Review Article

Formation and Repair of Tobacco Carcinogen-Derived Bulky DNA Adducts

Bo Hang

Life Sciences Division, Department of Cancer and DNA Damage Responses, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Correspondence should be addressed to Bo Hang, bo_hang@lbl.gov

Received 20 May 2010; Revised 16 July 2010; Accepted 17 September 2010

Academic Editor: Ashis Basu

Copyright © 2010 Bo Hang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

DNA adducts play a central role in chemical carcinogenesis. The analysis of formation and repair of smoking-related DNA adducts remains particularly challenging as both smokers and nonsmokers exposed to smoke are repetitively under attack from complex mixtures of carcinogens such as polycyclic aromatic hydrocarbons and N-nitrosamines. The bulky DNA adducts, which usually have complex structure, are particularly important because of their biological relevance. Several known cellular DNA repair pathways have been known to operate in human cells on specific types of bulky DNA adducts, for example, nucleotide excision repair, base excision repair, and direct reversal involving O6-alkylguanine DNA alkyltransferase or AlkB homologs. Understanding the mechanisms of adduct formation and repair processes is critical for the assessment of cancer risk resulting from exposure to cigarette smoke, and ultimately for developing strategies of cancer prevention. This paper highlights the recent progress made in the areas concerning formation and repair of bulky DNA adducts in the context of tobacco carcinogen-associated genotoxic and carcinogenic effects.

1. Introduction

Tobacco was traded from North America to the world about 500 years ago. Since then, tobacco use by smoking cigarettes, cigars, and pipes, or by chewing, has wreaked havoc on mankind. Nearly 1.3 billion people are active smokers worldwide [1], who also pose a threat of indirect exposure to even more nonsmokers through secondhand smoke (SHS, also known as environmental tobacco smoke, ETS). Cigarette smoke accounts for 30% of all cancer deaths. Based on the International Agency for Research on Cancer (IARC), cigarette smoking is associated with cancers in many organs/tissues such as lung, head, neck, and bladder [2]. The lung is particularly vulnerable as ~90% of lung cancer cases are caused by cigarette smoking. Cigarette smoke causes other diseases as well, including pulmonary disorders, cardiovascular diseases and stroke, and developmental defects. There is also sufficient evidence in recent years that SHS causes lung cancer [3]. In US nonsmokers, SHS is responsible for about 3,000 lung cancer deaths, 46,000 cardiac-related illnesses, and 430 sudden infant death syndrome (SIDS) per year [4].

Four types of smoke have been classified so far [5]: (1) Mainstream smoke (MSS), created by tobacco combustion at approximately 1,200–1,600°C, when smokers inhale the tobacco smoke from a burning cigarette, (2) Sidestream smoke (SSS), emanating from the smouldering end of a lit cigarette at ~900°C when no active smoking occurs while the smoker pauses before taking the next puff, (3) SHS, a mixture of about 85% of SSS and 15% of exhaled MSS, and (4) Thirdhand smoke (THS), a newly emerged type, defined as residual tobacco smoke adsorbed onto indoor surfaces after active smoking has ceased, where the semivolatile and nonvolatile components undergo chemical transformation to produce new toxicants [6–8].

In the last 50 years, many studies have been performed to identify chemical toxicants in cigarette smoke, which may represent the most rich resource of exogenous human mutagens and carcinogens. MSS contains more than 4,000 chemicals. Among them, over 60 have been classified by IARC as carcinogens [9]. These include 10 polycyclic aromatic hydrocarbons (PAHs), 6 hydrocarbons, 10 nitrosamines, 13 aromatic amines, 2 aldehydes, 3 phenolic compounds, 4 volatile hydrocarbons, 3 nitro compounds,
12 miscellaneous organic compounds, and 9 inorganic and metals compounds [10]. This list contains some of the strong animal and/or human carcinogens, such as PAHs, N-nitrosamines, and aromatic amines, all of which react with DNA to form adducts [11–13]. The most prevalent ones in the vapor phase are aldehydes, benzene, and butadiene. It should be emphasized that SSS or SHS also contains several thousand individual compounds as does MSS [5], and most of the above-mentioned carcinogens are also present in SSS/SHS [12]. Since such a carcinogenic source, that is, cigarette smoke, is preventable, and DNA adduct levels correlate with cigarette consumption [13], tobacco smoke provides a unique model for understanding the cause-effect or environment-gene relationship in smoking-related cancer development. However, the real assessment of such relationships is very difficult due to multiple reasons [14]. For example, Metabolic activation imposes an additional level of complexity for such assessment. In addition, cigarette smoke contains co-carcinogens and tumor promoters that are also crucial for tumorigenicity of smoke condensates [12, 15, 16].

Although certain carcinogens in cigarette smoke, such as formaldehyde and α,β-unsaturated aldehydes (enal)s, directly react with DNA to form covalent adducts, most of carcinogenic compounds are so-called procarcinogens that must be metabolically activated to form ultimate carcinogens [12]. These metabolites are usually electrophiles that react with the nucleophilic sites on DNA bases. The well-studied microsomal cytochrome P450 (CYP) system [17] activates many tobacco carcinogens such as PAHs, N-nitrosamines, aromatic amines, and benzene [18–20]. Therefore, carcinogen metabolism is often a double-edged sword in that it not only detoxifies and excretes toxicants but may also convert them into harmful reactive species. The individual variation in metabolic activation such as genetic polymorphisms in carcinogen-metabolizing genes is an important determinant of DNA adduct levels and is used to identify smokers with increased cancer risk [21, 22].

