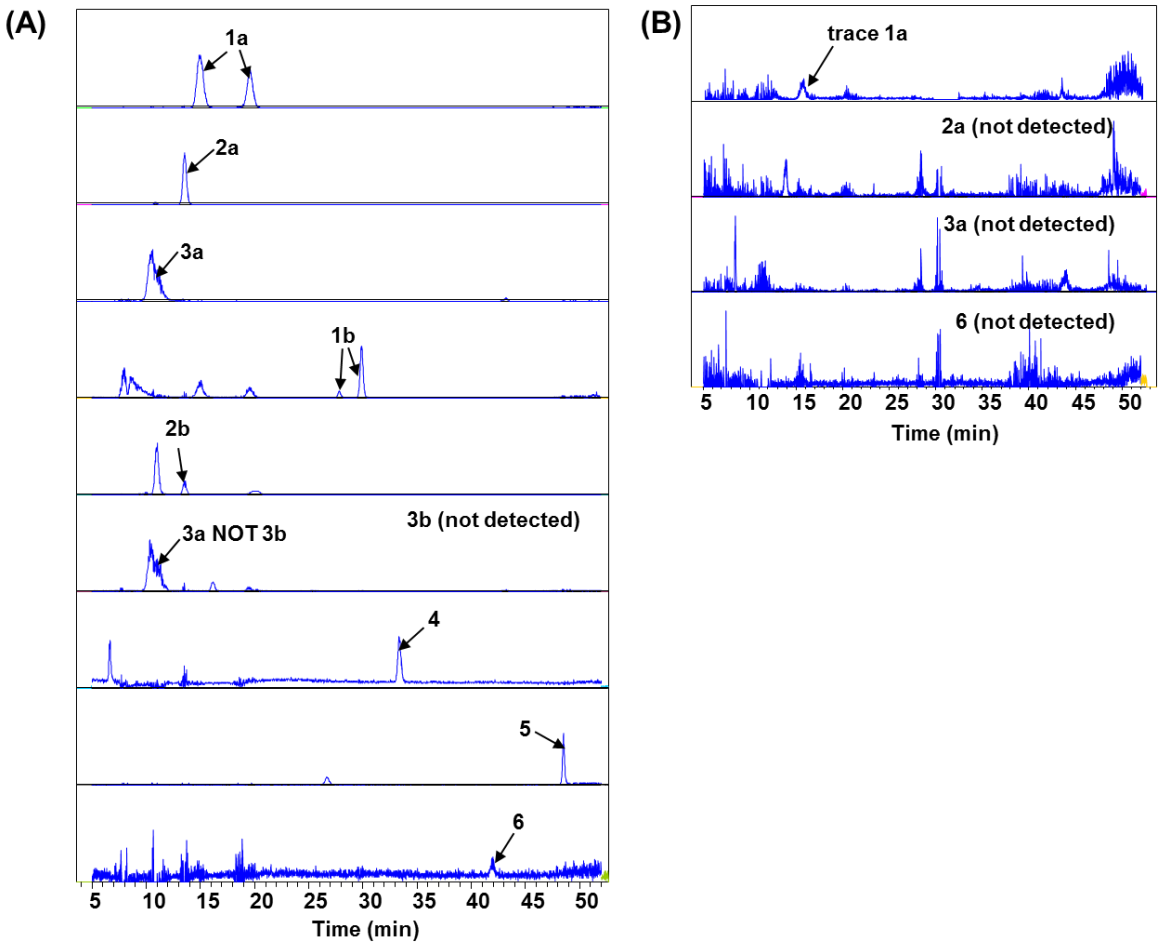


Figure 2.9. Metabolites from reaction in rat liver extracts**.** (A) Metabolites detected in microsomal fraction from the reaction of thymine propenal with GSH; the metabolism in this fraction is not extensive since 1a – 3a are detected. (B) 1a – 3a as well as 6 (starting material) are not detected in cytosolic fraction from the reaction of thymine propenal with GSH, indicating that the metabolism is very effective in this fraction

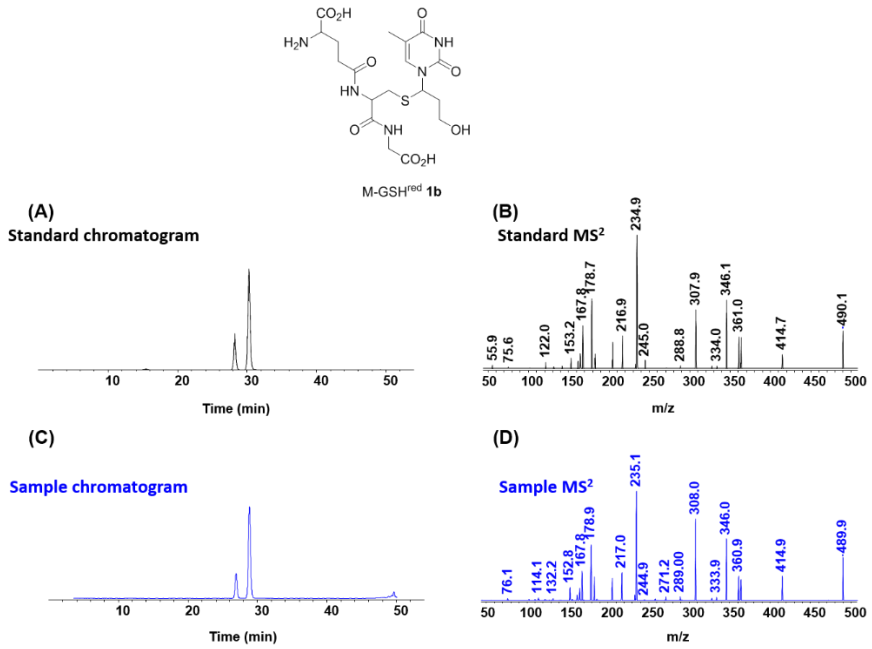


Figure 2.10. Comparison of retention time and MS^2 of 1b from standard to those of 1b from liver cytosol reaction

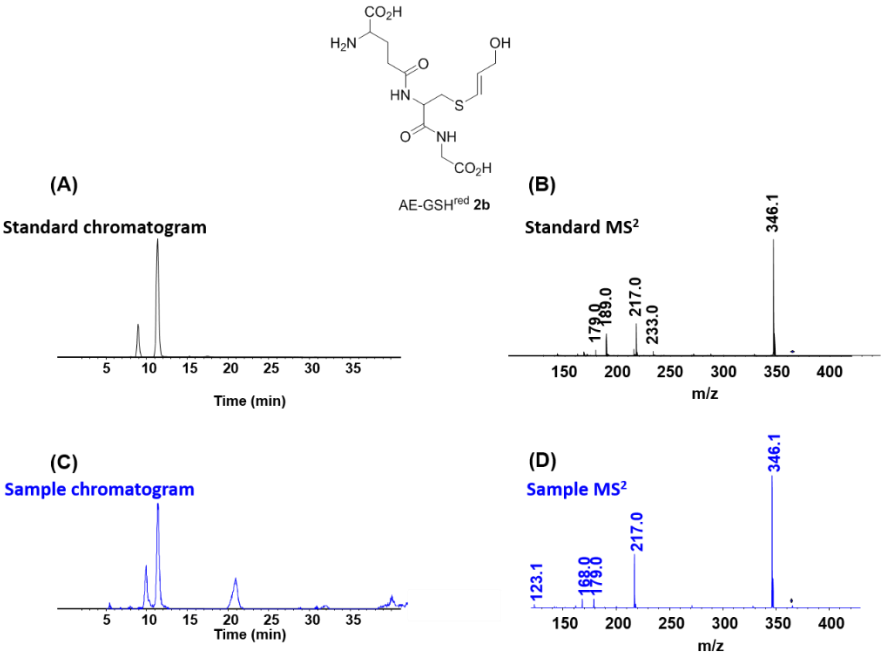


Figure 2.11. Comparison of retention time and MS^2 of 2b from standard to those of 2b from liver cytosol reaction

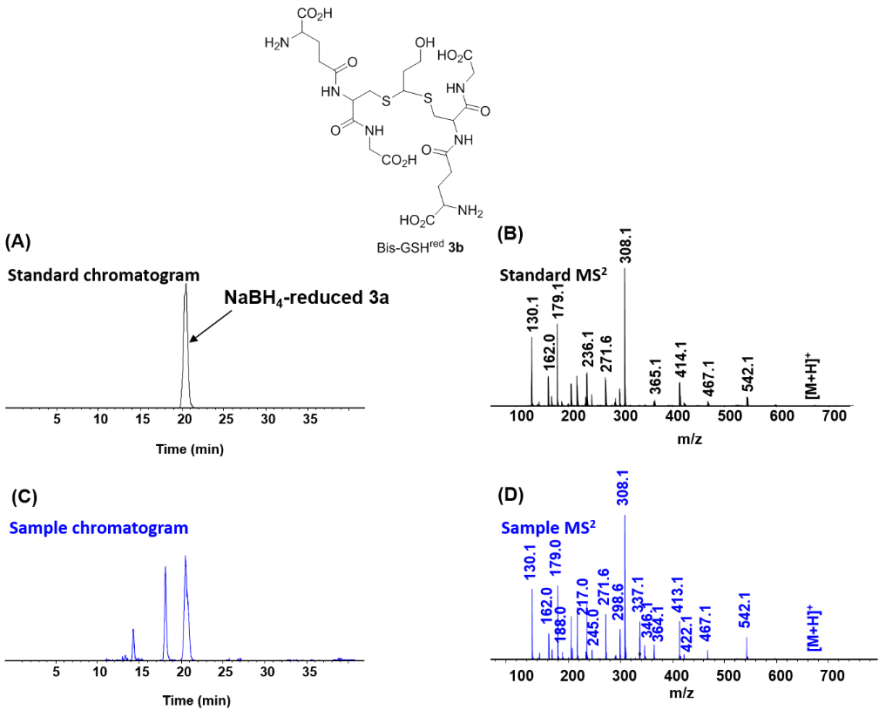


Figure 2.12. Comparison of retention time and MS² of **3b** from standard to those of **3b** from liver cytosol reaction

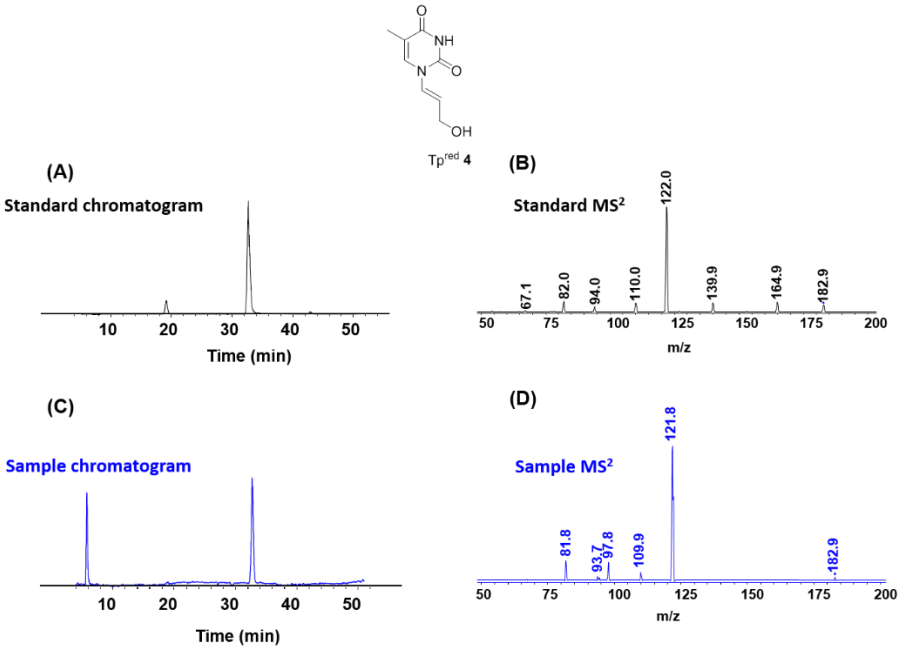


Figure 2.13. Comparison of retention time and MS² of **4** from liver cytosol reaction

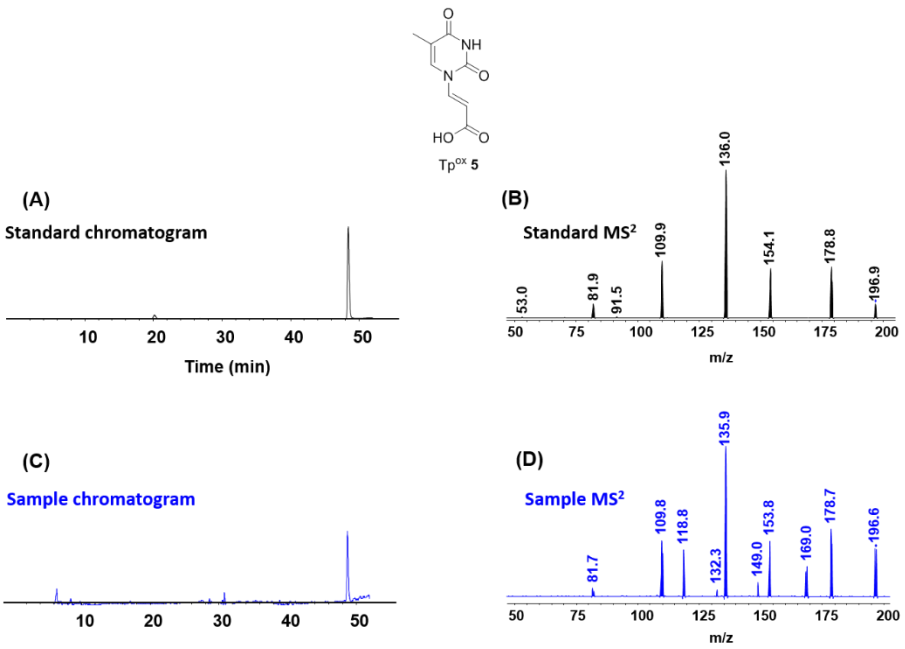
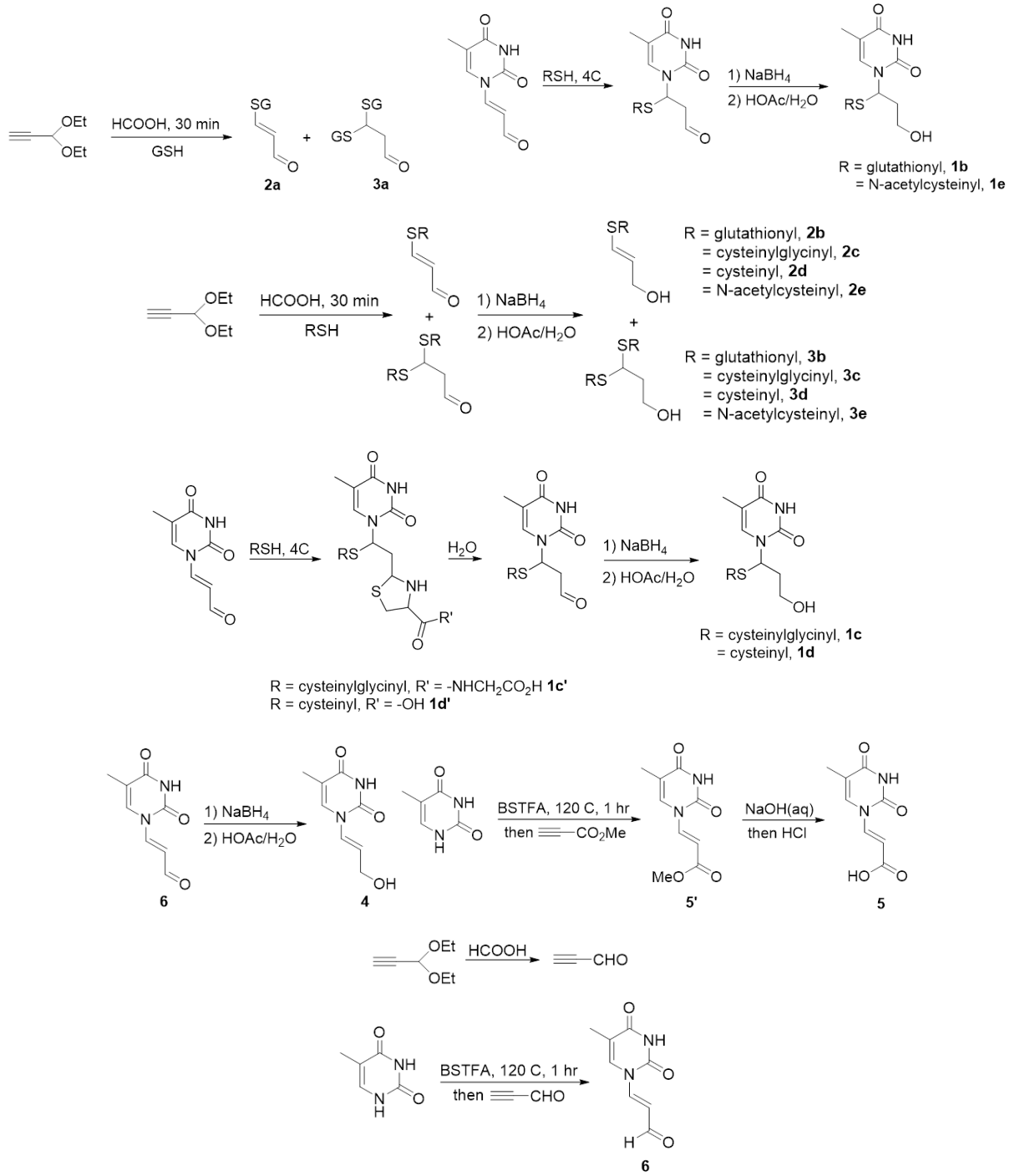


Figure 2.14. Comparison of retention time and MS^2 of 5 from standard to those of 5 from liver cytosol reaction

Synthesis of chemical standards. The synthesis procedures of the chemical standards for 1b-1e, 2b-2e, 3b-3e, 4-6 are described in Materials & Methods and illustrated in Scheme 2.2. LC-MS/MS methods were developed and the parameters are shown in Table 2.1.



Scheme 2.2. Chemical synthesis diagram for the standards

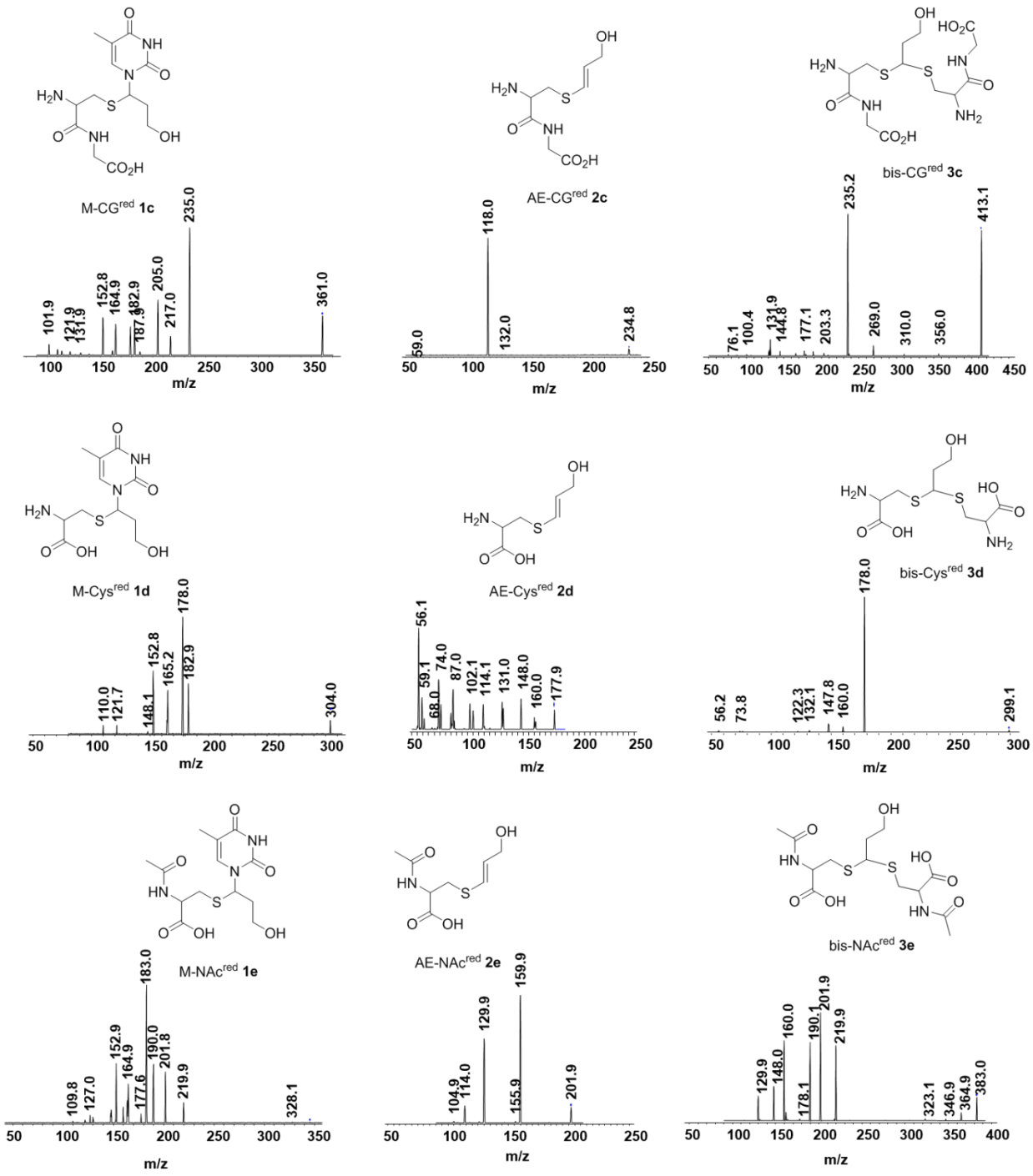


Figure 2.15. MS^2 of chemical standards of 1c – 3c, 1d – 3d and 1e – 3e.

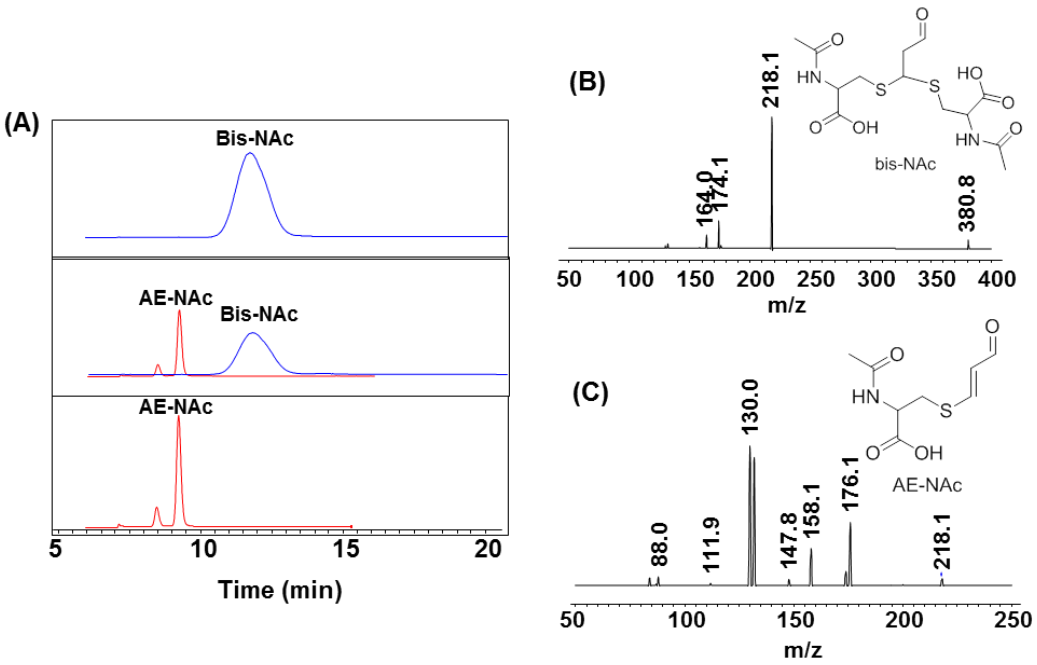


Figure 2.16. Structural confirmation of AE-NAc and bis-NAc. (A) bis-NAc in the aldehyde form is unstable and transforms to AE-NAc of which the structure is known by ^{1}H-NMR (D_2O); δ1.85 (CH3 at NAc, s, 3H), δ3.20 (-S*CH*$_2$CH(COOH)NHCOCH$_3$, dd, 2H), δ4.62 (-SCH$_2$*CH*(COOH)NHCOCH$_3$, t, 1H), δ6.29 (-SCH=*CH*CHO, dd, J_{HH} = 15 Hz (trans), J_{HCHO} = 8 Hz (long range), 1H), δ7.98 (-S*CH*=CHCHO, d, J_{HH} = 15 Hz (trans), 1H), δ9.25 (-SCH=CH*CHO*, d, J = 8 Hz (long range), 1H), . (B) MS/MS fragmentation of bis-NAc showing the loss of N-acetylcysteine to form AE-NAc (m/z = 218) as the base peak followed by carbonyl loss (hallmark of aldehyde functional group at m/z = 174) and N-acetylcysteine peak (m/z = 164). (C) MS/MS fragmentation of AE-NAc showing carbonyl loss (m/z = 176) followed by decarboxylation giving rise to m/z = 130.

Discussion

There are three steps in developing biomakers [24] involving target selection for biomarkers, method development for investigation of the targets and method validation, and application of the method in the pathophysiology-related condition for finding the correlation between the biomarkers and the diseases. This chapter of the thesis involves the first step and the half process of the second step. DNA damage products are chosen in this thesis projects due to the reasons described in Chapter 1. LC-MS provides a good choice for development of analytical tools because of its high sensitivity after optimization with targeted compounds.

In this chapter, we performed a systematic analysis of Tp metabolites based on chemical principles of *in vitro* reactivity and well characterized enzymatic metabolic reactions. For base propenals, the well established reactivity of the α,β-unsaturated carbonyl immediately provided candidate metabolites in the form of conjugates with GSH and other nucleophiles, as well as oxidation and reduction of both the aldehyde and carbon-carbon double bond. We firstly exploit scan mode of MS albeit the low sensitivity; however, this low sensitivity will provide us with helpful information in terms of the predominant metabolites generated in the direct reactions of Tp and GSH, where the concentration of Tp is high (0.1 mM). The observation of both **1a** and **2a** in direct reactions of Tp with GSH is contradictory to previous studies showing **2a** as the only product in this reaction and **1a** arising in GST-catalyzed reactions [4]. Reaction of Tp with GSH and equine GSTs revealed similar levels of **2a** compared to the direct reaction but a 2-fold higher level of **1a** (**Fig. 2A**). While these results confirm that Tp is an efficient GST substrate [4], they go further to demonstrate that **1a** is the major GST-catalyzed product and that both **1a** and **2a** form in direct reactions. Metabolism in liver cell extracts added another level of complexity, with the reduction of the aldehyde in compounds **1a**, **2a** and **3a** to an alcohol in **1b**, **2b** and **3b**, consistent with the activity of the aldose reductase superfamily enzymes [23], which in turn converts the unstable **1a** and **3a** to the stable products **1b** and **3b**. The fact that the reactive α,β-unsaturated carbonyl structure of Tp is lost in all of the metabolites produced *in vitro* and *in vivo* attests to the detoxification function of physiological metabolism. At the end of the *in silico* and *in vitro* analyses, we were left with 15 metabolite candidates that take into account the predicted processing of GSH conjugates to CG, Cys and mercapturic acid derivatives in liver and kidney.

After scanning through all the predicted metabolites arising *in vitro*, the chemical standards for candidates for analysis *in vivo* need to be synthesized. The reason is that high sensitivity of LC-MS/MS is required for the analysis in biological samples whose contains complex matrix and interference, which is however limited to the optimized parameter (i.e., fragmentation voltage,

collision energy, etc.) with known standards. Some of the synthesized standards were prepared from different reactions described in the direct reaction section. For example, AE-GSH 1b was prepared by the coupling reaction of propargyl aldehyde, prepared from deprotection of its acetal form, with GSH. This way of synthesis would provide us with more percent yield of the compound and make it easy to collect high amount of the standard for NMR confirmation.

Overall, we successfully identified the candidates for biomarker discovery by means of performing metabolite prediction through *in vitro* reactions and then synthesized chemical standards to optimize LC-MS/MS methods for investigation and quantitation *in vivo*. Furthermore, chemical standards would help with the retention time confirmation when analyzed with biological samples *in vivo* as described in Chapter 3.

Table 2.1. Summary of fragmentation voltage (V) and collision-induced dissociation energy (V) applied to LC-MS/MS in MRM mode.

Compound	Fragmentor (V)	CID (V)	MRM
M-GSH^{red} **1b**	128	16	490 → 235
AE-GSH^{red} **2b**	128	16	364 → 168
Bis-GSH^{red} **3b**	94	12	336 → 130
Tp^{red} **4**	95	12	183 → 122
Tp^{ox} **5**	80	17	197 → 136
Tp **6**	89	11	181 → 110
M-CG^{red} **1c**	90	4	361 → 235
AE-CG^{red} **2c**	78	16	235 → 118
Bis-CG^{red} **3c**	132	16	413 → 235
M-Cys^{red} **1d**	78	4	304 → 178
AE-Cys^{red} **2d**	86	8	178 → 148
Bis-Cys^{red} **3d**	84	8	299 → 178
M-NAc^{red} **1e**	78	4	346 → 183
AE-NAc^{red} **2e**	69	4	202 → 160
Bis-NAc^{red} **3e**	92	9	383 → 202

Table 2.2. Summary of the data for structural characterization by HR-MS, MS/MS and ^{1}H-NMR

Compound	$[M+H]^+_{obs}$	$[M+H]^+_{calc}$	Δmass (ppm)	MS/MS fragmentation	^{1}H-NMR (δ, ppm)	Comment
M-GSH, **1a**	488.1442	488.1446	0.67	470 (-H_2O), 362 (-thymine), 189 (methylene CysGly), 233 (-thymine, -glutamate)	-	Confirmed with **1b**. **1a** is unstable; so that it is confirmed with $NaBH_4$-reduced M-GSH (**1b**)
AE-GSH, **2a**	362.1018	362.1016	-0.37	287 (-CO_2H, -CHO), 215 (-thymine,-glutamate,-H_2O), 189 (methylene CysGly)	(D_2O) δ8.05 (GSC*H*=CH, d, 1H, J=15), δ6.35 (CH=C*H*CHO, dd, 1H, J=15), δ9.35 (C*H*O, d, 1H) δ2.25 ($HO_2CCH(NH_2)CH_2$- in glu, m, 2H), δ2.58 ($HO_2CCH(NH_2)CH_2CH_2$ in glu, t, 2H), δ3.22 (CH_2SH in cys, m, 2H), δ3.82 ($HO_2CCH(NH_2)CH_2CH_2$ in glu, t, 1H), δ4.02 ($NHCH_2CO_2H$ in gly, s, 2H)	λ_{max} = 290 nm, confirmed conjugation with sulfur attached to β-carbon.
Bis-GSH, **3a**	669.1852	669.1855	0.36	362 (-GSH), 130 (Glutamate), 189 (methylene CysGly)	-	Confirmed with **2a** by isolating **3a** and let it converts AE-GSH (**2a**). t_R is matched
M-GSH^{red}, **1b**	490.1598	490.1602	0.94	346 (-thymine,-H_2O), 308 (GSH), 361 (-pyroglutamate), 235 (-thymine,-pyroglutamate), 179 (CysGly)	(D_2O) δ1.90 (CH_3 in thymine, s, 3H), δ7.90 (H vinylic in thymine, s, 3H), δ2.10-2.30 ($HO_2CCH(NH_2)CH_2$- in glu and –$SCH(T)CH_2CH_2OH$, m, 4H), δ2.60 ($HO_2CCH(NH_2)CH_2CH_2$ in glu, t, 2H), δ3.10 (CH_2SH in cys, m, 2H), δ3.80 ($HO_2CCH(NH_2)CH_2CH_2$ in glu, t, 1H), δ3.90 ($NHCH_2CO_2H$ in gly, s, 2H), δ4.41 ($NHCHCH_2SH$) in cys, t, 1H), δ3.60 (-CH_2CH_2OH, t, 2H)	-
AE-GSH^{red}, **2b**	364.1177	364.1173	-1.1	346 (-H_2O), 217 (-pyroglutamate,-H_2O)	-	**2b** was generated by treating **2a** with $NaBH_4$
Bis-GSH^{red}, **3b**	336.1066 $[M+2H]^+$	336.1065 $[M+2H]^+$	-0.7	542 (-pyroglutamate), 308 (GSH), 364 (-GSH)	-	**3b** was generated by treating **3a** with $NaBH_4$
Tp^{red}, **4**	183.0763	183.0764	0.61	165 (-H_2O), 122 (methylene thymine), 110 (thymine – oxygen)	(D_2O) δ1.87 (CH_3 in thymine, s, 3H), δ7.95 (H vinylic in thymine, s, 3H), δ6.10 (-CH=C*H*CH_2OH, m, 1H), δ3.81 (=CHCH_2OH, t, 2H)	-
Tp^{ox}, **5**	197.0551	197.0557	2.94	179 (-H_2O), 154 (-CO_2), 136 (vinyl thymine-oxygen), 110 (thymine – oxygen)	(D_2O) δ1.87 (CH_3 in thymine, s, 3H), δ7.96 (H vinylic in thymine, s, 3H), δ6.20 (-CH=C*H*CO_2H, d, 1H), δ8.10 (-C*H*=CHCO_2H, d, 1H)	-
Tp, **6**	181.0607	181.0608	0.55	163 (-CO), 110 (thymine – oxygen), 55 (propenal)	(DMSO-d_6) δ1.87 (CH_3 in thymine, s, 3H), δ8.05 (H vinylic in thymine, s, 3H), δ6.10 (-CH=C*H*CHO, dd, 1H), δ8.15 (-C*H*=CHCHO, d, 1H), δ9.55 (CH=CHC*H*O, d, 1H)	-

Table 2.2. (Cont'd) Summary of the data for structural characterization by HR-MS, MS/MS and ^{1}H-NMR

Compound	$[M+H]^+_{obs}$	$[M+H]^+_{calc}$	Δmass (ppm)	MS/MS fragmentation	^{1}H-NMR (δ, ppm)	Comment
M-CG^{red}, **1c**	361.1178	361.1176	-0.83	235 (-thymine), 217 (-thymine,-H_2O), 205 (Vinyl CysGly)	-	**1c** is biochemically generated by specific cleavage of the peptide bond between glutamate and cysteine by means of treating **1b** with γ-glutamyl transferase
AE-CG^{red}, **2c**	235.0750	235.0747	-2.55	118 (Cysteinamide)	-	**2c** is biochemically generated by specific cleavage of the peptide bond between glutamate and cysteine by means of treating **2b** with γ-glutamyl transferase
Bis-CG^{red}, **3c**	413.1154	413.1159	1.21	235 (-CysGly)	-	**3c** is biochemically generated by specific cleavage of the peptide bond between glutamate and cysteine by means of treating **3b** with γ-glutamyl transferase
M-Cys^{red}, **1d**	304.0963	304.0962	-0.33	183 (thymine propenol), 178 (-thymine), 110 (thymine – oxygen)	-	**1d** is biochemically generated by specific cleavage of acetyl group from N-acetylcysteine by means of treating **1e** with acylase I
AE-Cys^{red}, **2d**	178.0529	178.0532	1.68	160 (-H_2O), 56 (allyl alcohol)	-	**2d** is biochemically generated by specific cleavage of acetyl group from N-acetylcysteine by means of treating **2e** with acylase I
Bis-Cys^{red}, **3d**	299.0735	299.0730	-1.67	178 (-Cys), 160 (-Cys,-H_2O), 56 (allyl alcohol)	-	**3d** was biochemically generated by specific cleavage of acetyl group from N-acetylcysteine by means of treating **3e** with acylase I
M-NAc^{red}, **1e**	346.1068	346.1067	-0.29	220 (-thymine), 202 (-thymine,-H_2O), 183 (thymine propenol), 127 (thymine), 110 (thymine – oxygen)	(CD_3OD) δ1.84-1.87 (CH_3 in thymine and CH_3CONH-Cys, 6H), δ8.10 (H vinylic in thymine, s, 3H), δ3.10 (CH_2SH in cys, dd, 2H), δ4.45 (-CONHCHCH_2SH) in cys, t, 1H), δ2.01-2.20 (-CHCH_2CH_2OH, m, 2H), δ3.31 (-CH_2CH_2OH, t, superimposed with CH_3OH)	-
AE-NAc^{red}, **2e**	202.0530	202.0532 (H_2O loss)	-2.47	160 (-NHCOCH_3), 130 (-NHCOCH_3,-CH_2OH)	-	**2e** was generated by treating AE-NAc (^{1}H-NMR described in **Figure S12**) with $NaBH_4$
Bis-NAc^{red}, **3e**	383.0915	383.0914	1.57	220 (-NAc), 202 (-NAc,-H_2O)	-	**3e** was generated by treating Bis-NAc (of which the structure confirmation described in **Figure S12**) with $NaBH_4$

References

1. Dedon, P.C., *The chemical toxicology of 2-deoxyribose oxidation in DNA.* Chem Res Toxicol, 2008. **21**(1): p. 206-19.
2. Alary, J., et al., *Mercapturic Acid Conjugates as Urinary End Metabolites of the Lipid-Peroxidation Product 4-Hydroxy-2-Nonenal in the Rat.* Chemical Research in Toxicology, 1995. **8**(1): p. 34-39.
3. Laurent, A., et al., *Metabolism of 4-hydroxynonenal, a cytotoxic product of lipid peroxidation, in rat precision-cut liver slices.* Toxicology Letters, 2000. **114**(1-3): p. 203-214.
4. Berhane, K., et al., *Detoxication of base propenals and other alpha, beta-unsaturated aldehyde products of radical reactions and lipid peroxidation by human glutathione transferases.* Proc Natl Acad Sci U S A, 1994. **91**(4): p. 1480-4.
5. Zhou, X., et al., *Quantification of DNA strand breaks and abasic sites by oxime derivatization and accelerator mass spectrometry: application to gamma-radiation and peroxynitrite.* Anal Biochem, 2005. **343**(1): p. 84-92.
6. Pouget, J.P., et al., *Formation of modified DNA bases in cells exposed either to gamma radiation or to high-LET particles.* Radiat Res, 2002. **157**(5): p. 589-95.
7. Frelon, S., et al., *High-performance liquid chromatography--tandem mass spectrometry measurement of radiation-induced base damage to isolated and cellular DNA.* Chem Res Toxicol, 2000. **13**(10): p. 1002-10.
8. Hayes, J.D., J.U. Flanagan, and I.R. Jowsey, *Glutathione transferases.* Annu Rev Pharmacol Toxicol, 2005. **45**: p. 51-88.
9. Mainwaring, G.W., J.R. Foster, and T. Green, *Nuclear and cellular immunolocalization of theta-class glutathione S-transferase GSTT1-1 in the liver and lung of the mouse.* Biochem J, 1998. **329 (Pt 2)**: p. 431-2.
10. Soboll, S., et al., *The content of glutathione and glutathione S-transferases and the glutathione peroxidase activity in rat liver nuclei determined by a non-aqueous technique of cell fractionation.* Biochem J, 1995. **311 (Pt 3)**: p. 889-94.
11. Tirmenstein, M.A. and D.J. Reed, *Role of a partially purified glutathione S-transferase from rat liver nuclei in the inhibition of nuclear lipid peroxidation.* Biochim Biophys Acta, 1989. **995**(2): p. 174-80.
12. Wu, G., et al., *Glutathione metabolism and its implications for health.* J Nutr, 2004. **134**(3): p. 489-92.
13. Cohen, G. and P. Hochstein, *Glutathione Peroxidase: The Primary Agent for the Elimination of Hydrogen Peroxide in Erythrocytes.* Biochemistry, 1963. **2**: p. 1420-8.
14. Shan, X.Q., T.Y. Aw, and D.P. Jones, *Glutathione-dependent protection against oxidative injury.* Pharmacol Ther, 1990. **47**(1): p. 61-71.
15. Strange, R.C., et al., *Glutathione-S-transferase family of enzymes.* Mutat Res, 2001. **482**(1-2): p. 21-6.
16. Klaassen, C.D., L.J. Casarett, and J. Doull, *Casarett and Doull's toxicology : the basic science of poisons*. 8th ed. 2013, New York: McGraw-Hill. p.
17. Gorgues, A. and A. Le Coq, *Préparation et utilisation d'un nouveau synthon acétylénique symétrique ou dissymétrique : le monoacétal de l'acétyl'˜enedicarbaldéhyde.* Tetrahedron Letters, 1979. **20**(50): p. 4825–4828.
18. Johnson, F., et al., *Synthesis and biological activity of a new class of cytotoxic agents: N-(3-oxoprop-1-enyl)-substituted pyrimidines and purines.* J Med Chem, 1984. **27**(8): p. 954-8.
19. Tang, S.S. and G.G. Chang, *Kinetic characterization of the endogenous glutathione transferase activity of octopus lens S-crystallin.* J Biochem, 1996. **119**(6): p. 1182-8.

20. Ganguly, S. and K.K. Kundu, *Protonation/Deprotonation Energetics of Uracil, Thymine, and Cytosine in Water from Emf/Spectrophotometric Measurements.* Canadian Journal of Chemistry-Revue Canadienne De Chimie, 1994. **72**(4): p. 1120-1126.
21. Lough, C.E., D.J. Currie, and H.L. Holmes, *Rates of Reaction of N-Butanethiol with Some Conjugated Heteroenoid Compounds.* Canadian Journal of Chemistry, 1968. **46**(5): p. 771-&.
22. Grollman, A.P., et al., *Origin and cytotoxic properties of base propenals derived from DNA.* Cancer Res, 1985. **45**(3): p. 1127-31.
23. Sobrattee, S., et al., *Metabolism of base propenals in the heart: The role of aldose reductase in detoxifying mutagenic and cytotoxic products of oxidative DNA damage.* Journal of Molecular and Cellular Cardiology, 2001. **33**(6): p. A112-A112.
24. Dedon, P.C., et al., *Challenges in developing DNA and RNA biomarkers of inflammation.* Biomark Med, 2007. **1**(2): p. 293-312.

Chapter 3

Metabolism of thymine propenal in Sprague-Dawley rats

Abstract

Oxidative stress and inflammation generate reactive chemical species that damage all types of cellular molecules, with this pathology mechanistically linked to disease. Yet damaged molecules have not emerged as significant biomarkers of these stresses in the same way that C-reactive protein, for example, has been exploited as a marker for systemic inflammation and cardiovascular disease risk. In addition to poor understanding of the damage actually arising in human tissues, the limited use of DNA lesions as biomarkers arises in part from a lack of appreciation of the fate of DNA damage products from the moment of formation in a tissue to release from cells to final excretion from the body. Here we have applied the principles of metabolite profiling to define the *in vitro* and *in vivo* biotransformation of a common product of DNA oxidation, thymine propenal (Tp), one of the base propenals arising from 4'-oxidation of DNA. Of 15 compounds predicted from *a priori* knowledge of chemical reactivity and demonstrated *in vitro* in rat liver cell extracts, analysis of urine samples from rats given intravenous doses of Tp revealed three major metabolites: two mercapturic acid derivatives and thymine propenoic acid, which accounted for ~6% of the dose. An additional four metabolites, including conjugates with glutathione, cysteinylglycine and cysteine, were observed in bile and accounted for ~22% of injected dose. While the parent Tp was never detected in blood, urine or bile, one of the major metabolites of Tp, a bis-mercapturic acid derivative of reduced Tp, appeared as a background excretory product in unstressed rats and was significantly elevated during bleomycin/CCl_4-induced oxidative stress. This points to the importance of defining the metabolism of DNA damage products in the effort to define biomarkers linking cancer to inflammation.

Introduction

Endogenous DNA damage has long been considered a potentially useful source of biomarkers of diseases associated with inflammation and oxidative stress given the mechanistic links to diseases such as cancer and the broad representation of pathophysiological chemistries, such as alkylation, oxidation, deamination, nitration and halogenation [1-6]. Major limitations to the development of DNA damage products as biomarkers involve the choice of a sampling compartment, with invasive sampling of tissues preventing large-scale human studies, and a lack of appreciation for the biological fate of DNA lesions following their formation in a cell. The latter problem is reflected in the potential for rapid DNA repair to maintain relatively low steady-state levels of DNA damage even in severely inflamed tissues [1]. Although there have been attempts to characterize DNA damage products released from cells into accessible compartments such as urine, blood or feces [7, 8], the biotransformational fates of the damage products, such as the metabolism after release by DNA repair, have not been rigorously assessed, which has the potential to allow important biomarker candidates to escape detection.

One notable exception arises in the recent studies of the metabolic fate of the 3-(2-deoxy-β-D-erythro-pentofuranosyl)pyrimido[1, 2-α]purin-10(3H)-one (M_1dG) adduct arising from reaction of of 2-deoxyguanosine (dG) with the endogenous electrophiles malondialdehyde and base propenals. When released from cells presumably by nucleotide excision repair, M_1dG is found at extremely low levels (10–20 fmol/kg) in human urine [9]. However, Marnett and coworkers discovered that M_1dG is oxidized by hepatic xanthine oxidase to generate 6-oxo-M_1dG, which is excreted to a small extent in urine and mainly in bile [10-12]. On the basis of this example, we have developed a logical and systematic approach of *in vitro* metabolic profiling to predict the *in vivo* metabolic fate of endogenous DNA lesions and other damage products, using the base propenal products of DNA oxidation to illustrate the approach.

In Chapter 3, we validate our LC-MS/MS methods developed from 15 synthesized metabolites mentioned in Chapter 2 by virtue of exogenously administering thymine propenal (Tp) to Sprague-Dawley rats via an intravenous (IV) route. The collected biological samples including urine and bile are subject to analysis by LC-MS/MS. The metabolites detected in those samples are confirmed with the synthesized standards. As described below in this Chapter, our metabolite prediction is relatively accurate and shown that the analytical method can be used with the biological samples from rat. At the end of this Chapter, endogenous induction of base propenals using bleomycin and CCl_4 is performed and the result shows that bis-NAc^{red} is the only metabolite found in rat urine.

Materials & Methods

Materials. The following chemicals were commercially obtained and used without further purification: thymine; 3,3-diethoxy-1-propyne; glutathione (GSH); potassium monohydrogen phosphate (K_2HPO_4); potassium dihydrogen phosphate (KH_2PO_4); *N,O-bis*-(trifluoromethylsilyl)trifluoroacetamide (BSTFA); liver microsomes from female Sprague-Dawley rats; human liver cytosol; β-nicotinamide adenine dinucleotide phosphate, reduced, sodium salt (NADPH); and glutathione S-transferases from equine liver were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO). Deionized water was further purified with a Milli-Q system (Millipore Corporation, Bedford, MA) and was used in all experiments. Chemical standards, 1a to 1e, 2a to 2e, 3a to 3e and 4 to 6, were synthesized and characterized as described in Chapter 2.

Instrumentation. UV-vis measurements were made on an HP8452 diode-array spectrometer (Agilent Technologies). HPLC-UV analyses were carried out on Agilent 1290 series with binary pumps and degassers. Full scan and collision-induced dissociation (CID) mass spectra of in vitro reaction products were obtained from an Agilent Technologies 6430 ion trap LC/MS coupled to Agilent HPLC 1200 series. Quantitation data is obtained from multiple reaction monitoring mode (MRM) from Agilent Technology 6430 QQQ coupled to Agilent HPLC 1290 series.

HPLC methods. Four HPLC methods were applied in the experiment and described as follows. *HPLC method 1.* An HPLC column was equilibrated with 100% solvent A (0.1% formic acid in H_2O) and 0%B (acetonitrile) at a flow rate of 0.200 mL/min at 30°C. The solvent was programmed as follows: a linear gradient from the starting solvent to 5% B in 30 min; a linear gradient increasing from 5%B to 46%B for 12 min; increasing to 100% B in 0.1min, holding for 10 min; decreasing to 0% B in 0.1 min; and reequilibrating at initial conditions for 15 min. *HPLC method 2.* ZORBAX Eclipse RRHD, XDB-C18 (2.1 × 100 mm), 1.8 μm (Agilent Technologies) was equilibrated with 95.5% solvent A (0.1% formic acid in H_2O) and 0.5%B (acetonitrile) at a flow rate

of 0.250 mL/min at 20°C. The solvent was programmed as follows: a linear gradient from the starting solvent to 13% B in 13 min; increasing to 100% B in 0.1min, holding for 10 min; decreasing to 0.5% B in 0.1 min; and reequilibrating at initial conditions for 5 min. *HPLC method 3.* ZORBAX Eclipse, XDB-C8 (4.6 × 250 mm), 5 μm (Agilent Technologies) was equilibrated with 95.5% solvent A (0.1% formic acid in H_2O) and 0.5%B (acetonitrile) at a flow rate of 0.300 mL/min at 20°C. The solvent was programmed as follows: a linear gradient from the starting solvent to 7% B in 12 min; increasing to 100% B in 0.1min, holding for 10 min; decreasing to 0.5% B in 0.1 min; and reequilibrating at initial conditions for 15 min. *HPLC method 4.* Dionex Acclaimed Polar Advantage, C18 (2.1 × 150 mm), 3 μm (Thermo Scientifics) was equilibrated with 100 % solvent A (0.1% formic acid in H_2O) and 0%B (acetonitrile) at a flow rate of 0.250 mL/min at 20°C. The solvent was programmed as follows: an isocratic elution of 0%B for 6 min; a linear gradient to 3% B in 13 min; a linear gradient from 3%B to 20%B in 10 min; increasing to 100% B in 0.1min, holding for 10 min; decreasing to 0% B in 0.1 min; and reequilibrating at initial conditions for 15 min.

Administration of Tp to Sprague-Dawley rats. Animal experiments were performed in accordance with the protocols approved by MIT Committee on Animal Care (CAC) and with the NIH *Guide for the Care and Use of Laboratory Animals*. Female Sprague-Dawley rats (225–250 g) with catheters surgically implanted into the jugular vein and bile duct were obtained from Charles River Laboratories and housed in shoebox cages. Prior to the experiment the animals were transferred into metabolic cages and fasted for 8 h before the beginning of dosing. The dosing solution was prepared in sterile saline solution, administered through the jugular vein an approximate volume of 0.25 ml, and flushed with 0.25 ml of sterile heparinized saline. Six animals were used in the experiment (three animals dosed with sterile saline as vehicle and three animals dosed with 1 mg/kg thymine propenal in sterile saline). All animals were housed in metabolic cages through the duration of the experiment to collect urine over the following intervals: pre-dose (8 h), 0 – 8, 8 –16, 16 – 24 h. Urine was collected in sterile polypropylene tubes over dry ice to reduce artifacts. Bile

was collected through biliary catheter over the following time points: 0.25, 0.5, 1, 2, 4 and 6 h. Samples were stored at –80 °C before analysis.

Sample workup. A sample of urine (50 □L) or bile (3 μL for the Tp-treated group and 20 μL for the saline-treated group) was filtered through 10 kD filter to remove proteins. The filtrate was then further pre-purified using HPLC method 1 with a ZORBAX Eclipse XDB C8 (4.6 × 150 mm), 5 □m (Agilent Technologies) analytical column. The fractions were divided into two portions: very polar fraction collected earlier and less polar fraction collected later. Each fraction was lyophilized by Speed Vac. The residuals were reconstituted in water prior to injection to LC-QQQ using HPLC method 2. The mass spectrometric method was described below.

Administration of bleomycin and carbon tetrachloride (CCl_4). Animal experiments were performed in accordance with the protocols approved by MIT Committee on Animal Care (CAC) and with the NIH *Guide for the Care and Use of Laboratory Animals*. Female Sprague-Dawley rats (225–250 g) were obtained from Charles River Laboratories and housed in shoebox cages. Prior to the experiment, the animals were transferred into metabolic cages and fasted for 8 h before dosing. Bleomycin solution (1 mg/kg) was prepared in sterile saline and administered intraperitoneally (IP) in a volume of 0.25 mL, while 800 mg/kg CCl_4 was administered in corn oil vehicle by IP injection. Six rats were used in this experiment: two dosed with sterile saline; two with bleomycin; and two with CCl_4. All animals were housed in metabolic cages through the duration of the experiment to collect urine over the following intervals: pre-dose (8 h), 0 – 4, 4 – 8, 8 –16, 16 – 24 h. Urine was collected in sterile polypropylene tubes over dry ice to reduce artifacts. Samples were stored at –80 °C before analysis.

LC-MS analysis of *in vivo* metabolites. LC-MS/MS analyses of urine and bile were conducted with a ZORBAX Eclipse XDB C8 (4.6 × 250 mm, 5 μm, Agilent Technologies) analytical column and ZORBAX Eclipse RRHD, XDB-C18 (2.1 × 100 mm, 1.8 μm, Agilent Technologies) for polar fraction and non-polar fraction with HPLC method 3 and 2, respectively. The HPLC was

coupled to an Agilent 6430 series QQQ mass spectrometer operating in positive ion mode. Nitrogen was the nebulizing and drying gas (15 psi, 10 L/min) and had a temperature set at 325 °C. Analyses were performed with the mass spectrometer set to multiple reaction monitoring (MRM) mode with the fragmentation voltages and collision energies shown in Table 2.2 in Chapter 2.

Results

Identifying Tp metabolites in urine and bile from Tp-treated rats. With the set of predicted Tp metabolites, we proceeded to assess the metabolism of Tp in rats exposed by intravenous (IV) administration through a jugular vein catheter. Sampling of blood and bile was performed through jugular vein and bile duct cannulae, respectively, with urine collected over dry ice in 8 h intervals in metabolic cages without food or water to contaminate the collection.

Of the 15 targeted compounds, analysis of urine samples revealed three major metabolites excreted within 8 h of IV administration of Tp: the mercapturic acid derivatives 1e (M-NAc^{red}) and 3e (bis-NAc^{red}) and the Tp oxidation product 5 (Figure 3.1). Tp was notably absent in the urine and, as expected, the cysteine, cysteinylglycine and GSH conjugated metabolites were not detected likely due to the high activity of γ-glutamyl transferases, CysGly dipeptidases and N-acetyl transferases in kidney. Quantitative analysis using external calibration curves based on synthetic standards revealed that 1e, 3e, and 5 accounted for ~6% of the IV injected dose. Interestingly, bis-NAc^{red} 3e is the only metabolite detected in the urine of saline-treated rats (Figure 3.1A, bottom panel) and was present at levels 40-times lower than the bis-NAc^{red} 3e adduct level detected in Tp-treated rats during 24 h of urine collection (Tables 3.1 and 3.2).

Given the low recovery of urinary Tp metabolites, we considered the possibility that Tp and its metabolites could be excreted in bile. Using biliary catheters, bile was collected periodically over 6 h following IV injection of Tp and analyzed by LC-MS/MS for the presence of the 15 targeted metabolites. Peaking in the first 30 min following injection, we detected not only the mercapturic acid metabolites 1e and 3e, but also the GSH, CG and Cys conjugation products 1b, 1c, 1d, 3d and 3e. Surprisingly, Tp^{ox} 5 was detected at levels ~36-times higher than the next most abundant metabolite 1b, which contrasts with the low level of 5 in urine (Figure 3.1C, 3.2C). In addition, M-NAc^{red} 1e and bis-NAc^{red} 3e were detected in the saline-treated rats (Figure 3.2A, bottom panel).

The biliary level of M-NAc^{red} 3e from Tp-injected rats relative to those treated with saline increases ~70-fold during 6 h of bile collection (Tables 3.3 and 3.4). In contrast to urine, the amount of bis-NAc^{red} 3e excreted in bile from the Tp-treated rats was not different from saline-treated controls. This might be due to the faster clearance rate of bis-NAc^{red} 3e than the Michael-type adducts (1b-e) in the biliary system. Again using external calibration curves, quantitative analysis of the biliary excretion products revealed that they accounted for ~22% of the IV injected dose.

Finally, analysis of blood samples as early as 5 min after IV injection revealed no detectable Tp (limit of detection ~10 fmol). This raised questions about the stability of base propenals in blood, so we undertook a series of *in vitro* experiments to assess the reactivity of Tp and Ap with serum components. Addition of Ap to commercial preparations of rat serum recapitulated the observations in the blood of Tp treated rats, with no detectable Ap in as little as 2-3 minutes after addition of Ap to serum. To assess the reactivity of base propenals with serum nucleophiles, the second-order rate constants (37 °C) for the reaction of Ap with bovine serum albumin, GSH, cysteinylglycine (CG) and cysteine (Cys) were found to be 0.15, 34, 38, and 14 $M^{-1}s^{-1}$, respectively. These results raise the possibility of a catalytic activity consuming base propenals in blood and serum or some kind of sequestration of the base propenals in lipids or proteins.

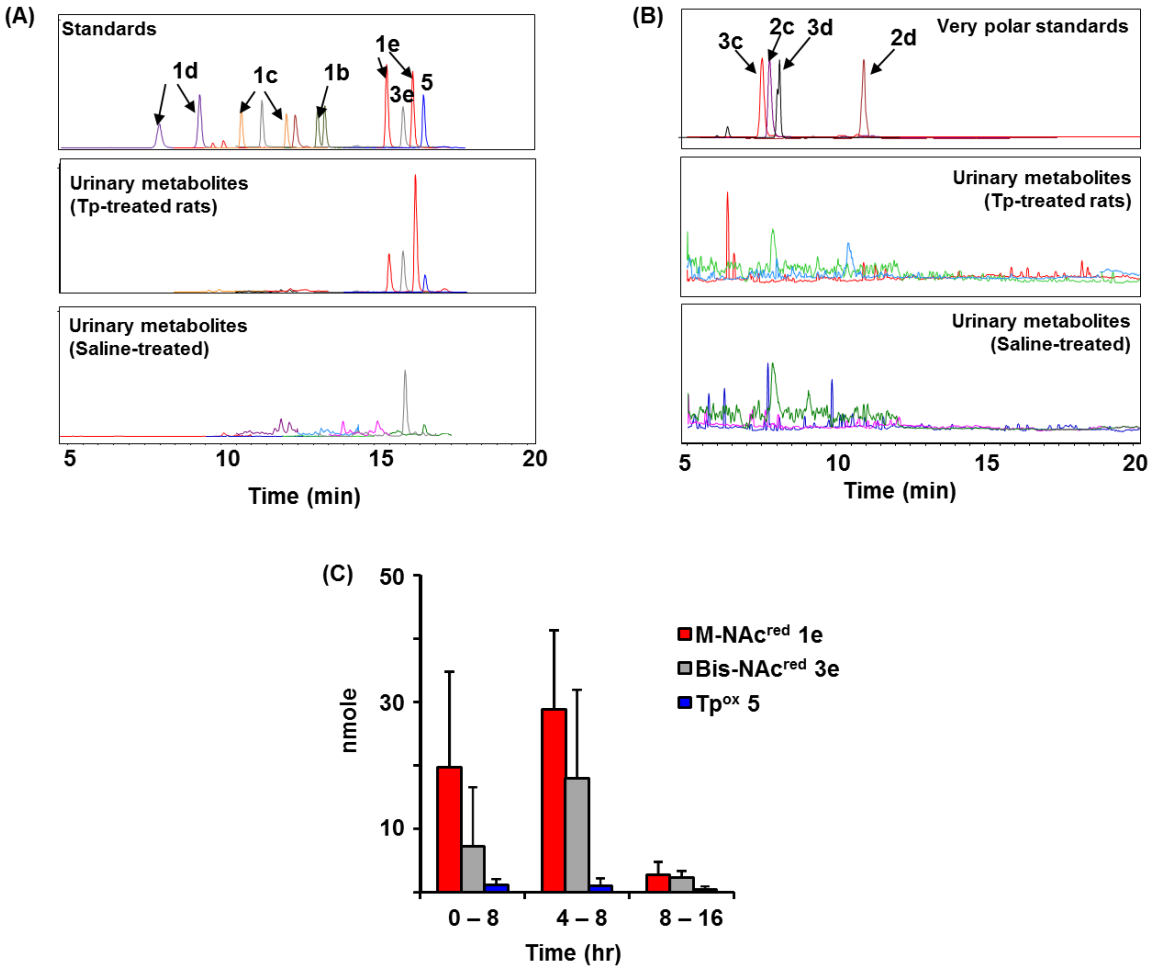


Figure 3.1. Urinary metabolites recovered following administration of thymine propenal 6. (A) MRM-EIC of relatively nonpolar compounds including 1b – 1e, 3e and 5; from synthesized standards (top); from rats treated with thymine propenal 6 (middle), three metabolites out of fifteen targeted metabolites detected in urine from rats (n = 3) treated with 6, including 1e, 3e and 5; and from rats treated with saline (bottom), 3e detected as background from saline-treated rats (n = 2). (B) MRM-EIC of polar compounds including 2c, 2d, 3c and 3d; from synthesized standards (top); from rats treated with thymine propenal 6 (middle); and from rats treated with saline (bottom). (C) Estimated quantitation of 1e, 3e and 6 from rat urine treated with 6 from three fraction of urine (0 – 8, 8 – 16 and 16 – 24 hr).

Table 3.1. Urinary excretion (nmole) of Tp metabolites in rats (n = 3) treated with Tp[1]

Metabolites	Metabolite abundance (nmole) for three durations of urine collection		
	0 – 8 hr	**8 – 16 hr**	**16 – 24 hr**
M-NAc^{red}	20 ± 15	29 ± 12	2.7 ± 2.0
bis-NAc^{red}	7.2 ± 9.3	18 ± 14	2.3 ± 1.0
Tp^{ox}	1.1 ± 1.0	1.0 ± 1.2	0.33 ± 0.55

[1]Tp (~1 mg/kg) was administered IV and urine was collected over dry ice for the noted periods after T_P administration. Metabolite concentration was determined by LC-MS/MS using external calibration curves. The error is standard deviation (SD).

Table 3.2. Urinary excretion (nmole) of Tp metabolites in rats treated with saline

Compound	Metabolite abundance (nmole) for three durations of urine collection		
	0 – 8 hr	**8 – 16 hr**	**16 – 24 hr**
Bis-NAc^{red}	0.25 ± 0.060	0.22 ± 0.17	0.23 ± 0.080

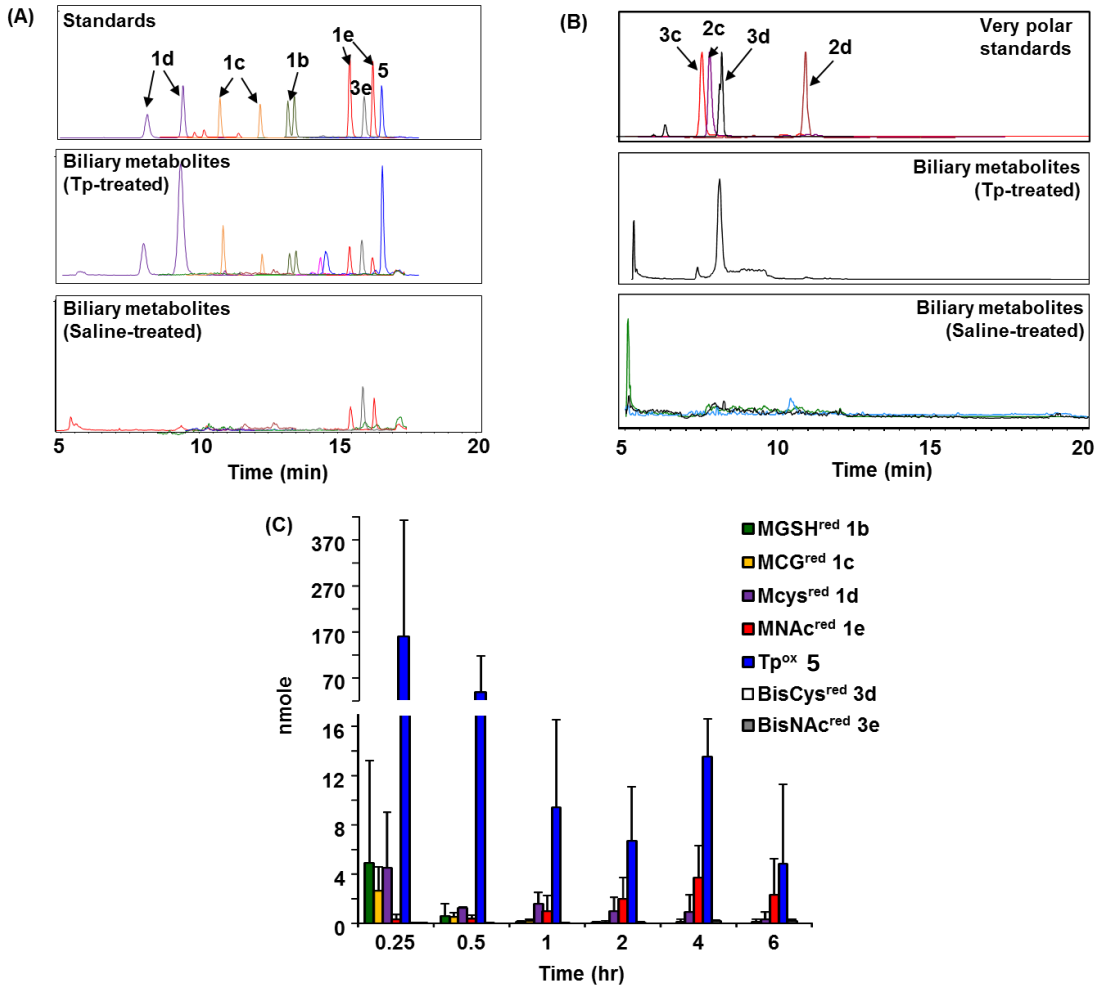


Figure 3.2. Biliary metabolites recovered following administration of thymine propenal 6. (A) MRM-EIC of relatively nonpolar compounds including 1b – 1e, 3e and 5; from synthesized standards (top); from rats treated with thymine propenal 6 (middle), six metabolites detected including 1b – 1e, 3e and 5; and from rats treated with saline (bottom), 1e and 3e detected as background from saline-treated rats (n = 2). (B) MRM-EIC of polar compounds including 2c, 2d, 3c and 3d; from synthesized standards (top); from rats treated with thymine propenal 6 (middle), 3d detected; and from rats treated with saline (bottom). (C) Estimated quantitation of 1b – 1e, 3e and 5 from rat urine treated with 6 from six fractions of bile (0.25, 0.5, 1, 2, 4 and 6 hr).

Table 3.3. Biliary excretion (nmole) of Tp metabolites in rats treated with Tp (n = 3)[1]

Compound	Metabolite abundance (nmole) for three duritions of bile collection					
	0.25 hr	**0.5 hr**	**1 hr**	**2 hr**	**4 hr**	**6 hr**
M-GSH^{red}	4.9 ± 8.3	0.61 ± 0.96	0.13 ± 0.078	0.043 ± 0.075	0	0
M-CG^{red}	2.6 ± 2.0	0.52 ± 0.34	0.19 ± 0.17	0.075 ± 0.13	0.13 ± 0.23	0.13 ± 0.23
M-Cys^{red}	4.5 ± 4.6	1.3 ± 0.10	1.6 ± 0.92	0.96 ± 1.2	0.92 ± 1.4	0.35 ± 0.57
M-NAc^{red}	0.30 ± 0.39	0.36 ± 0.32	0.96 ± 1.3	2.0 ± 1.7	3.7 ± 2.6	2.3 ± 3.0
Tp^{ox}	177 ± 252	55 ± 80	9.4 ± 7.1	6.7 ± 4.4	14 ± 3.1	2.3 ± 6.5
Bis-NAc^{red}	0.022 ± 0.017	0.021 ± 0.011	0.045 ± 0.032	0.073 ± 0.057	0.17 ± 0.11	0.17 ± 0.14
Bis-Cys$^{red, 2}$	0.043 ± 0.022	0.0033 ± 0.0057	0	0	0	0

[1]Tp (~1 mg/kg) was administered IV and urine was collected over dry ice for the noted periods after Tp administration. Metabolite concentration was determined by LC-MS/MS using external calibration curves. The error is standard deviation (SD).

[2]Ethyl cysteine is used for quantitation in external calibration.

Table 3.4. Biliary excretion (pmole) of Tp metabolites in rats treated with saline (n = 2)

Compound	Metabolite abundance (pmole) for three duritions of bile collection					
	0.25 hr	**0.5 hr**	**1 hr**	**2 hr**	**4 hr**	**6 hr**
Bis-NAc^{red}	162 ± 59	122 ± 29	248 ± 201	107 ± 18	93 ± 2.5	100 ± 0.89
M-NAc^{red}	13 ± 10	25 ± 32	44 ± 46	15 ± 11	23 ± 13	14 ± 9.7

***Tp metabolites as biomarkers of oxidative stress*.** The evidence for Tp metabolism in rats and the detection of a variety of Tp metabolites as urinary and biliary excretion products, including two putative endogenous species present in saline-treated rats, suggested that urinary Tp metabolites could serve as biomarkers of oxidatively damaged DNA and more broadly of oxidative stress. To test this hypothesis, rats were intraperitoneally (IP) treated with a DNA-cleaving and base propenal-generating anticancer drug, bleomycin, or with carbon tetrachloride (CCl_4) as a toxicant that induces widespread inflammation and oxidative stress. Analysis of urine samples reveals only one metabolite, bis-NAc^{red} 3e in urine from the saline-treated, bleomycin-treated and CCl_4-treated rats (Figure 3.3A). Quantitation using external calibration reveals that the urinary level of bis-NAc^{red} from the bleomycin- and CCl_4-treated rats is ~3 – 4 times higher than that from the saline-treated group (Figure 3.3B). Coupled with the observation of increased levels of bis-NAc^{red} in the urine from rats treated with Tp, these results suggest that the increase bis-NAc^{red} detected in the urine from oxidatively stressed arises from base propenals generated during DNA damage by the stressors.

Discussions

At the end of the *in silico* and *in vitro* analyses in Chapter 2, we were left with 15 metabolite candidates that take into account the predicted processing of GSH conjugates to CG, Cys and mercapturic acid derivatives in liver and kidney. These 15 molecules proved to be predictive of excreted metabolites produced from both exogenous and endogenous base propenals. Of the 15, 3 were found in urine: small quantities of the carboxylic acid derivative of Tp (Tp^{OX}, 5) and larger quantities of the mercapturic acid derivatives 1e and 3e (Figure 3.4), with 3e detected in saline-treated rats as a candidate metabolite of endogenously produced base propenals. In spite of rapid

IV administration of Tp, excretion of 1e and 3e peaked in the second 8-16 h urine collection period (Figure 3.1), suggesting a relatively slow (hours) mercapturic acid metabolism process. This is consistent with the biliary excretion of 1e and 3e (Figure 3.2). However, precursors to the mercapturic acid 1e, the GSH, CG and Cys conjugates 1b, 1c and 1d, respectively, were almost entirely excreted in bile within one hour after Tp injection. The presence of M-GSH^{red} 1b in bile could occur as a result of overwhelming the relatively low level of γ-glutamyltransferase in rat liver [13] by the relatively large Tp dose. However, the absence of other GSH conjugates in bile suggests that γ-glutamyltransferase activity is adequate to completely convert GSH conjugates to CG and Cys metabolites, which is consistent with the observation of these metabolites in bile. These results suggest that the relatively low γ-glutamyltransferase activity in rat liver, relative to the high level in rat kidney [13], is the predominant route for initial metabolism of GSH conjugates of Tp, while the kidney is the major site for slower N-acetylation of the resulting Cys conjugates to form the mercapturic acid products observed in urine. This type of interorgan transfer of GSH conjugate metabolites during mercapturic acid processing is well established [13], though the established hepatic capacity for mercapturic acid formation appears to be minimal for Tp metabolism in rats.

We must now ask if the metabolites lacking the thymidine nucleobase, 3a-e, are possibly derived from metabolism of endogenous molecules other than Tp or other base propenals, such as malondialdehyde or acrolein (a lipid peroxidation product). This is highly unlikely given the absence of the metabolites, except for 3e, in saline-treated rats and their appearance only on rats treated with Tp. Further, 3e increased dramatically upon Tp injection. Even the background production of 3e is unlikely to involve malondialdehyde or acrolein. Malondialdehyde is metabolized mainly to CO_2 [14, 15], while acrolein was mainly excreted as the monoacetylcysteine adduct *in vivo* [16, 17]. Furthermore, addition elimination products (2a-e) are not detected in either urine or bile. This suggests that GSH initially adds to the β-carbon of α, β-unsaturated aldehyde

species faster than the reduction/oxidation of the aldehyde, which suggests that overall Tp metabolic pathway shown in Figure 3.4.

While kidney cannot be ruled out as a source of the oxidized Tp (5), the fact that it is excreted into the bile at levels significantly higher (~440-fold increase) than urine (Tables 3.1A, 3.2A) suggests that the oxidized Tp is formed in the liver and excreted mainly in bile but trace amounts in urine analogous to the appearance of 1e and 3e in bile. The high level of Tp^{ox} in bile is at odds with the general rule of biliary clearance for compounds with molecular weights higher than 300 g/mol. It is therefore highly likely that the negative charge of the carboxylate at physiological pH makes Tp^{ox} a substrate for organic anion transporters or organic anion transporting peptides in bile canaliculis [18].

In spite of the identification of seven metabolites in urine and bile, these potential biomarker candidates accounted for only 25% of the injected dose of Tp, which raises the question about the fate of the bulk of the Tp. Certainly there may be other metabolic pathways such as conjugation with glucuronic acid, sulfate or amino acids. The inability to detect Tp in serum *in vitro* or in blood immediately following IV injection of Tp in rats suggests the possibility of a metabolic activity in blood that intercepts most Tp, a nucleophilic reaction with serum proteins -- though the low reactivity with serum albumin argues that protein reactions must involve sulfur-rich proteins -- or sequestration of Tp by a high-affinity binding site in a protein. The latter would bias our analyses since proteins were removed by ultrafiltration. Ultimately, an understanding of the complete fate of base propenals will rely on tracking with radiolabeled compounds.

A major requirement for biomarker development is the existence of a dose-response relationship between the marker and the pathology, in this case formation of base propenals in response to inflammation and oxidative stress. The observation of bis-NAc^{red} 3 in the urine of saline-treated rats made it the best candidate for a base propenal-derived biomarker of oxidative stress. To induce the oxidative stress, rats were treated with CCl_4, which is well established to cause

acute and chronic tissue injury and inflammation [19] by virtue of hepatic metabolism to trichloromethyl and peroxy trichloromethyl radicals [20, 21]. One piece of evidence that CCl_4 would lead to formation of base propenals involved the observation by Marnett and coworkers that the level of M_1dG increased ~2-fold in rats treated with CCl_4 [22], with strong evidence M_1dG arises from reactions of DNA with base propenals [23]. The 3-fold increase in urinary bis-NAc^{red} 3e rats urine treated with CCl_4 further strengthens the link between M_1dG and base propenals, and, more importantly, establishes base propenal metabolites as biomarkers of inflammation, with strong implications for biomarkers of disease risk.

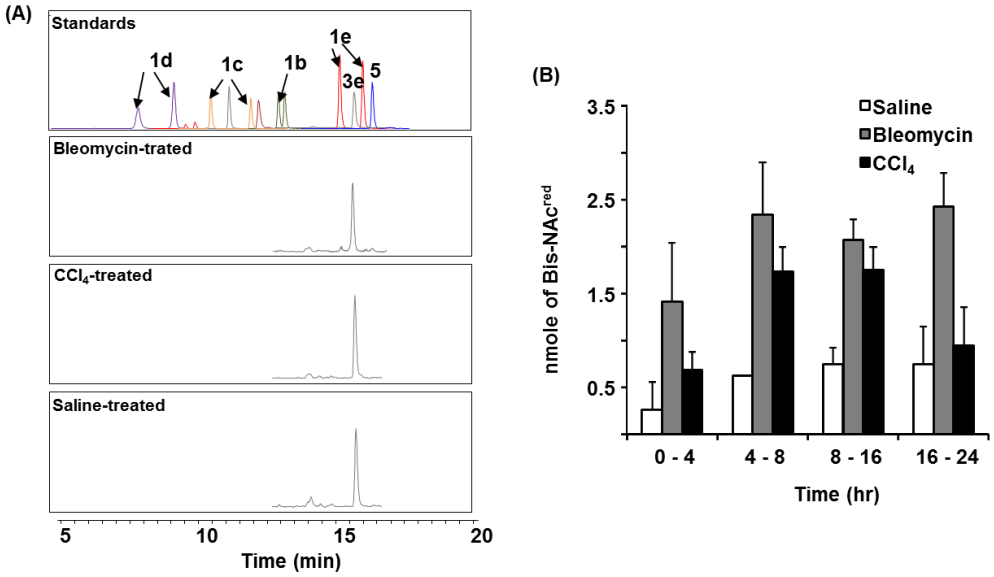


Figure 3.3. Urinary metabolites detected following administration of bleomycin and CCl_4. (A) Bis-NAc^{red} 3e is the only urinary metabolites detected in rats treated with saline (n = 2), bleomycin (n = 2) and CCl_4 (n = 2). (B) The level of Bis-NAc^{red} 3e excreted in urine from three groups of rats: saline-treated (white), bleomycin-treated (grey) and CCl_4-treated (black).

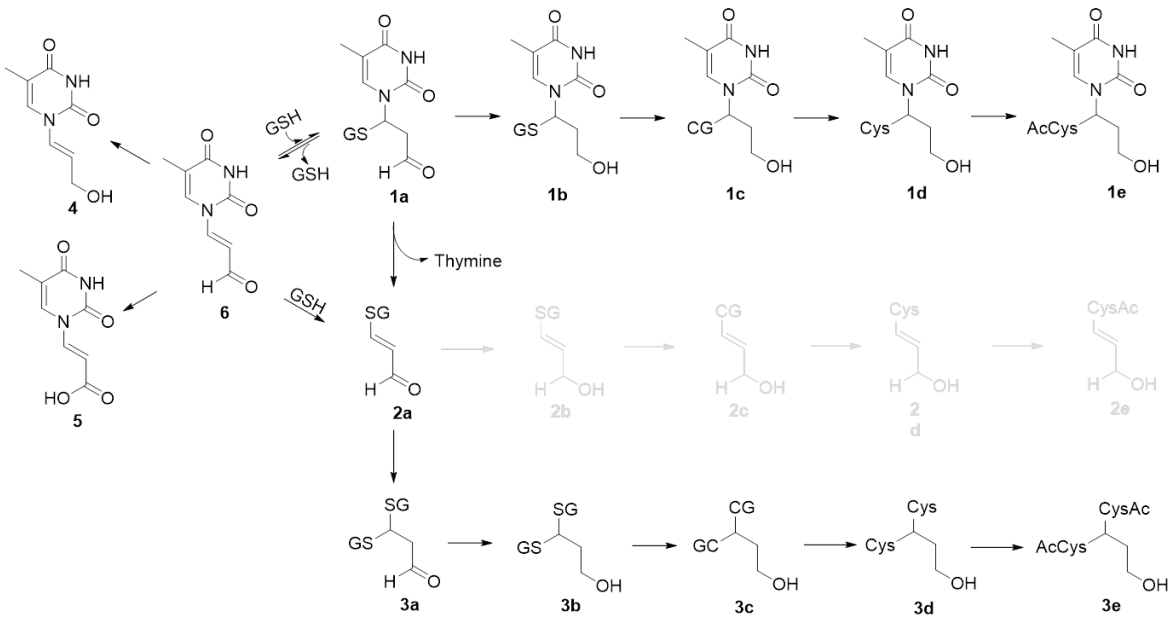


Figure 3.4. Proposed metabolic fates of thymine propenal 6 via GSH conjugations. The addition-elimination product series, highlighted in grey, is not detected in both urine and bile.

References

1. Dedon, P.C., et al., *Challenges in developing DNA and RNA biomarkers of inflammation.* Biomark Med, 2007. **1**(2): p. 293-312.
2. Loeb, L.A. and C.C. Harris, *Advances in chemical carcinogenesis: a historical review and prospective.* Cancer Res, 2008. **68**(17): p. 6863-72.
3. De Bont, R. and N. van Larebeke, *Endogenous DNA damage in humans: a review of quantitative data.* Mutagenesis, 2004. **19**(3): p. 169-85.
4. Epe, B., *Role of endogenous oxidative DNA damage in carcinogenesis: what can we learn from repair-deficient mice?* Biol Chem, 2002. **383**(3-4): p. 467-75.
5. Marnett, L.J., *Oxyradicals and DNA damage.* Carcinogenesis, 2000. **21**(3): p. 361-70.
6. Dedon, P.C. and S.R. Tannenbaum, *Reactive nitrogen species in the chemical biology of inflammation.* Arch Biochem Biophys, 2004. **423**(1): p. 12-22.
7. Marnett, L.J., *Oxy radicals, lipid peroxidation and DNA damage.* Toxicology, 2002. **181-182**: p. 219-22.
8. Olinski, R., et al., *Urinary measurement of 8-OxodG, 8-OxoGua, and 5HMUra: a noninvasive assessment of oxidative damage to DNA.* Antioxid Redox Signal, 2006. **8**(5-6): p. 1011-9.
9. Hoberg, A.M., et al., *Measurement of the malondialdehyde-2'-deoxyguanosine adduct in human urine by immuno-extraction and liquid chromatography/atmospheric pressure chemical ionization tandem mass spectrometry.* J Mass Spectrom, 2004. **39**(1): p. 38-42.
10. Otteneder, M.B., et al., *In vivo oxidative metabolism of a major peroxidation-derived DNA adduct, M1dG.* Proc Natl Acad Sci U S A, 2006. **103**(17): p. 6665-9.
11. Knutson, C.G., et al., *Monitoring in vivo metabolism and elimination of the endogenous DNA adduct, M1dG {3-(2-deoxy-beta-D-erythro-pentofuranosyl)pyrimido[1,2-alpha]purin-10(3H)-one}, by accelerator mass spectrometry.* Chem Res Toxicol, 2008. **21**(6): p. 1290-4.
12. Knutson, C.G., et al., *Metabolism and elimination of the endogenous DNA adduct, 3-(2-deoxy-beta-D-erythropentofuranosyl)-pyrimido[1,2-alpha]purine-10(3H)-one, in the rat.* J Biol Chem, 2007. **282**(50): p. 36257-64.
13. Hinchman, C.A. and N. Ballatori, *Glutathione conjugation and conversion to mercapturic acids can occur as an intrahepatic process.* J Toxicol Environ Health, 1994. **41**(4): p. 387-409.
14. Marnett, L.J., et al., *Distribution and oxidation of malondialdehyde in mice.* Prostaglandins, 1985. **30**(2): p. 241-54.
15. Siu, G.M. and H.H. Draper, *Metabolism of Malonaldehyde Invivo and Invitro.* Lipids, 1982. **17**(5): p. 349-355.
16. Parent, R.A., H.E. Caravello, and D.E. Sharp, *Metabolism and distribution of [2,3-C-14]acrolein in Sprague-Dawley rats.* Journal of Applied Toxicology, 1996. **16**(5): p. 449-457.
17. Parent, R.A., et al., *Metabolism and distribution of [2,3-C-14]acrolein in Sprague-Dawley rats - II. Identification of urinary and fecal metabolites.* Toxicological Sciences, 1998. **43**(2): p. 110-120.
18. International Transporter, C., et al., *Membrane transporters in drug development.* Nat Rev Drug Discov, 2010. **9**(3): p. 215-36.
19. Ogeturk, M., et al., *Caffeic acid phenethyl ester protects kidneys against carbon tetrachloride toxicity in rats.* J Ethnopharmacol, 2005. **97**(2): p. 273-80.
20. Weber, L.W.D., M. Boll, and A. Stampfl, *Hepatotoxicity and mechanism of action of haloalkanes: Carbon tetrachloride as a toxicological model.* Critical Reviews in Toxicology, 2003. **33**(2): p. 105-136.
21. Berhane, K., et al., *Detoxication of base propenals and other alpha, beta-unsaturated aldehyde products of radical reactions and lipid peroxidation by human glutathione transferases.* Proc Natl Acad Sci U S A, 1994. **91**(4): p. 1480-4.

22. Chaudhary, A.K., et al., *Detection of Endogenous Malondialdehyde-Deoxyguanosine Adducts in Human Liver.* Science, 1994. **265**(5178): p. 1580-1582.
23. Dedon, P.C., et al., *Indirect mutagenesis by oxidative DNA damage: Formation of the pyrimidopurinone adduct of deoxyguanosine by base propenal.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(19): p. 11113-11116.

This page is intentionally left blank

Chapter 4

The genetic toxicology of intermediary metabolism: Development of a method to quantify the N^2-carboxyhydroxymethyl 2'-deoxyguanosine adduct derived from the reaction of glyoxylate with DNA

Abstract

Intermediary metabolism produces a variety of electrophiles with the potential for reacting with DNA to form mutagenic adducts, as illustrated by the glyxoylate cycle in plants, bacteria, protists, and fungi. The goal of this chapter was to assess the genetic toxicology of glyoxylate by characterizing its reactivity with dG and DNA, and then developing and validating an analytical method for identifying and quantifying glyoxylate-mediated adduct formation in DNA. In reactions of glyoxylate with dG at pH 7.4, the only product formed was identified by chromatography-coupled mass spectrometry as N^2-carboxyhydroxymethyl 2'-deoxyguanosine (N^2-CHMdG). Characterization of N^2-CHMdG revealed that it is unstable and degrades to release dG with a half-life of 17 min. Reduction of N^2-CHMdG led to the formation of N^2-carboxymethyl-2'-dexyguanosine (N^2-CMdG), which proved to be stable at pH 7.4 for days-weeks. With this stable adduct, an isotope-dilution, chromatography-coupled tandem mass spectrometric method was then developed to quantify N^2-CHMdG as N^2-CMdG in calf thymus DNA treated with glyoxylate. The method proved to be extremely sensitive (1 fmol limit of detection, ~10 fmol limit of quantitation) and revealed a dose-response relationship for the formation of N^2-CHMdG in glyoxylate-treated DNA. This method was then applied in Chapter 3 to study the genetic toxicology of metabolic shifts in mycobacteria.

Introduction

Intermediary metabolism gives rise to many types of electrophilic species, including thioesters, phosphate esters, α,β-unsturated carbonyl compounds, and aldehyde-containing compounds, which are capable of reacting with nucleophilic sites in DNA and proteins. For example, it has recently been shown that acetyl CoA and succinyl CoA can act as electrophiles and undergo nucleophilic acyl substitution with the N_{ε}-amino group of lysine and the resulting modification exerts biological functions in terms of regulation of enzymes in central metabolic pathways [1-3]. Electrophiles arising from metabolism also have the potential to react with DNA to form mutagenic adducts, by analogy to DNA reactions with endogenous electrophiles arising from lipid peroxidation products [4]. As a prelude to studies of mycobacterial pathogenesis, this chapter focuses on the genetic toxicology of glyoxylate, an aldehydic species derived from the glyoxylate cycle found in plants, fungi, bacteria, archaea, protists, and nematodes. In plant seedlings, cells utilize the glyoxylate pathway to convert stored lipids to carbohydrates necessary for plant growth and development [5-7]. The glyoxylate cycle also allows mycobacteria to survive nutrient deprivation by switching to β-oxidation of stored fatty acids, since it conserves two carbon in the form of C2-glyoxylate instead of losing it as two molecules of CO_2 per turn that happens during the TCA cycle [8-10].

Toward to the goal of characterizing the genetic toxicology of glyoxylate in mycobacteria, we explored the reactivity of glyxoylate with DNA in this chapter and developed an analytical method to quantify the resulting N^2-carboxyhydroxymethyl 2'-deoxyguanosine (N^2-CHMdG) adducts. The method validated method proved to be high sensitivity, with a limit of quantification of 10 fmol for the stable derivative of N^2-CHMdG, N^2-CMdG. This method will be applied to mycobacterial metabolism in Chapter 3.

Materials & Methods

Materials. Sodium glyoxylate, ethyl glyoxylate 2'-deoxyguanosine, potassium monohydrogenphosphate, potassium dihydrogenphosphate, Tris-HCl, phenol/chloroform in isoamyl alcohol, bovine serum albumin fraction V, sodium chloride, D-(+)-glucose, acetic acid, sodium acetate, formic acid, calf thymus DNA, DNase I, benzonase, calf intestine phosphatase, were purchased from Sigma-Aldrich (St. Louis, MO). Phosphodiesterase I was purchased from Affymetrix (Cleveland, Ohio). [^{15}N(U)]-labeled 2'-deoxyguanosine was purchased from Cambridge Isotopes (Andover, MA). Isopropanol, acetonitrile, hygromycin and 7H9 Middlebrook media were purchased from VWR Scientific (Franklin, MA). Deionized water was further filtered through MilliQ systems (Millipore corporation, Bedford, MA) and used in the whole experiment.

Instruments. UV–vis measurements were made on an HP8452 diode-array spectrophotometer (Agilent Technologies). HPLC analyses were carried out on an Agilent 1100 HPLC system with binary pumps, a degasser, and an autoinjector. Quantitative LC-MS/MS analyses were conducted on an Agilent 1290 HPLC pump system interfaced with an Agilent 6430 triple quadrupole mass spectrometer. Electrospray ionization (ESI)-MS experiments for isotopic purity and structural confirmation were performed on anAgilent MSD-TOF mass spectrometer.

HPLC. There are two HPLC methods.

HPLC method A. The mobile phases are 5 mM ammonium acetate in water (solvent A) and acetonitrile (solvent B). The gradient of the solvents is programmed as follows: 0% B for 5 min; increase to 20% B for 25 min; increase to 100% B for 0.1 min and hold at 100% B for 10 min; decrease to 0% B for 1 min and re-equilibrate the column for 15 min for the next run.

HPLC method B. The mobile phases are 0.1% formic acid in water (solvent A) and acetonitrile (solvent B). The gradient of the solvents is programmed as follows: 0%B for 5 min;

increase to 20% B for 25 min; increase to 100% B for 0.1 min and hold at 100% B for 10 min; decrease to 0% B for 1 min and re-equilibrate the column for 15 min for the next run.

Synthesis of N^2-carboxymethyl 2-deoxyguanosine (N^2-CMdG) and the internal standard. 2'-Deoxyguanosine (0.267 g, 1 mmol) was reacted with 100 mmol of ethyl glyoxylate in 1 M NaOAc/HOAc buffer pH 5.5 and the mixture was stirred for 15 min at ambient temperature. Then, 25 μL of H_2SO_4 was added to the stirred mixture and the suspension was stirred for 30 min. Another 50 μL of H_2SO_4 was added and the suspension would turn into solution which was further stirred for 1 hr. The Schiff base intermediate was purified by HPLC method A. The fractions from HPLC were combined and evaporated *in vacuo*. The Schiff base was then reconstituted in 50% EtOH/H_2O and 0.380 mg (10 mmol) of sodium borohydride ($NaBH_4$) was added to reduce the Schiff base to amine. The solution was further stirred for 2 hr and then neutralized by 50% HOAc until no H_2 gas was developed. The crude of the final product N^2-CHMdG was purified with HPLC method A with a Hypersil GOLD C18 column (3.0 × 250 mm, 5 μm, Thermo Scientific). UV (λ_{max}, pH 7) = 260 nm. $[M+H]^+_{obs}$ = 326.1103; $[M+H]^+_{calc}$ = 326.1111; error = 2.39 ppm. MS/MS (ESI+): 210 (nucleobase + H), 164 (nucleobase – CO_2 + H), 117 (oxonium ion of 2'-dR). ^{1}H-NMR (D_2O): δ2.53 (1H at C2', m), δ2.22 (1H at C2', m), δ3.55 (1H at C4', t), δ3.98 (2H methylene, s,NH–C<u>*H*₂</u>–COOH), δ4.38 (1H at C3', s), δ6.15 (1H at C1', dd), δ6.50 (1H at N^2-exocyclic NH_2, t), δ7.95 (1H at C8 guanine, s).

Reaction of ethyl glyoxylate with 2'-deoxyguanosine. 2'-Deoxyguanosine (1 mM final concentration) was reacted with 1 mM (final concentration) of ethyl glyoxylate in 50 mM phosphate buffer, pH 7.4 for 4 h at 37 °C. The reaction mixture was then analyzed by LC-MS using Agilent 1100 system interfaced with Agilent 1100 series LC/MSD on a Dionex Acclaimed Polar Advantage C18 column (2.1 × 150 mm, 3 μm particle size, Thermo Scientific) with the gradient in HPLC method B.

Reaction of glyoxylate with 2'-deoxyguanosine. 2'-Deoxyguanosine (1 mM final concentration) was reacted with 1 mM (final concentration) of sodium glyoxylate in 50 mM phosphate buffer, pH 7.4 for 4 h at 37 °C. The reaction mixture was then analyzed by LC-MS using

Agilent 1100 system interfaced with Agilent 1100 series LC/MSD on a Dionex Acclaimed Polar Advantage C18 column (2.1 × 150 mm, 3 μm particle size, Thermo Scientific) with the gradient in HPLC method B.

Determination of N^2-CHMdG stability. N^2-CHMdG was purified from the reaction of 1 mM 2'-deoxyguanosine with 100 mM sodium glyoxylate using HPLC method B at 4 °C. The purified fraction of N^2-CHMdG was incubated with 50 mM phosphate buffer, pH 7.4 at 37 °C. Aliquots of the reaction mixture were sampled at 15, 30, 45, 60 and 120 min, and N^2-CHMdG and dG were quantified by HPLC at λ_{max} = 260 nm. The extinction coefficients (ε_{260}) for N^2-CHMdG and dG are 10.1 $mM^{-1}cm^{-1}$ [11] and 11.7 $mM^{-1}cm^{-1}$ [12] , respectively.

Determination of N^2-CHMdG stability in DNA. Calf thymus DNA (100 g) was reacted with 5 mM glyoxylate at 37 °C for 18 h. The DNA was then precipitated with 3 M NaOAc/HOAc buffer pH 5.5 followed by an equal volume of isopropanol. The pellet was washed with 75% EtOH/H_2O twice, dried *in vacuo* and dissolved in phosphate buffer pH 7.4. After 0.5, 1 and 2 h at 37 °C, samples were removed and reduced with 2 mM $NaCNBH_3$ at 37 °C for 16 h. The mixtures were acidified with 3 M NaOAc/HOAc buffer pH 5.5, and DNA was precipitated with isopropanol. The DNA pellets were washed with 75% EtOH/H_2O twice and dried *in vacuo*. The resulting DNA was reconstituted in 10 mM Tris-HCl buffer (pH 7.9) with 1 mM $MgCl_2$ and digested with 10 U benzonase, 5 U DNase I, 17 U calf intestine phosphatase, 0.1 U phosphodiesterase I, and 10 pmol of [^{15}N(U)]-labeled CMdG. The digestion reaction was incubated for 12 h at 37 °C. Following removal of proteins with a 10 kD exclusion, the collected filtrates were evaporated and reconstituted in water. The solution was purified by HPLC method A to collect the N^2-CMdG fraction. The collected fractions were evaporated, reconstituted in water and analyzed by LC-MS/MS.

Reaction of glyoxylate with calf thymus DNA. Calf thymus DNA (100 μg) was reacted with glyoxylate at concentrations of 0.5, 1, 2 and 4 mM at 37 °C for 4 h. The reaction was stopped and products stabilized by adding 2 mM purified $NaCNBH_3$ and further incubating at 37 °C for 16 h.

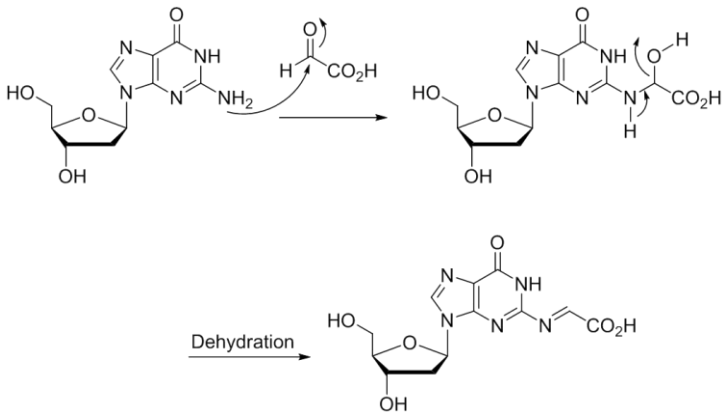
The mixtures were acidified with 3 M NaOAc/HOAc buffer pH 5.5, and DNA was precipitated with isopropanol. The DNA pellets were washed with 75% EtOH/H_2O twice and dried *in vacuo*. The resulting DNA was reconstituted in 10 mM Tris-HCl buffer pH 7.9, 1 mM $MgCl_2$ and digested with 10 U benzonase, 5 U DNase I, 17 U calf intestine phosphatase, 0.1 U phosphodiesterase I, and 10 pmol of [^{15}N(U)]-labeled CMdG. The digestion reaction was incubated for 12 h at 37 °C. The reaction was then worked up using 10 kD exclusion filter to remove proteins. The collected filtrates were evaporated and reconstituted in water. The solution was purified by HPLC method A to collect N^2-CMdG fraction. The collected fractions were evaporated, reconstituted in water and analyzed by LC-MS/MS.

Results

Formation of a Schiff's base adduct in reactions of dG with ethyl glyoxylate but not with glyxoylate. We had predicted that glyoxylate would react with the N^2-position of dG to form an initial carbinolamine by nucleophilic attack of the N^2-exocyclic amine of dG at the aldehyde of glyxoylate to form the N^2-CHMdG product shown in Scheme 4.1, followed by dehydration to form a stable Schiff's base. To test this prediction, sodium glyoxylate (1 mM) was reacted with dG (1 mM) in 50 mM phosphate buffer (pH 7.4) at 37 °C and products analyzed by LC-MS/MS. The reaction generated only the adduct later characterized to be N^2-CHMdG (Figure 4.1). Unexpectedly, the Schiff's base did not form. We argue that carboxylate anion forms an intramolecular hydrogen bond with the hydroxyl group of the adduct, which prevents the dehydration of the hydroxyl group, as shown in Scheme 4.2. The failure to dehydrate N^2-CHMdG to form the Schiff's base was proposed to involve stabilization of the carbinolamine to prevent its dehydration. To test this hypothesis, dG was reacted with the ethyl ester of the glyoxylate carboxylic acid, ethyl glyoxylate, and products again analyzed by LC-MS/MS. As shown in Figure 4.2, the Schiff's base was generated in abundance in this reaction.

The structure of N^2-CHMdG was confirmed by high-resolution mass spectrometry with collision-induced dissociation. The MS^2 fragmentation pattern is shown in Figure 1B. The difference in the observed and calculated masses of N^2-CHMdG is 2.49 ppm, which falls into the acceptable range of error where the mass difference should be lower than 5 ppm. The CID spectrum shows signals at *m/z* 226 and 208, which correspond to N^2-carboxyhydroxymethylguanine (N^2-CHMGua) and the Schiff's base of N^2-CHMGua derived from the loss of water from N^2-CHMGua. Due to the instability of N^2-CHMdG (discussed below), the structure of N^2-CHMdG could not be directly confirmed by ^{1}H-NMR without stabilizing by reduction with $NaCNBH_3$ to form N^2-CMdG (Figure 4.3A, 4.3B). ^{1}H-NMR confirms the N^2-adduct product structure by showing the chemical shift of

methylene group attached to N^2-position at δ3.98 as described in Materials & Methods. Furthermore, integration of the proton on the exocyclic amine indicates only one proton and shows an expected triplet peak at δ6.50, indicating that the exocyclic amine proton is located adjacent to the two protons of the methylene group, which confirms that glyoxylate modifies the N^2-position of dG (see Materials & Methods in this chapter).



Scheme 4.1

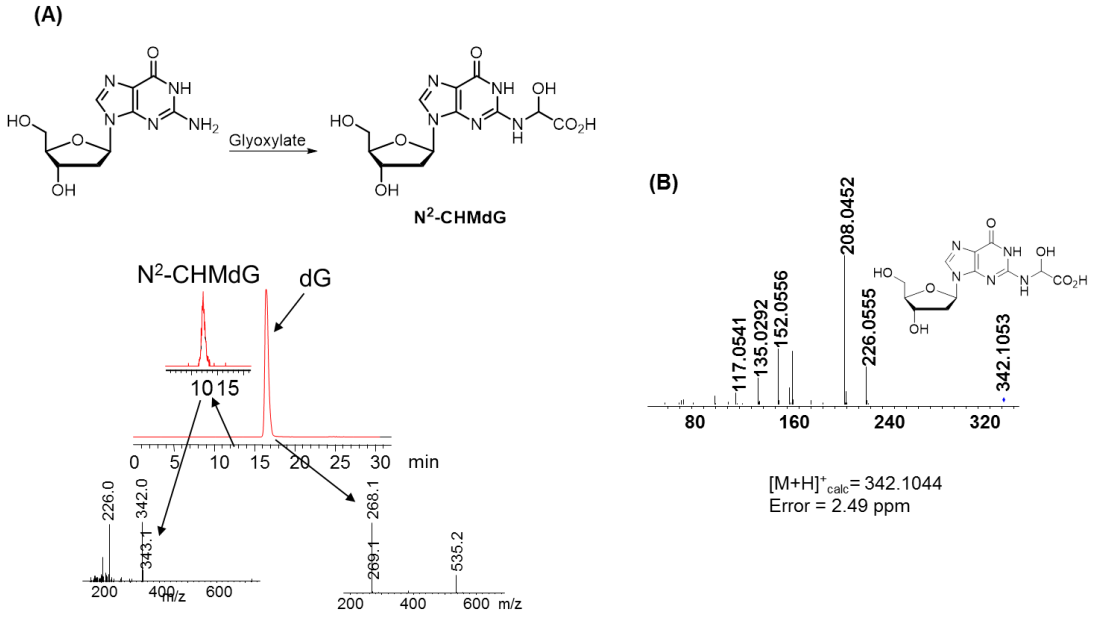


Figure 4.1. Generation of the N²-CHMdG adduct from dG and glyoxylate (A) and structural characterization by high-resolution MS (HR-MS) with 2.40 ppm error and MS^2 fragmentation revealing the fragmentation of the nucleobase from the adduct, N²-CHMGua (m/z = 226). Please note that m/z = 535 is corresponding to the dG dimer.

Instability of the N²-CHMdG adduct. The stability of N²-CHMdG nucleoside adduct was assessed following purification of N²-CHMdG from the reaction of glyoxylate with dG using HPLC. The purified adduct was analyzed by HPLC at λ_{max} = 260 nm and the peak area was used to determine the relative abundance of N²-CHMdG and dG (Figure 4.3A). The half-life ($t_{1/2}$) of N²-CHMdG at 37 °C in phosphate buffer (pH 7.4) was 17 min, with N²-CHMdG undergoing degradation to release dG as shown in Figure 4.3A.

Stabilization of N²-CHMdG to N²-CMdG nucleoside. Since we need to quantify the level of the N²-CHMdG adduct as a nucleoside from DNA, N²-CHMdG nucleoside needs to be stabilized before LC-MS/MS analysis. Sodium cyanoborohydride ($NaCNBH_3$) has been widely used to stabilize Schiff's bases and carbionalamines by transforming them to their corresponding stable amines. The

goal here was to reduce N^2-CHMdG to N^2-CMdG. Initial attempts at reducing N^2-CHMdG were hampered by contaminating CN^-, a product of $NaCNBH_3$ degradation in commercial preparations. CN- competes with hydride ion (H^-) reduction to the amine by forming a cyanohydrin species that prevents reduction [13-15]. Without purification, the yield of N^2-CMdG was only 2%. By purifying $NaCNBH_3$ by crystallization, we were able to quantitatively convert N^2-CHMdG to N^2-CMdG (Figure 4.3B) with 78% yield. The resulting N^2-CMdG adduct proved to be stable for more than 24 h and likely for days to weeks.

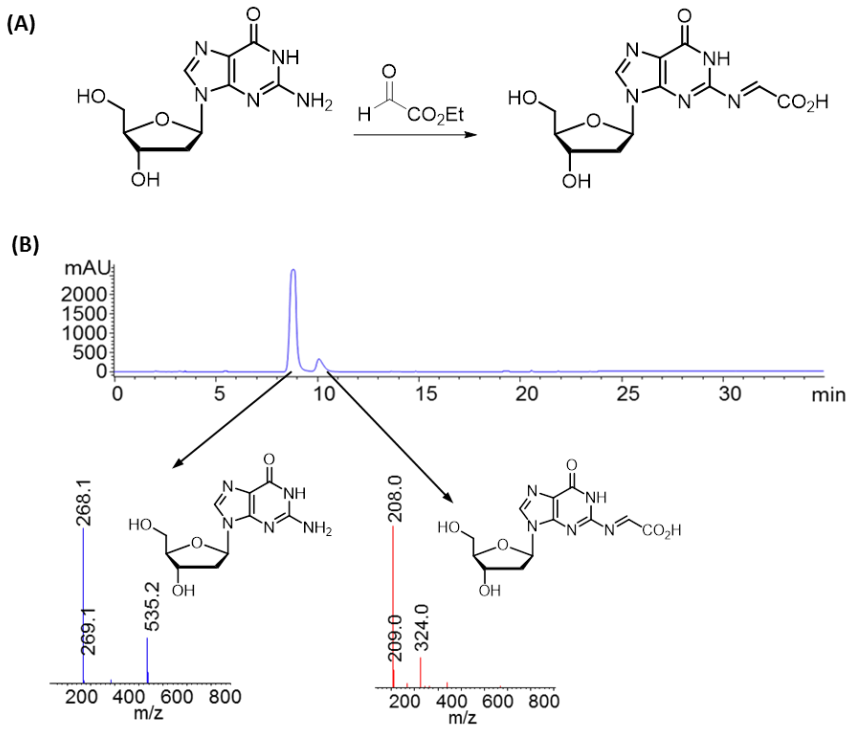
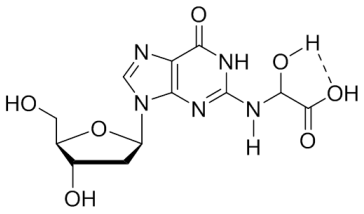


Figure 4.2. Generation of the Schiff's base adduct from dG and glyoxylate (A) and structural characterization by LC-MS (B).

Stability of N²-CHMdG in DNA. The instability of N²-CHMdG in the nucleoside form raised concerns about the stability of the adduct in double-stranded DNA. The experimental conditions for assessing the stability of N²-CHMdG was described in Materials & Methods. As shown in Figure 4.4, the level of the adduct decreases to the background level within 15 min.



Scheme 4.2. Proposed five-membered intramolecular hydrogen bonding in N²-CHMdG, resulting in preventing N²-CHMdG from dehydration to Schiff's base.

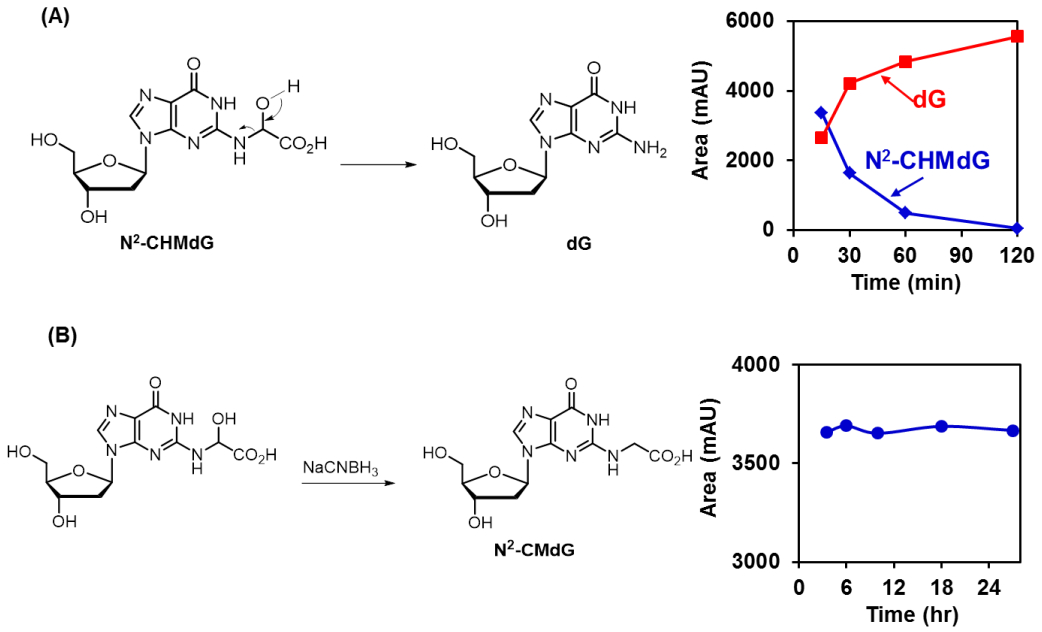


Figure 4.3. N²-CHMdG is unstable and transforms to dG (A). Stabilization of N²-CHMdG to N²-CMdG using $NaCNBH_3$

Development of an LC-MS/MS method for quantifying N^2-CMdG as a stabilized form of N^2-CHMdG in calf thymus DNA. The instability of N^2-CHMdG both as a nucleoside and in DNA necessitated the development of an analytical method based on the more stable reduced form of the adduct, N^2-CMdG. To this end, calf thymus DNA was reacted with different concentrations of sodium glyoxylate and then treated with $NaCNBH_3$ as described in Materials & Methods to convert unstable N^2-CHMdG to stable N^2-CMdG. After hydrolysis to 2'-deoxynucleosides with nucleases and alkaline phosphatase, the putative reduction product N^2-CMdG was purified by HPLC and identified and quantified by isotope-dilution LC-MS/MS in positive-ESI multiple reaction monitoring mode (MRM) with the transition of the *m/z* value 326 to 210 and 331 to 215 for the N^2-CMdG in the DNA sample and [^{15}N(U)]-labeled N^2-CMdG in the internal standard, respectively (Figure 4.5A). Secondary MRM transitions are also used -- 326 to 164 for N^2-CMdG and 331 to 169 for [^{15}N(U)]-labeled N^2-CMdG internal standard -- to ensure that the identified peaks were not the product from interference in the matrix. The detection limit for N^2-CMdG in this assay is ~1 fmol. In addition, the reaction of calf thymus DNA and sodium glyoxylate exhibits a strong dose-response behavior (Figure 4.5B), which suggests that glyoxylate reacts not only with dG nucleoside and exocyclic amine of dG in DNA but also modifies DNA in an enzyme-independent fashion.

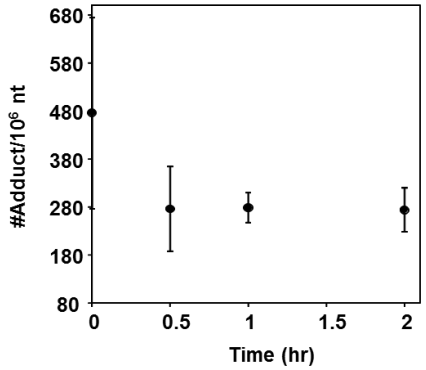


Figure 4.4. Stability of N^2-CHMdG in calf thymus DNA.

Discussion

The goal of Chapter 4 in this thesis is to initially test the hypothesis that aldehyde-containing glyoxylate can act as an electrophile by reacting with biological nucleophiles, in this case dG and subsequently develop LC-MS/MS methods for analysis of this adduct in *M. smegmatis* (Chapter 5). We started by reacting glyoxylate with dG but to our surprise, the expected Schiff's base was not generated. Instead, N^2-CHMdG is the product (Figure 4.1A). The prevention from hydration of N2-CHMdG is proposed by the intramolecular hydrogen bonding between carboxylate and hydrogen in the hydroxyl group (Scheme 4.2). This argument is confirmed by the reaction of ethyl glyoxylate (carboxylic acid is derivatized to an ethyl ester) with dG and the Schiff's base is the predominant product (Figure 4.2).

N^2-CHMdG is unstable in both nucleoside form and in DNA. As to the nucleoside form, general bases such as phosphate group can generate dehydration of N^2-CHMdG, subsequently transforming it to dG. However, the instability of N^2-CHMdG in DNA was unexpected, since Swenberg *et al.* observed that an acetaldehyde-derived N^2-dG adduct was stable in DNA [16]. We speculate that anion from carboxylate group forms intramolecular hydrogen bond with hydrogen in hydroxyl group, which in turn makes the hydroxyl hydrogen more acidic and deprotonated easier at pH 7.4 in the phosphate buffer.

Similar to Chapter 2 in terms of unavailability of commercial N^2-CMdG, we need to synthesize this compound in our lab. The reaction for the synthesis of N^2-CMdG is not performed via the pathway of glyoxylate and dG but instead ethyl glyoxylate with dG. It is very obvious that the reactivity of ethyl glyoxylate towards dG is much more than glyoxylate since the carboxylate anion is esterified, which in turn makes the adjacent aldehydic carbonyl of ethyl glyoxylate more electrophilic. In this way, the reaction of ethyl glyoxylate with dG provides us with much more percent yield. The Schiff's base intermediate is subject to reduction by $NaBH_4$. To this end, a higher

amount of N^2-CMdG standards could be obtained for use of LC-MS/MS optimization and structural confirmation by NMR spectroscopy.

Obtaining the N^2-CMdG standard, an MRM method was developed through MRM mode. MRM provides the best sensitivity for the analysis. This ultra-sensitive method is required for DNA adduct analysis due to the unpredictable amount of the adduct in DNA and generally low abundance of the adduct [17]. In our method, the limit detection of N^2-CMdG is ~1 fmol. Therefore, in this chapter, we showed that glyoxylate can react with dG and form N^2-CHMdG adduct. Stabilization of N2-CHMdG was successfully executed by reduction of purified $NaCNBH_3$ to generate N2-CMdG. To this end, the synthesized N^2-CMdG was used for LC-MS/MS method development to gain the best sensitivity (~1 fmol) for further analysis of the adduct in *M. smegmatis* analysis that is described in the next chapter.

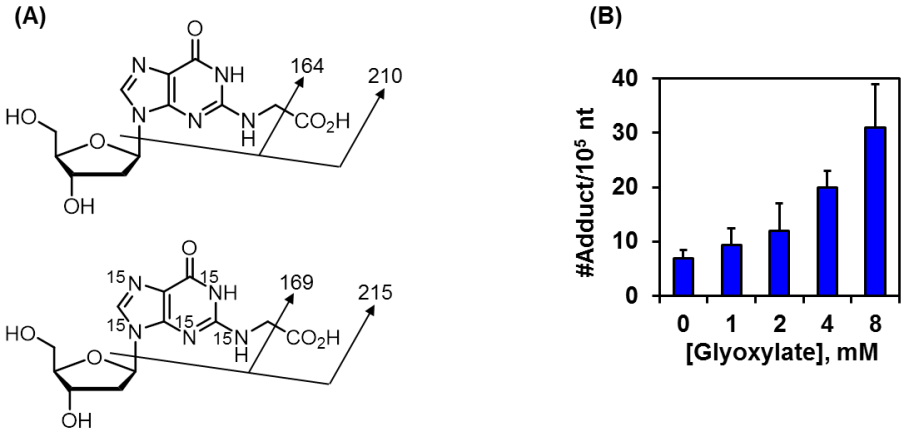


Figure 4.5. Two MRM transitions are used for single compounds to prevent interfering signals: 326→210 and 326 →164 for N^2-CMdG and 331→215 and 331→169 for [^{15}N(U)]-labeled N^2-CMdG (A). Dose-response of the N^2-CMdG level from calf thymus DNA according to glyoxylate concentrations (B)

References

1. Moellering, R.E. and B.F. Cravatt, *Functional Lysine Modification by an Intrinsically Reactive Primary Glycolytic Metabolite.* Science, 2013. **341**(6145): p. 549-553.
2. Park, J., et al., *SIRT5-mediated lysine desuccinylation impacts diverse metabolic pathways.* Mol Cell, 2013. **50**(6): p. 919-30.
3. Weinert, B.T., et al., *Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation.* Cell Rep, 2013. **4**(4): p. 842-51.
4. Blair, I.A., *DNA adducts with lipid peroxidation products.* J Biol Chem, 2008. **283**(23): p. 15545-9.
5. Kornberg, H.L. and H. Beevers, *The Glyoxylate Cycle as a Stage in the Conversion of Fat to Carbohydrate in Castor Beans.* Biochimica Et Biophysica Acta, 1957. **26**(3): p. 531-537.
6. Kornberg, H.L. and H. Beevers, *A mechanism of conversion of fat to carbohydrate in castor beans.* Nature, 1957. **180**(4575): p. 35-6.
7. Eastmond, P.J., et al., *Postgerminative growth and lipid catabolism in oilseeds lacking the glyoxylate cycle.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5669-74.
8. Berg, J.M., J.L. Tymoczko, and L. Stryer, *Biochemistry*. 7th ed. 2012, New York: W.H. Freeman. xxxii, 1054, 43, 41, 48 p.
9. Murthy, P.S., M. Sirsi, and T. Ramakrishnan, *Effect of age on the enzymes of tricarboxylic acid and related cycles in Mycobacterium tuberculosis H37Rv.* Am Rev Respir Dis, 1973. **108**(3): p. 689-90.
10. Wayne, L.G. and K.Y. Lin, *Glyoxylate metabolism and adaptation of Mycobacterium tuberculosis to survival under anaerobic conditions.* Infect Immun, 1982. **37**(3): p. 1042-9.
11. Wang, H., H. Cao, and Y. Wang, *Quantification of N2-carboxymethyl-2'-deoxyguanosine in calf thymus DNA and cultured human kidney epithelial cells by capillary high-performance liquid chromatography-tandem mass spectrometry coupled with stable isotope dilution method.* Chem Res Toxicol, 2010. **23**(1): p. 74-81.
12. Andrus, A. and R.G. Kuimelis, *Base composition analysis of nucleosides using HPLC.* Curr Protoc Nucleic Acid Chem, 2001. **Chapter 10**: p. Unit 10 6.
13. Jentoft, N. and D.G. Dearborn, *Labeling of proteins by reductive methylation using sodium cyanoborohydride.* J Biol Chem, 1979. **254**(11): p. 4359-65.
14. Jentoft, N. and D.G. Dearborn, *Protein labeling by reductive methylation with sodium cyanoborohydride: effect of cyanide and metal ions on the reaction.* Anal Biochem, 1980. **106**(1): p. 186-90.
15. Heacock, C.S., et al., *In vitro labeling of proteins by reductive methylation: application to proteins involved in supramolecular structures.* J Cell Biochem, 1982. **19**(1): p. 77-91.
16. Moeller, B.C., et al., *Biomarkers of exposure and effect in human lymphoblastoid TK6 cells following [13C2]-acetaldehyde exposure.* Toxicol Sci, 2013. **133**(1): p. 1-12.
17. Dedon, P.C., et al., *Challenges in developing DNA and RNA biomarkers of inflammation.* Biomark Med, 2007. **1**(2): p. 293-312.

Chapter 5

The genetic toxicology of intermediary metabolism: Analysis of N^2-carboxyhydroxymethyl 2'-deoxyguanosine during metabolic shifts in *Mycobacterium smegmatis*

Abstract

In the previous chapter, the concept of genetically toxic electrophiles from intermediary metabolism was illustrated with glyxoylate, which plays a central role in the alternative to the tricarboxylic acid cycle in plants, bacteria, protists, and fungi. It was demonstrated that glyoxylate reacts with dG in DNA to form N^2-carboxyhydroxymethyl 2'-deoxyguanosine (N^2-CHMdG), an unstable adduct for which a reduction-based analytical method was developed. In this chapter, the method is applied to explore the formation of N^2-CHMdG in mycobacteria subjected to nutrient deprivation that induces a shift to the glyoxylate cycle. The motivation for these studies lies in the higher-than-expected mutation rate in mycobacteria induced to enter the non-replicative state during nutrient deprivation or hypoxia. To test this model, adduct formation was compared in (1) wild type *M. smegmatis* (Msmeg) grown in rich or acetate media, the latter inducing a switch from the TCA cycle to the glyoxylate cycle; and (2) the isocitrate dehydrogenase (ICD)-deficient mutant of Msmeg. Growth of Msmeg in acetate medium caused a 2-fold increase in N^2-CHMdG compared to growth in rich medium (Middlebrook 7H9). Similarly, N^2-CHMdG increased in 2-fold in the ICD mutant compared to wild-type Msmeg grown in 7H9. The results support the idea that shifts in intermediary metabolism can lead to DNA damage that may cause mutations associated with nutrient deprivation in mycobacteria, with implications for the genetic toxicology of other metabolism-derived electrophiles. At the end of this chapter, the level of glyoxylate in Msmeg by cytosol was quantified using aldehyde reactive probe (ARP) in the four types of samples mentioned above. Unexpectedly, the abundances of glyoxylate were not significantly different despite in an ICD-deficient strain. This might be due to the steady-state of glyoxylate governed by the binding of electrophilic site of glyoxylate to proteins or amino acids. Moreover, other fates of glyoxylate in terms of conversion to glycine and aminolevulinate can serve as an alternative explanation for the unchanged level of the compound. Eventually, to test whether glyoxylate can cause mutation, blue-white assay was performed based on the hypothesis that glyoxylate could mutate *LacZ* gene in the pUC19 plasmid. After treatment of various concentrations of glyoxylate with pUC19 plasmid DNA, only blue colonies were observed regardless of the glyoxylate concentration. It is still inconclusive that glyoxylate cannot cause mutation since the method might not be sensitive enough for showing the *LacZ*-mutated white colony caused by glyoxylate.

Introduction

Intermediary metabolism in all organisms gives rise to many types of potentially genotoxic electrophilic species, including thioesters, phosphate esters, α,β-unsturated carbonyl compounds, and aldehyde-containing compounds, which are capable of reacting with nucleophilic sites in DNA and proteins. In the Chapter 4, we explored glyxoylate, the central metabolite in the glyoxylate cycle that serves as an alternative to the tricarboxylic acid cycle in plants, bacteria, protists, and fungi. The glyoxylate cycle is an anapleurotic pathway to the TCA cycle, which means that glyoxylate can replenish one of the intermediates in the TCA cycle. In this case, glyoxylate combines with acetyl CoA to generate malate and yield the total two molecules of oxaloacetate per turn, instead of losing two carbon atoms as CO_2. It was demonstrated in Chapter 4 that glyoxylate is capable of reacting with dG in DNA to form N^2-carboxyhydroxymethyl 2'-deoxyguanosine (N^2-CHMdG), an unstable adduct for which a reduction-based analytical method was developed. Here we apply the method to explore the formation of N^2-CHMdG in mycobacteria subjected to nutrient deprivation that induces a shift to the glyoxylate cycle.

The motivation for these studies lies in the higher than expected mutation rate in mycobacteria induced to enter the non-replicative state of dormancy induced by nutrient deprivation or hypoxia. More specifically, we test the hypothesis that shifting mycobacteria from the TCA cycle to the glyoxylate cycle (Figure 1.8, Chapter 1), either by nutrient deprivation or mutation, leads to elevated levels of potentially mutagenic DNA adducts, including N^2-CHMdG. Nutrient deprivation, hypoxia and nitric oxide, among other environmental changes during mycobacterial infection, all activate the Dos regulon that eventually causes a shift from TCA cycle-based metabolism to the glyoxylate cycle as the bacterium begins to utilize endogenous fatty acids as a fuel source [1]. The glyoxylate cycle plays a crucial role in mycobacterial dormancy and virulence, as indicated by the loss of virulence in murine pulmonary infections of *M. tuberculosis*

lacking the gene for isocitrate lyase (ICL) (Figure 1.8, Chapter 1) [2] and high levels of ICL gene expression dormant mycobacteria [3, 4].

Coupled with the shift to the glyxoylate cycle in dormant mycobacteria, the non-replicative bacterium also experiences a higher than expected mutation rate. Recently, it has been shown that the mutation rate of dormant *M. tuberculosis*, where it was expected to be very low due to the shift of bacteria to the non-replicating state, is as high as that of actively replicating *M. tuberculosis* [5]. This mutation rate may play a role in clinical drug resistance since it is believed that a significant portion of drug resistance in *M. tuberculosis* results from mutations during infection [5]. Here we test the hypothesis that potentially mutagenic glyxoylate adducts in DNA correlate with the shift to glyxoylate cycle in non-replicative mycobacteria.

For these studies, we employ *M. smegmatis* (Msmeg) as a BSL2 surrogate for BSL3-level *M. tuberculosis*. Msmeg is generally used for a model organism for *M. tuberculosis* due to its fast growth rate (generation time 3 hr) and very high similarity both in terms of genome and dormancy phenotype, with Msmeg sharing more than 2000 homologs with *M. tuberculosis* and the same unusual cell wall structure of *M. tuberculosis* and other mycobacterial species [6]. To test the DNA adduct hypothesis, we induced the glyoxylate cycle in Msmeg in two ways: using minimal media (M9) supplemented with acetate in order to reprogram the metabolism from the TCA cycle to the glyoxylate shunt [7], and using an ICD-deficient mutant strain grown in the 7H9 media. Both conditions led to an increase in N^2-CHMdG in support of the potential genetic toxicology of intermediary metabolism in mycobacteria.

Materials & Methods

Materials. Sodium glyoxylate, ethyl glyoxylate 2'-deoxyguanosine, potassium monohydrogenphosphate, potassium dihydrogenphosphate, Tris-HCl, phenol/chloroform in isoamyl alcohol, bovine serum albumin fraction V, sodium chloride, D-(+)-glucose, acetic acid, sodium acetate, formic acid, calf thymus DNA, DNase I, benzonase, calf intestine phosphatase, were purchased from Sigma-Aldrich (St. Louis, MO). Phosphodiesterase I was purchased from Affymetrix (Cleveland, Ohio). [^{15}N(U)]-labeled 2'-deoxyguanosine was purchased from Cambridge Isotopes (Andover, MA). Isopropanol, acetonitrile, hygromycin and 7H9 Middlebrook media were purchased from VWR Scientific (Franklin, MA). Deionized water was further filtered through MilliQ systems (Millipore corporation, Bedford, MA) and used in the whole experiment.

Instruments. UV–vis measurements were made on an HP8452 diode-array spectrophotometer (Agilent Technologies). HPLC analyses were carried out on an Agilent 1100 HPLC system with binary pumps, a degasser, and an autoinjector. Quantitative LC-MS/MS analyses were conducted on an Agilent 1200 HPLC pump system interfaced with an Agilent 6430 ion trap mass spectrometer.

***M. smegmatis* culture and genomic DNA isolation.** Wild type Msmeg was cultured in two media: Middlebrook 7H9 media supplemented with albumin-dextrose saline (ADS) and 0.5% Tween-80 and an M9 medium supplemented with 2.5 mM sodium acetate. The isocitrate dehydrogenase mutant (ΔICD) of Msmeg was cultured in Middlebrook 7H9 media supplemented with ADS and 0.5% Tween-80. The ICD knock-out mutant of *M. smegmatis* was given as a gift by Richa Gawande from Harvard School of Public Health (S. Fortune lab). Cells were cultured in 50 mL sterile culture tubes and maintained in RPM 250 at 37 °C to mid-log phase (OD ≈ 0.6). DNA was then extracted under carefully controlled conditions to minimize loss of the adducts. Chilled (4 °C) phenol/chloroform in isoamyl alcohol acidified with acetic acid (pH 5) was added to the cell pellets

followed by bead-beating at speed 5.3x (FastPrep, FP120) for 1 min. Purified $NaCNBH_3$ was then added to the mixture to a final concentration of 4 mM to convert the unstable N^2-CHMdG to the stable N^2-CMdG adduct, and the mixture was allowed to shake 16 h at 4 °C in a cold room. The purification of sodium cyanoborohydride was performed as mentioned in Chapter 4 [8]. A volume of 500 □L of Tris-HCl buffer (10 mM, pH 8.0) was added and, following vortexing and centrifugation at 15,000xg, the aqueous upper phase was removed. The Tris-HCl buffer extraction was repeated again and NaOAc/HOAc buffer pH 5.5 was added to the combined upper aqueous phases to a final concentration of 100 mM. 2x volume of isopropanol was subsequently added to precipitate the DNA, with incubation at –20 °C for 2 h. The DNA pellets were washed twice with 75% ethanol and then dried under vacuum.

Enzymatic hydrolysis of reduced genomic DNA. Reduced genomic DNA from Msmeg was reconstituted in 10 mM Tris-HCl buffer, pH 7.9, with 1 mM $MgCl_2$ and digested for 12 h at 37 °C with 10U benzonase, 5U DNase I, 17U calf intestine phosphatase, and 0.1U phosphodiesterase I, with 10 pmol of [^{15}N(U)]-labeled N^2-CMdG added immediately after the enzymes. The reaction was then worked up using a 10 kD exclusion filter to remove proteins. The collected filtrates were evaporated and reconstituted in 100 □L of water. The solution was purified by HPLC method A to collect the N^2-CMdG fraction. The collected fractions were evaporated, reconstituted in water and analyzed by LC-MS/MS.

Analysis of N^2-CMdG by LC-MS/MS. An aliquot of the sample was analyzed by HPLC-ESI-MS/MS using an Agilent 1290 series HPLC system interfaced with an Agilent 6430 triple quadrupole mass spectromer. Samples were resolved on a Dionex Acclaimed Polar Advantage C_{18} column (3 μm particle size, 2.1 × 150 mm, Thermo Scientific), using 0.1% formic acid in water (solvent A) and acetonitrile (solvent B) delivered at 0% B for 5 min; increased to 15% B for 15 min; increase to 100% B in 0.1 min and held at 100% B for 10 min; decrease to 0% B in 1 min and re-equilibrated for 15 min. The flow rate used is 0.200 mL/min. The HPLC effluent from the first 5 min

was diverted to waste to minimize the contamination of the ESI source. The MS was operated in the positive ion mode. Operating parameters were as follows: ESI capillary voltage, 4000 V; gas temperature, 300 °C; drying gas flow, 8 L/min; nebulizer pressure, 35 psi. Samples were analyzed in multiple reaction monitoring (MRM) mode, with the following transitions: *m/z* 326 → 210 and 326 → 164 for N^2-CMdG and *m/z* 331 → 215 and 331 → 169 for [^{15}N(U)]-labeled N^2-CMdG. Calibration curves for the labeled and unlabeled forms were constructed by plotting the MRM signal ratios between the labeled and unlabeled forms against their corresponding concentration ratios. Quantitation of N^2-CMdG in each sample was achieved using the MRM signal ratio between analyte of interest and its isotope-labeled internal standard and the response curve.

Synthesis of the glyoxylate-ARP standard. 1 mmol sodium glyoxylate was reacted with 100 mmol ARP in 10 mL acetate/acetic acid buffer pH 5. The reaction was allowed to stir for 6 h. The product from the reaction was then purified by HPLC with water/acetonitrile gradient. The structure of the standard was characterized by MS^2.

Quantitation of glyoxylate in cytosol of M. smegmatis. 3×10^8 cells for each sample: WT_7H9, WT_acetate, ICD_7H9 and ICD_acetate were subject to cell lysis using phenol chloroform (1 mL) and bead-beating as described earlier in genomic extraction. Then, 500 □L of Tris-HCl buffer (10 mM, pH 8.0) was added and, following vortexing and centrifugation at 15,000xg, the aqueous upper phase was removed. The Tris-HCl buffer extraction was repeated again. The extracts were then reacted with biotinylated aldehyde reactive probe (ARP) and subject to pre-purification using HPLC. The fractions were collected according to the retention time of the standard glyoxylate-ARP and evaporate *in vacuo* prior to reconstitution in water for LC-MS/MS analysis.

Analysis of glyoxylate-ARP adduct by LC-MS/MS (ion trap). An aliquot of the sample was analyzed by HPLC-ESI-MS/MS using an Agilent 1200 series HPLC system interfaced with an Agilent 6430 ion trap mass spectrometer. Samples were resolved on a Dionex Acclaimed Polar Advantage C_{18} column (3 μm particle size, 2.1 × 150 mm, Thermo Scientific), using 0.1% formic acid in water

(solvent A) and acetonitrile (solvent B) delivered at 0% B for 5 min; increased to 30% B for 25 min; increase to 100% B in 0.1 min and held at 100% B for 10 min; decrease to 0% B in 1 min and re-equilibrated for 15 min. The flow rate used is 0.200 mL/min. The HPLC effluent from the first 5 min was diverted to waste to minimize the contamination of the ESI source. The MS was operated in the positive ion mode. Operating parameters were as follows: ESI capillary voltage, 4000 V; gas temperature, 325 °C; drying gas flow, 10 L/min; nebulizer pressure, 20 psi. Samples were analyzed in multiple reaction monitoring (MRM) mode, with the following transitions: *m/z* 388 → 370. External calibration curve was constructed by plotting the MRM signal peak area. The optimized parameter are shown in Table 5.3.

Treatment of pUC19 plasmid with glyoxylate. The pUC19 plasmid containing *LacZ* (10 μg) was treated with each concentration of glyoxylate (0, 1, 2 and 4 mM) in 10 mM sodium acetate (pH 7.4) at 37 °C for 16 h in a total volume of 100 mL.

Transformation of glyoxylate-treated plasmid into E. coli. ~10 ng of the plasmid was transformed into E. coli by heat shock at 42 °C for 45 s. Then, the cells were placed on ice bath (4 °C) for 2 min. 500 mL of Super Optimal Broth with Catabolite repression media (SOC) was added to E. coli. Each tube was then place at 37 °C and rotated for 60 min. The first portion of the cells was then plated onto 1.5% agar plates. The others were plated onto 1.5% agar plates containing 100 mg/mL ampicillin, 20 mg/mL IPTG (Roche), and 20 mg/mL X-gal (Roche). TE was defined as the number of colony forming units (cfu) produced by 10 *n*g of pUC19 DNA in a transformation reaction normalized by the number of total cells plated on 1.5% agar plates without ampicillin.

Results

Optimizing the recovery of N^2-CHMdG in genomic DNA by reductive stabilization. However, for analysis of glyxoylate adducts in genomic DNA isolated from cells, the instability of N^2-CHMdG needed to be managed in the steps prior to $NaCNBH_3$ reduction, which are more involved with mycobacteria due to the extremely thick and waxy cell wall. To this end, DNA isolation from Msmeg was conducted under acidic (pH 6) and cold (0-4 °C) conditions, with standard phenol/$CHCl_3$ solutions adjusted to pH 6. Disruption of the mycobacterial cell wall was accomplished by bead beating with the modification of adding the acidified phenol/$CHCl_3$ to the cells in the bead beater and performing a highly energetic beating for only one minute. This was followed immediately by addition of $NaCNBH_3$ to reduce the N^2-CHMdG to N^2-CMdG. The sensitivity of the analytical method for detection of potentially reduced levels of N^2-CHMdG was enhanced by HPLC pre-purification of N^2-CMdG (Figure 5.1) to remove interference from the canonical nucleosides (A, C, G and T) and other debri that could cause ion suppression during LC-MS/MS analysis. The HPLC pre-purification also allowed quantification of the canonical nucleosides to increase precision of the assay by controlling for sample variation.

The shift to glyxoylate metabolism in Msmeg is accompanied by increases in N^2-CMdG adducts in the genome. Optimization of the LC-MS/MS analytical tool allowed us to proceed to test the hypothesis that a shift in Msmeg metabolism to the glyoxylate cycle would lead to increases in glyxoylate adduct formation. To this end, Msmeg was cultured in three conditions, including wild-type in rich 7H9 medium, wild-type in the minimal M9 medium

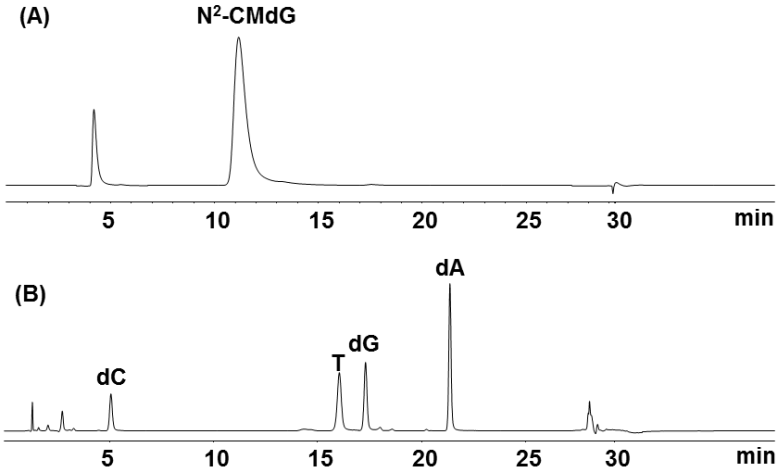


Figure 5.1. Pre-purification of N^2-CMdG at its retention time (A) out of the canonical nucleosides (B)

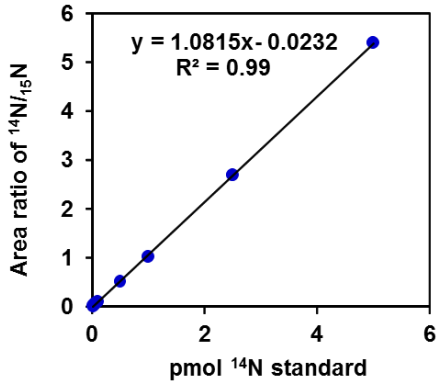


Figure 5.2. Isotope-dilution calibration curve of N^2-CMdG for quantitation of the adduct in *M. smegmatis* genome

supplemented with sodium acetate, and the ΔICD mutant in 7H9 medium, was extracted in the mildly acidic condition (pH ≈ 5) at 4°C in order to reduce the conversion of N^2-CHMdG to dG in basic condition. *M. smegmatis* was cultured in the minimal media containing only acetate because we wanted to test the hypothesis that glyoxylate can affect the level of the adduct by shifting the metabolic flux to the glyoxylate cycle. LC-MS/MS analysis revealed that N^2-CMdG is detected in *M. smegmatis* in all the three conditions (Figure 5.3A). The background level of the N^2-CMdG adduct in wild type cultured in the 7H9 medium is ≈100 adducts/10^6 nucleotides (Figure 5.3B). As expected, the N^2-CMdG adduct level in wild type *M. smegmatis* in acetate media is increased ≈2-3 fold relative to wild type cultured in the 7H9 medium. The level of the adduct detected in wild type cultured in the acetate medium is as much as that detected in the ΔICD mutant. The data indicate that electrophilic metabolites generated endogenously, in this case glyoxylate, can modify nucleophilic molecules in biological systems like thioester (succinyl CoA, malonyl CoA), phosphate ester (1,3-BPG) and α, β-unsaturated carbonyls (fumarate).

The shift to glyxoylate metabolism in Msmeg is not accompanied by increase in the steady-state level of glyoxylate. To test the hypothesis that elevated glyxoylate levels from the shift to the glyoxylate cycle would drive DNA adduct formation, a method was developed to quantify glyoxylate in cytosolic extracts from Msmeg. To do so, biotinylated aldehyde reactive probe (ARP) is reacted with glyoxylate to generate the Schiff's base for the synthesis of the standard (Figure 5.4A). The reaction product is characterized by MS^2 (Figure 5.4A). Then, an LC-MS/MS method was developed in the MRM mode for optimizing the parameters for the Schiff's base (Table 5.3), with the limit of detection ~5 fmol. Then, ARP was added to the cell extracts from Msmeg and the reaction was pre-purified by HPLC. The fractions from each sample were then analyzed by the LC-MS/MS developed method. Unexpectedly, the levels of glyoxylate in each sample: WT-7H9, WT-Ac^+, ICD-7H9 and ICD-Ac^+ are not significantly different (Table 5.2, Figure 5.4).

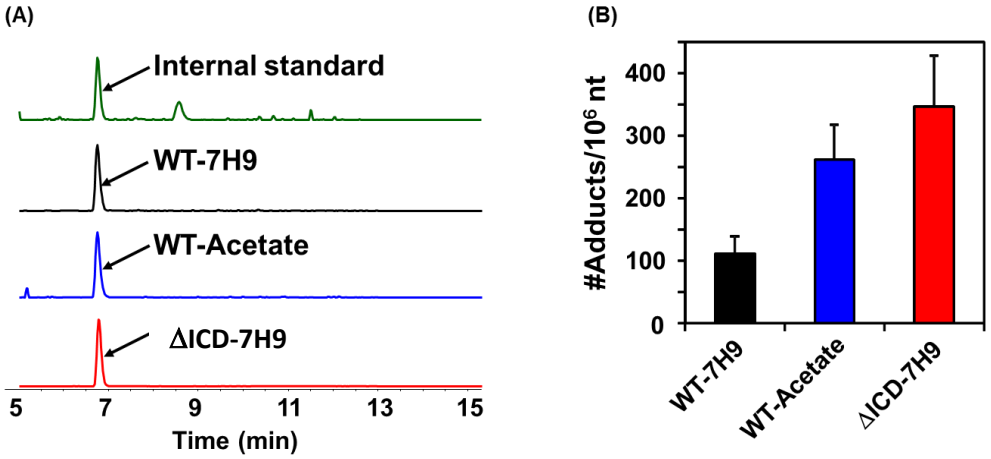


Figure 5.3. Detection of the adduct in WT_7H9, WT_Acetate and ΔICD_7H9 shown in the chromatogram (A); Quantitation data of the adduct that is increased ~2-3 times in WT_Acetate and ΔICD_7H9 relative to WT_7H9.

Table 5.1. Isotope dilution-based quantitation of the adduct in WT cultured in 7H9 and acetate media vs. ICD-dificient mutant

Strain	^{14}N	^{15}N	ratio	pmol/100uL	pmol DNA	#adduct/10^6
7H9_WT1	1602.00	607.00	2.64	28.28	453679.26	62.33
7H9_WT2	1672.00	712.00	2.35	25.13	281406.37	89.30
7H9_WT3	1725.00	446.00	3.87	41.57	329379.06	126.20
Ac_WT1	1072.00	423.00	2.53	27.14	133471.63	203.35
Ac_WT2	1100.00	297.00	3.70	39.79	128522.88	309.62
Ac_WT3	2289.00	517.00	4.43	47.62	161821.61	294.30
7H9_ICD1	1984.00	236.00	8.41	45.34	146437.43	309.61
7H9_ICD2	8700.00	284.00	30.63	165.58	178352.16	428.54
7H9_ICD3	2249.00	290.00	7.76	41.81	138498.15	301.90

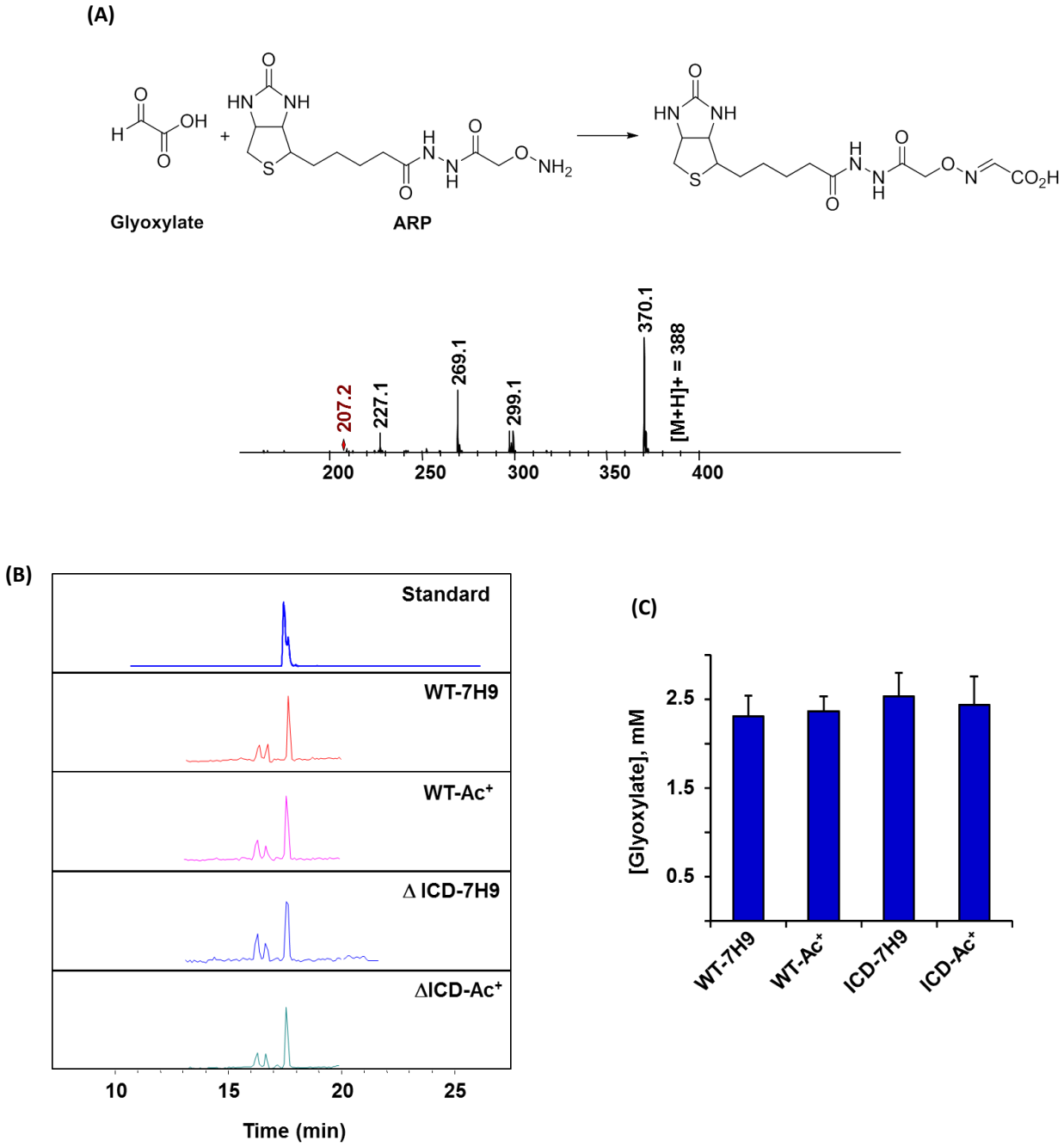


Figure 5.4. Using aldehyde reactive probe (ARP) to characterize glyoxylate in Msmeg cytosol. (A) The reaction of glyoxylate with ARP and MS2 characterization of the product. (B) Qualitative analysis of glyoxylate in WT-7H9, WT-Ac+, ΔICD-7H9 and ΔICD-Ac+. (C) Quantitation of glyoxylate using ARP by external calibration in WT-7H9, WT-Ac+, ΔICD-7H9 and ΔICD-Ac+.

Table 5.2. Glyoxylate quantitation (mM) by derivatizing with ARP via external calibration method

Strain-media	[Glyoxylate], mM[1]
WT-7H9	2.3 ± 0.23
WT-Ac$^+$	2.4 ± 0.17
ΔICD-7H9	2.5 ± 0.27
ΔICD-Ac$^+$	2.4 ± 0.32

[1]The concentration is normalized to the number of cells (3×10^8) and assume that the volume of single cells is 0.65 μm^3
Ac$^+$ is M9 supplemented with NaOAc, proline, glutamine and α-ketoglutarate

Table 5.3 Optimization parameter of glyoxylate-ARP adduct

Compound	Compound stability	Fragmentation Amplitude	MRM transition	LOD
Glyoxylate-ARP adduct	20%	70%	$388 \rightarrow 370$	5 fmol

Transformation efficiency and the blue-white assay. As shown in Table 5.4 and Figure 5.5, the transformation efficiency of incorporating the glyoxylated-treated plasmid into E. coli is increased at 1 mM and then decreased when the concentration of glyoxylate is increased. The phenomenon here is empirical, and Essigmann et al. also observed the same behavior of transformation [9]. However, all the transformants observed for each glyoxylate concentration are all blue without white colony forming on the ampicillin plate. The observation is discussed below in the discussion section.

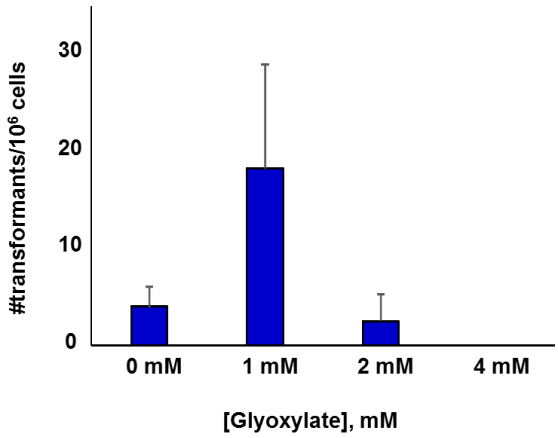


Figure 5.5. Transformation efficiency of incorporating glyoxylated-pUC19 plasmid in *E. coli*

Table 5.4 Transformation efficiency of pUC19 plasmid DNA treated with series concentration of glyoxylate

[Glyoxylate], mM	#Transformants/10^6 cells
0	4.0 ± 2.1
1	18 ± 11
2	2.5 ± 2.8
4	0

Discussion

Building on the observation in Chapter 4 that the electrophilic glyoxylate product of intermediary metabolism reacts with DNA to form N^2-CHMdG adducts, we proceed in this chapter to explore the formation of N^2-CHMdG in Msmeg subjected to nutrient deprivation that induces a shift to the glyoxylate cycle. The motivation for these studies lies in the higher than expected mutation rate in mycobacteria induced to enter the non-replicative state of dormancy induced by nutrient deprivation or hypoxia. While the LC-MS/MS method developed in Chapter 4 proved extremely sensitive, the approach to DNA purification from Msmeg had to be optimized to minimize

loss of unstable N^2-CHMdG during cell disruption and DNA purification. This was accomplished by working at low temperatures throughout and by acidifying the cell disruption and DNA workup conditions to prevent general base catalyzed hydrolysis of N^2-CHMdG. The sensitivity and precision of the method was further enhanced by HPLC pre-purification of the N^2-CHMdG.

The first major point is that the glyoxylate adduct is indeed detected in cells under all conditions. The level of detected N^2-CMdG adducts seems high at 1 per 10^4 nt, perhaps due to a background signal contributing to the apparent adduct signal, such as the existence of N^2-CMdG adducts in the Msmeg genome prior to reduction of N^2-CHMdG adducts during DNA isolation. Alternatively, it could simply be that Msmeg has a high level of N^2-CHMdG adducts due to a high dependence on glyoxylate metabolism (i.e., fatty acid metabolism) even under normal nutrient conditions. These two explanations for high glyoxylate adduct levels could be tested by comparing N^2-CMdG levels in human, yeast and other bacterial cells.

The second major point is that the N^2-CHMdG adduct level increases two-fold in Msmeg under both nutrient restriction conditions with acetate supplementation and in mutants lacking ICD to force metabolism through the glyoxylate cycle. This observation supports the idea that shifts in intermediary metabolism can lead to an increased burden of potentially mutagenic adducts and also supports the observation of unexpectedly high mutation rates in mycobacteria in non-replicative states caused by metabolic shifts, as discussed shortly. However, the lack of change in the steady-state level of glyoxylate in the two conditions studied here was unexpected in light of the increase in glyoxylate DNA adducts. It is possible that the glyoxylate measurement includes protein-bound or otherwise “sequestered” glyoxylate that could form due to its electrophilic nature. Indirect evidence from UV spectroscopy showed that glyoxylate can bind to albumin [10], lysine, arginine and glucosamine [11]. Alternatively, glyoxylate can be further converted to other metabolites such as glycine by aminotransferases where glyoxylate acts as electron acceptor during

anaerobic condition during the shutdown of electron transport chain [12], amino levulinate that is a precursor for heme synthesis [13].

Another important point is that the mutagenesis caused by glyoxylate assayed by transforming pUC19 plasmid DNA treated with various concentrations of glyoxylate into E. coli. As mentioned earlier, all the transformants are blue regardless of the concentration of glyoxylate. One may argue that glyoxylate may not cause mutation due to no white colony formation, resulting from the mutation in *LacZ* gene. However, it can also be argued that at 4 mM glyoxylate, no transformant is observed (Figure 5.5, Table 5.4), which implies that glyoxylate at high concentration might cause random mutation to ampicillin resistant gene in pUC19 plasmid. In this way, we would like to conclude that this blue-white assay may not be sensitive enough to test the mutation caused by metabolic-derived glyoxylate.

References

1. Kumar, A., et al., *Mycobacterium tuberculosis DosR regulon gene Rv0079 encodes a putative, 'dormancy associated translation inhibitor (DATIN)'.* PLoS One, 2012. **7**(6): p. e38709.
2. McKinney, J.D., et al., *Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase.* Nature, 2000. **406**(6797): p. 735-8.
3. Balcells, M.E., et al., *Isoniazid preventive therapy and risk for resistant tuberculosis.* Emerg Infect Dis, 2006. **12**(5): p. 744-51.
4. Cattamanchi, A., et al., *Clinical characteristics and treatment outcomes of patients with isoniazid-monoresistant tuberculosis.* Clin Infect Dis, 2009. **48**(2): p. 179-85.
5. Ford, C.B., et al., *Use of whole genome sequencing to estimate the mutation rate of Mycobacterium tuberculosis during latent infection.* Nat Genet, 2011. **43**(5): p. 482-6.
6. King, G.M., *Uptake of carbon monoxide and hydrogen at environmentally relevant concentrations by mycobacteria.* Appl Environ Microbiol, 2003. **69**(12): p. 7266-72.
7. de Carvalho, L.P., et al., *Metabolomics of Mycobacterium tuberculosis reveals compartmentalized co-catabolism of carbon substrates.* Chem Biol, 2010. **17**(10): p. 1122-31.
8. Jentoft, N. and D.G. Dearborn, *Labeling of proteins by reductive methylation using sodium cyanoborohydride.* J Biol Chem, 1979. **254**(11): p. 4359-65.
9. Kim, M.Y., et al., *AlkB influences the chloroacetaldehyde-induced mutation spectra and toxicity in the pSP189 supF shuttle vector.* Chem Res Toxicol, 2007. **20**(8): p. 1075-83.
10. Martinez-Rodriguez, R., et al., *Synaptic immunolocalization of glyoxylate-complex molecules in the striate areas of the rat cerebral cortex: light and electron microscopic studies.* J Neurosci Res, 1998. **51**(2): p. 268-74.
11. Dutta, U., et al., *Non-enzymatic interactions of glyoxylate with lysine, arginine, and glucosamine: a study of advanced non-enzymatic glycation like compounds.* Bioorg Chem, 2007. **35**(1): p. 11-24.
12. Wayne, L.G. and K.Y. Lin, *Glyoxylate metabolism and adaptation of Mycobacterium tuberculosis to survival under anaerobic conditions.* Infect Immun, 1982. **37**(3): p. 1042-9.
13. de Carvalho, L.P., et al., *Activity-based metabolomic profiling of enzymatic function: identification of Rv1248c as a mycobacterial 2-hydroxy-3-oxoadipate synthase.* Chem Biol, 2010. **17**(4): p. 323-32.

Chapter 6

Conclusions

Goal of this thesis

The goal of the studies is to develop biomarkers of inflammation and oxidative stress. We hypothesize that DNA damage products, once generated at the site of damage, are metabolized before excretion into urine and bile. The rationale of considering metabolism as another variable playing an important role in the biomarker discovery is that M_1dG is oxidized to 6-oxo-M_1dG prior to excretion into urine and feces [1-3]. In this thesis, we use base propenal, a 4'-oxidation product from 2'-deoxyribose oxidation, as a representative damage product for the studies. Then, LC-MS/MS methods were developed through the in vitro studies (Chapter 2) and served as an analytical tool for characterizing the metabolites arising from base propenal metabolism in vivo (chapter 3).

DNA damage serves as the target for biomarker development

The first step of biomarker development is to choose the target for sampling [4]. Here, we chose DNA damage products as the target for our studies. DNA damage causes mutation; so that it fits with mechanistic models associated with pathophysiology of diseases such as cancer [5-7]. The product spectrum of DNA damage covers all types of damage chemistries: oxidation, nitration, nitrosative deamination, alkylation and halogenation [8, 9] (Figure 1.1). In addition, quantitation data of DNA damage products correlates with oxidative stress in animal models. For these three reasons as well as its α,β-unsaturated aldehyde electrophilcity, we chose base propenal as our model compound for the studies of biomarker discovery.

Development of LC-MS/MS analytical platform for DNA damage analysis in vivo

The next step is to choose and develop analytical methods. In this thesis, we choose LC-MS as our analytical method given its high sensitivity in limit of detection that is suitable for analyzing low level of metabolites the biological sample. To establish the LC-MS/MS method (Chapter 2), we

first performed the in vitro reactions of thymine propenal to obtain possible metabolites of thymine propenal with GSH, given that α,β-unsaturated aldehyde serves as ideal substrates for GSH/GSTs systems. After that, chemical standards were synthesized due to lack of commercially available GSH, CysGly, Cys and N-acetylcysteine standards we identified as potential candidates for the studies in vivo. Then, the LC-MS/MS analytical platform in the MRM mode were developed based on the synthesized standards to obtain the best sensitivity prior to analysis in biological samples. Then, the analytical methods were validated by analysis of rat urine and bile injected with thymine propenal. The recovery of thymine propenal shown as various metabolites of mercapturic acid conjugates with its oxidized/reduced products was ~25%, total amount detected in urine and bile. Here, we succeeded in developing and validating the LC-MS/MS analytical tool for searching biomarkers as well as proving the concept that metabolism is another crucial factor to be considered for biomarker sampling.

Application of the LC-MS/MS method for discovering biomarkers of inflammation and oxidative stress

The last step is to investigate the change in the level of the markers associated with pathophysiology of inflammation and oxidative stress. Here, animal models for our oxidative stress/biomarker sampling studies were generated by administering bleomycin and CCl_4, the toxicants known for generation of 4'-oxidation of deoxyribose and biomolecule-damaging agents, respectively. According to our results, Bis-NAc^{red} was the only adduct detected in rat urine from bleomycin, CCl_4 and saline-treated rats, with the ~3-4 fold increase in the first two groups. To this end in this first thesis project, we developed DNA damage biomarkers based on in vitro to in vivo studies, along with LC-MS/MS method development and validation. Eventually, we could show that base propenals can be metabolized to different compounds including mercapturic acid derivatives

and oxidized/reduced base propenals; and only one compound, Bis-NAc^{red}, was detected and increased in the endogenously induced animal model.

Metabolic switching-mediated DNA modification in mycobacteria

The other thesis project is related to DNA modification caused by electrophilic species arising from central metabolism. In this thesis, *M. smegmatis* is the model organism for the studies. There was evidence showing that the glyoxylate cycle is upregulated during mycobacteria dormancy [10, 11]. Glyoxylate is the key compound in this shunt, where mybacteria exploit this compound for energy conservation during its dormancy for their survival using lipids as the source [12]. Glyoxylate is an aldehydic compound, which is apparently electrophilic and potentially able to modify nucleophilic biomolecules in cells. There was also a lot of evidence showing that other electrophilic metabolites from primary metabolism, such as succinyl CoA, malonyl CoA, fumarate, etc., can modify lysine and cysteine side chain in proteins and in turn regulate the function of those proteins [13-15]. Therefore, we hypothesized that glyoxylate could also modify nucleophiles in cells, in this case, DNA is the target for glyoxylate modification. This leads to our goal to characterize glyoxylate adduct in DNA and we expect that the adduct should be higher in *M. smegamatis* that is forced to switch the metabolism to the glyoxylate cycle.

To characterize the glyoxylate adduct, we first established LC-MS/MS methods by analyzing the reaction products between glyoxylate and dG. However, the product from the reaction was unstable due to its carbinolamine nature and needs stabilizing by $NaCNBH_3$. Then, LC-MS/MS (MRM) was developed to gain the best sensitivity, along with chemically synthesized standard for method development and isotope-dilution quantitation. The validation of the a developed LC-MS/MS analytical method was necessary prior to analysis in M. smegmatis genome. In this case, calf thymus DNA reacting with glyoxylate was used as reference for validation. Finally, the DNA from the mycobacteria was analyzed and we found that the level of the adduct was increased around ~2-3 times in the wild type forced to go to the glyoxylate pathway by culturing in M9 medium

supplemented with acetate. The increase (2 – 3 times) is also similar in the isocitrate dehydrogenase-deficient strain of *M. smegmatis*.

References

1. Otteneder, M.B., et al., *In vivo oxidative metabolism of a major peroxidation-derived DNA adduct, M1dG.* Proc Natl Acad Sci U S A, 2006. **103**(17): p. 6665-9.
2. Knutson, C.G., et al., *Monitoring in vivo metabolism and elimination of the endogenous DNA adduct, M1dG {3-(2-deoxy-beta-D-erythro-pentofuranosyl)pyrimido[1,2-alpha]purin-10(3H)-one}, by accelerator mass spectrometry.* Chem Res Toxicol, 2008. **21**(6): p. 1290-4.
3. Knutson, C.G., et al., *Metabolism and elimination of the endogenous DNA adduct, 3-(2-deoxy-beta-D-erythropentofuranosyl)-pyrimido[1,2-alpha]purine-10(3H)-one, in the rat.* J Biol Chem, 2007. **282**(50): p. 36257-64.
4. Halliwell, B., *Effect of diet on cancer development: is oxidative DNA damage a biomarker?* Free Radic Biol Med, 2002. **32**(10): p. 968-74.
5. Barnes, D.E. and T. Lindahl, *Repair and genetic consequences of endogenous DNA base damage in mammalian cells.* Annu Rev Genet, 2004. **38**: p. 445-76.
6. Delaney, J.C. and J.M. Essigmann, *Biological properties of single chemical-DNA adducts: a twenty year perspective.* Chem Res Toxicol, 2008. **21**(1): p. 232-52.
7. You, Y.H. and G.P. Pfeifer, *Similarities in sunlight-induced mutational spectra of CpG-methylated transgenes and the p53 gene in skin cancer point to an important role of 5-methylcytosine residues in solar UV mutagenesis.* J Mol Biol, 2001. **305**(3): p. 389-99.
8. De Bont, R. and N. van Larebeke, *Endogenous DNA damage in humans: a review of quantitative data.* Mutagenesis, 2004. **19**(3): p. 169-85.
9. Epe, B., *Role of endogenous oxidative DNA damage in carcinogenesis: what can we learn from repair-deficient mice?* Biol Chem, 2002. **383**(3-4): p. 467-75.
10. Schnappinger, D., et al., *Transcriptional adaptation of Mycobacterium tuberculosis within macrophages: Insights into the phagosomal environment.* Journal of Experimental Medicine, 2003. **198**(5): p. 693-704.
11. Tailleux, L., et al., *Probing Host Pathogen Cross-Talk by Transcriptional Profiling of Both Mycobacterium tuberculosis and Infected Human Dendritic Cells and Macrophages.* Plos One, 2008. **3**(1).
12. Eisenreich, W., et al., *Carbon metabolism of intracellular bacterial pathogens and possible links to virulence.* Nature Reviews Microbiology, 2010. **8**(6): p. 401-412.
13. Moellering, R.E. and B.F. Cravatt, *Functional Lysine Modification by an Intrinsically Reactive Primary Glycolytic Metabolite.* Science, 2013. **341**(6145): p. 549-553.
14. Park, J., et al., *SIRT5-Mediated Lysine Desuccinylation Impacts Diverse Metabolic Pathways.* Molecular Cell, 2013. **50**(6): p. 919-930.
15. Weinert, B.T., et al., *Lysine Succinylation Is a Frequently Occurring Modification in Prokaryotes and Eukaryotes and Extensively Overlaps with Acetylation.* Cell Reports, 2013. **4**(4): p. 842-851.

Appendix A

Preliminary identification of DNA modifications in S. Montevideo: 2'-deoxy-7-deazaguanosine in DNA

Introduction

Queuosine is one of the 7-deazapurines found in tRNA at the wobble position [1]. However, comparative genomic data from our collaborator, Dr. Jennifer Thiaville from de Crecy lagard lab in the University of Florida, indicated that a novel gene—tgtA5—functionally analogous to tRNA guanine transglycosylases (TGTs) is clustered with DNA-binding proteins such as a DndB-like protein, SNF/SWII helicases, and DNA repair proteins. TGT is the protein functions by cleaving C – N glycosidic bond of rG and constructs the bond of depurinated ribose with Q at the wobble position of tRNA (G^{34}) [2]. In addition, organisms such as mycobacteria have incomplete queuosine biosynthesis pathway and these intermediates were not detected in tRNA, which raises the question why they have these genes in their genomes [3]. Furthermore, other classes of 7-deazapurines, i.e. toyocamycin, were detected as mono-, di- and tri- 2'-deoxyribonucleotides in the nucleotide pool [4]. These pieces of evidence led us to characterizing the structure of 2'-deoxyqueuosine and its derivatives in DNA from S. Montevideo, one of the organisms having tgtA5 clustered with the DNA-binding proteins. We hypothesize that DNA-binding proteins (e.g., DndB-like protein, helicases) may help TgtA5 bind to DNA, and TgtA5 will in turn catalyze the substitution of guanine with queuine and its derivatives in DNA via transglycosylation reaction similar to the function of Tgt.

Materials & methods

Materials. Formic acid, calf thymus DNA, DNase I, benzonase, calf intestine phosphatase, were purchased from Sigma-Aldrich (St. Louis, MO). Phosphodiesterase I was purchased from Affymetrix (Cleveland, Ohio). Deionized water was further filtered through MilliQ systems (Millipore corporation, Bedford, MA) and used in the whole experiment.

DNA digestion. ~100 μg of DNA from S. Montevideo and S. Typhimurium was dissolved in 10 mM Tris-buffer, pH 7.9 and 1 mM $MgCl_2$ in two different tubes. Benzonase (20U), DNase I (30U), calf intestine phosphatase (34U) and phosphodiesterase (0.5U) were added to perform DNA digestion. The reaction was allowed to complete in 6 hours and then filtered through 10 kDa filter to separate the digested nucleosides from proteins. Then, the filtrate was collected and pre-purified (described below) to get rid of salt and Tris-buffer.

Pre-purification. The digested nucleosides were subject to pre-purification using Agilent 1100 with this following solvent program: 99.5% solvent A (0.1% formic acid in water) and 0.5% solvent B (acetonitrile) for 5 min, a linear gradient to 25% solvent B in 20 min, increasing to 100% solvent B in 0.1 min, holding for 10 min, decreasing to 0.5% solvent B in 1 min, and reequilibrating at 0.5% solvent B for 15 min. The fraction of the 2'-deoxynucleosides was collected from 14 min (the time point at which dC was eluted) to 35 min (the time point at which solvent B wash was finished) in order to ensure that all the types of 2'-deoxynucleosides were collected. After that, the fraction was concentrated *in vacuo* and reconstituted in water.

Sample analysis by LC-MS/MS. Samples (20 mL) were loaded onto a Dionex Acclaimed Polar Advantage II C18 column (2.1 x 250 mm, 3 mm) equilibrated with 95.5% solvent A (0.1% formic acid in water) and 0.5% solvent B (acetonitrile) at a flow rate of 0.25 mL/min. The solvent was programmed as follows: a linear gradient from the starting solvent to 22% B in 15 min, increasing to 100% B in 0.1 min, holding for 10 min, decreasing to 0.5% solvent B in 1 min, and reequilibrating at initial conditions for 10 min. Eluting compounds were detected by mass spectrometer (Agilent 6430, triple quadruple). Positive ion electrospray ionization was used with the following parameters: spray voltage 3.8 kV, capillary temperature 325°C, nebulizer gas 30 psi, fragmentor voltage 76 V, and 4 V collision energy. Samples were analyzed by selected reaction monitoring, which followed these ion transition: 326→210, 311→195, 292→176, 296→180, 394→278 and 309→193 for 2'-deoxyCPH_4, 2'-deoxyCDG, 2'-deoxypreQ_0, 2'-deoxypreQ_f, 2'-deoxyQ and 2'-deoxyArch, respectively.

Results & Discussion

Due to the lack the chemical authentic standard of Q and its derivatives and the difficulty in chemical synthesis, LC-MS/MS method development is impossible to be developed. We first proposed to optimize the parameter of mass spectrometer using dG as the ballpark based on the similar N-glycosidic bond of dG and 7-deazapurine species. After optimization, the parameters including fragmentation voltage and collision energy were obtained: 76 V and 4 V, respectively.

We then performed selected reaction monitoring (SRM) in the mass spectrometer with the transition of parent ions to daughter ions (Figure 1) in order to characterize those six deoxynucleosides in DNA samples of *S. Montevideo* and *S. Typhimurium LT2*. These six target compounds were selected according to intermdiates in the queuosine biosynthetic pathway (scheme 1 in attachment). The targeted analysis (SRM) reveals three deoxynucleosides in DNA of *S. Montevideo*: 2'-deoxyCDG (figure 2a), 2'-deoxyArch (figure 2b), and 2'-deoxypreQ_0 (figure 2c) while no signal is detected in DNA of *S. Typhimurium LT2* (figure 2d-f). Although 2'-deoxypreQ_1 and 2'-deoxyQ were not detected by this mode of analysis, we could not rule out the possibility that they are present in DNA. Therefore, we changed the mass transition to $[M+2H]^{2+} \rightarrow [bases+2H]^{2+}$ on the grounds that primary and secondary amine in the bases of 2'-deoxypreQ_1 and 2'-deoxyQ are possibly further protonated. However, very weak signals were detected for both compounds. Moreover, we also tried to target the mass transition of 2'-deoxyQ (parent ion) → 2'-deoxypreQ_1 (daughter ion generated from loss of cyclopentenediol ring). This transition was selected according to H. Kasai et al. from GC-MS data. Nevertheless, the intensity of the peak of 2'-deoxyQ was only moderately improved. To this end, it is still ambiguous if the hypermodified 2'-deoxyQ is present in DNA. In this short report, we shows that three potential 2'-deoxy-7-deazaguanosine derivatives including 2'-deoxyCDG, 2'-deoxypreQ_0 and 2'-deoxyArch were detected only in *S. Montevideo* DNA using SRM mass spectrometry. SRM appears to be a relatively specific detection method because

the SRM signal can only be shown only if two specific ions (parent and daughter) are detected. However, chemical synthesis of at least the three compounds (2'-deoxyCDG, 2'-deoxypreQ_0 and 2'-deoxyArch) should be conducted to ensure the structures. Please note that we cannot ignore the fact that 2'-deoxypreQ_1 and 2'-deoxyQ can be present in DNA, since the SRM method used here is not optimized due to lack of chemical standards to optimize the method.

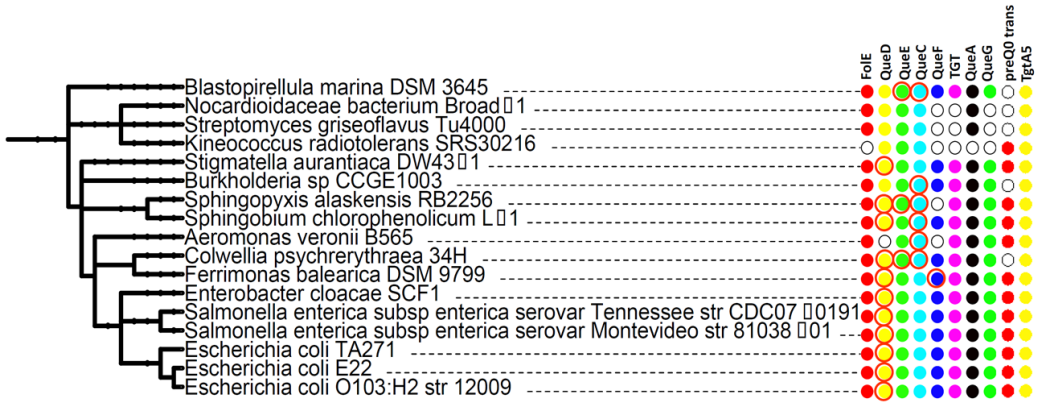


Figure A1. Gene clusters of tgtA5 with Q biosynthetic pathways. The organism of analysis in this appendix, S. Montevideo, shows the gene clustering between tgtA5 and Q biosynthesis. Red circles indicate presence of a two copies of the gene in the organism: one is a Q synthesis enzyme and the second is a paralog located in the tgtA5 cluster. (Data taken from Dr. Jennifer Thiaville from Crecy Lagard lab in the University of Florida)

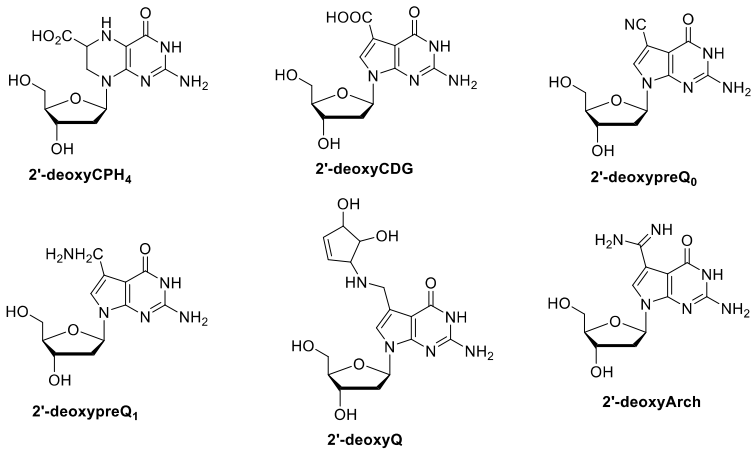


Figure A2. Shows the structure of 2'-deoxy-7-deazaguanosines that are targeted in SRM based on the intermediates in queuosine biosynthesis. The mass transition of SRM, $[M+H]^+ \rightarrow [base+H]^+$, is as follows: 326→210, 311→195, 292→176, 296→180, 394→278 and 309→193 for 2'-deoxyCPH$_4$, 2'-deoxyCDG, 2'-deoxypreQ$_0$, 2'-deoxypreQ1, 2'-deoxyQ and 2'-deoxyArch, respectively. The mass transitions are chosen based on the general trend of neutral loss of 2'-deoxyribose of nucleosides.

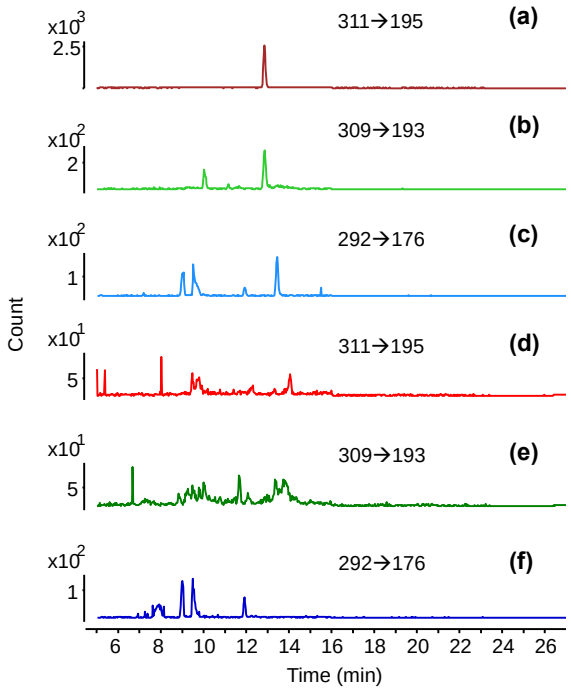


Figure A3. SRM chromatogram of putative 2'-deoxyCDG (mass transition: 311→195), 2'-deoxyArch (mass transition: 309→193) and 2'-deoxypreQ$_0$ (mass transition: 292→176) in DNA of S. Montevideo ((a), (b), (c), respectively) and of 2'-deoxyCDG, 2'-deoxyArch and 2'-deoxypreQ$_0$ in DNA of S. Typhimurium ((d), (e), (f), respectively). The signals of the three compounds are detected only in DNA of S. Montevideo but not S. Typhimurium LT2.

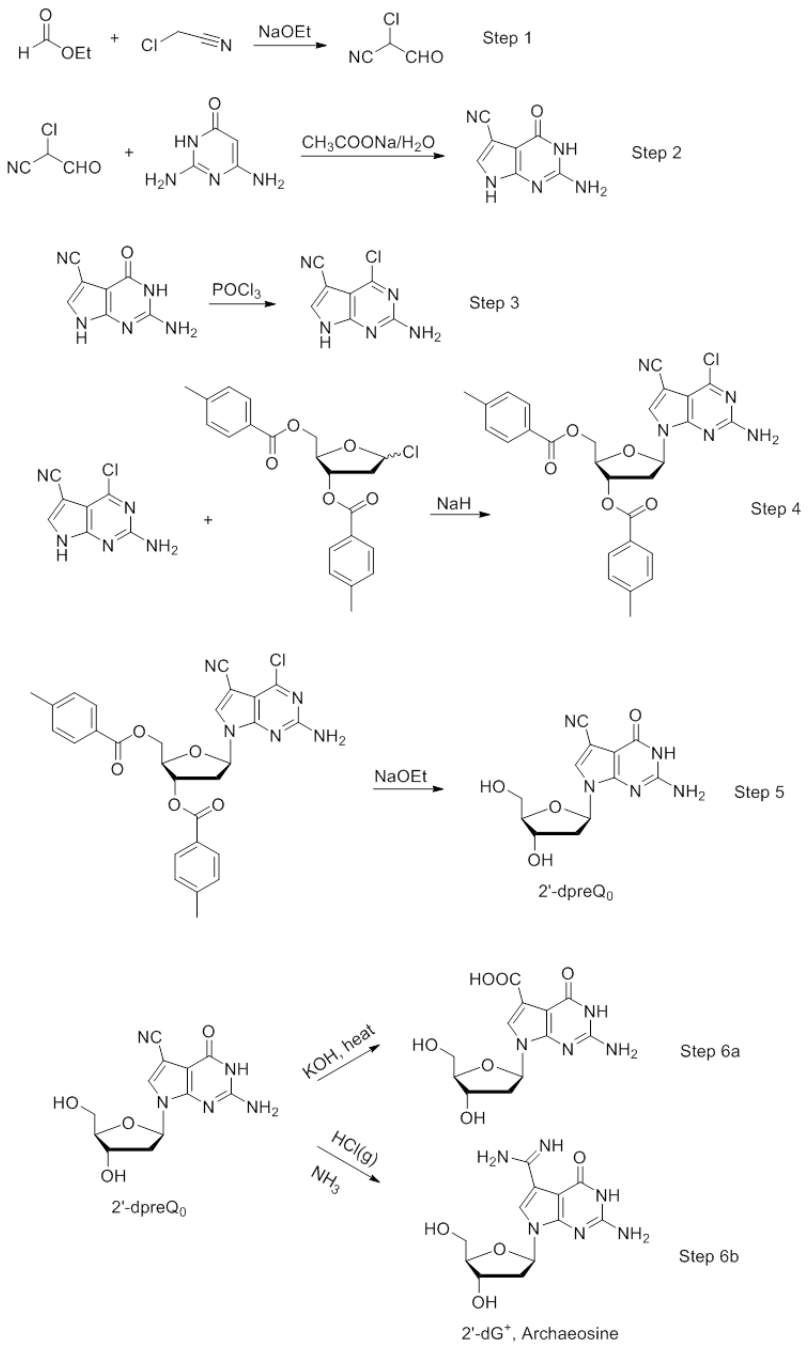


Figure A4. Proposed synthetic scheme for Q derivative candidates. Chemical synthesis of Q and its derivatives has been proposed in this thesis: step 1and 1 [5] step 3 [6], step 4, 5 and 6a [7] and step 6b [8]

References

1. Phillips, G., et al., *Biosynthesis of 7-Deazaguanosine-Modified tRNA Nucleosides: a New Role for GTP Cyclohydrolase I.* Journal of Bacteriology, 2008. **190**(24): p. 7876-7884.
2. El Yacoubi, B., M. Bailly, and V. de Crecy-Lagard, *Biosynthesis and Function of Posttranscriptional Modifications of Transfer RNAs.* Annual Review of Genetics, Vol 46, 2012. **46**: p. 69-95.
3. Iyer, L.M., et al., *Computational identification of novel biochemical systems involved in oxidation, glycosylation and other complex modifications of bases in DNA.* Nucleic Acids Research, 2013. **41**(16): p. 7635-7655.
4. Suhadoln.Rj, T. Uematsu, and H. Uematsu, *Toyocamycin - Phosphorylation and Incorporation into Rna and DNA and Biochemical Properties of Triphosphate.* Biochimica Et Biophysica Acta, 1967. **149**(1): p. 41-&.
5. Moeller, K., et al., *Expression and characterization of the nitrile reductase queF from E. coli.* Enzyme Microb Technol, 2013. **52**(3): p. 129-33.
6. Wilding, B., et al., *Nitrile Reductase from Geobacillus kaustophilus: A Potential Catalyst for a New Nitrile Biotransformation Reaction.* Advanced Synthesis & Catalysis, 2012. **354**(11-12): p. 2191-2198.
7. Ramasamy, K., et al., *Total and Stereospecific Synthesis of Cadeguomycin, 2'-Deoxycadeguomycin, Ara-Cadeguomycin, and Certain Related Nucleosides.* Journal of the Chemical Society-Perkin Transactions 1, 1989(12): p. 2375-2384.
8. Bruckl, T., et al., *A short and efficient synthesis of the tRNA nucleosides PreQ(0) and archaeosine.* Organic & Biomolecular Chemistry, 2007. **5**(23): p. 3821-3825.