Most chemical carcinogens react with cellular DNA as the ultimate target. Cigarette smoke condensates were initially known to have mutagenic activity by the 1970s [23], and adducts were detected in cellular DNA from smokers in the 1980s [24]. Since then, many tobacco carcinogen-derived adducts have been identified in vitro and in vivo, owing to the development of highly sensitive analytical detection methods, such as 32P-postlabeling and mass spectrometry (MS) [25, 26]. It has been shown that cigarette smokers have higher levels of DNA adducts than nonsmokers [12, 13, 27, 28]. Studies have also shown that current smokers have higher adduct levels compared with former smokers [13]. With some exceptions of inconsistency, many epidemiologic and clinical studies have shown an association between the in vivo levels of DNA adducts resulting from cigarette smoke and the occurrence of tobacco-related cancers in lung, head and neck, and bladder [13]. There are numerous reviews specifically related to the relationships between tobacco carcinogen exposure, DNA adduct formation, carcinogen/adduct mutagenic potential, and increased cancer risk related to smoking [13, 19, 26, 27, 29–32].

Tobacco carcinogens generate a broad spectrum of DNA lesions ranging from sugar damage, apurinic/apyrimidinic (AP) sites, small modified bases (e.g., O6-mG and 8-oxoG), and bulky base adducts to more deleterious lesions such as DNA crosslinks and strand breaks. The so-called bulky DNA adducts are formed by the covalent binding of those chemical carcinogens with large size, such as PAHs and aromatic amines, to various sites on DNA bases. These adducts also include exocyclic DNA bases such as the etheno, propano, and benzetheno adducts formed by respective bifunctional compounds [33, 34]. These bulky adducts represent a major and important class of DNA damage originating from exposure to cigarette smoke. One characteristic of these bulky adducts is that they tend to significantly disrupt the DNA helical structure and block Watson-Crick base pairing [35, 36]. They are usually highly mutagenic, as exemplified by the PAH-DNA adducts [30] and exocyclic DNA adducts [34, 37]. Some of them may not be repaired (e.g., benzo[c]phenanthrene N6-dA adducts [38]) or only poorly repaired (e.g., two dibenzo[a,l]pyrene-induced DNA adduct [39]), thus leading to their persistence in genomic DNA. Smokers with high levels of these bulky adducts have been shown to be associated with an increased risk of cancers [40, 41]. In fact, most of the compelling data on the connection of DNA adducts with cancer have been obtained with bulky DNA adducts and their respective carcinogens. For example, PAH- and acrolein-DNA adducts are preferentially formed in the same mutational hotspots of p53 in the lung cancers of smokers [30, 42, 43]. This tumor suppressor gene is frequently mutated in ∼40% of lung cancer cases. There is also evidence that a high level of bulky DNA adducts in tissues, such as those caused by PAHs and vinyl chloride (VC), is associated with an increased risk of tumor in humans and animals [40, 44–46]. It should be pointed out that tobacco smoke also produces reactive oxygen species (ROS) and induces oxidative stress [11, 47]. Those lesions that arise directly from ROS attack of a base (e.g., 8-oxoG) or deoxyribose (e.g., base propenals) [48, 49] could also play a role in tobacco carcinogenesis.

Cigarette smoking can cause complex biological responses. If unrepair ed, DNA adducts may block replication and transcription. There is evidence that only a single BPDE-DNA adduct can effectively block expression of a reporter gene [50]. DNA damage can either activate checkpoint signaling pathways leading to cell cycle arrest or induce cell apoptosis by recruitment of immunologic and inflammatory responses [31]. More importantly, persistence of DNA adducts such as those formed by tobacco carcinogens PAHs and N-nitrosamines plays a central role in tobacco-induced carcinogenesis [27]. These adducts not only represent a very early event by inducing specific genetic changes that are a prerequisite to the initiation of cancer, but also occur during the continuum of the carcinogenic process [31]. DNA adducts can lead to nucleotide misincorporation, thus causing gene mutations. Mutations in the p53 gene are more commonly observed in lung cancers from smokers than nonsmokers [51, 52]. Exposure to smoking has been associated with activating mutations in proto-oncogenes (e.g., the ras gene family) and inactivation of tumor
Figure 1: Schematic presentation of the formation, repair and mutagenic potential of tobacco carcinogen-induced bulky adducts in genomic DNA. The formation of a bulky DNA adduct (marked in red), if not repaired or poorly repaired, results in multiple genetic changes such as cell death and gene mutation. TLS (translesion DNA synthesis) can prevent a mutation by error-free incorporation opposite the adduct. Error-prone incorporation at the adduct site by replicative or translesional DNA pols is the source for mutation. Only those adducts that finally escape all the defense mechanisms may lead to biologically important mutations.

Suppressor genes (e.g., \( p53 \) and \( p16 \)) in cancers [53–56]. Microarray-based analyses also reveal that cigarette smoking alters expression of many genes involved in other functions [57]. In addition to point mutations, there are correlations between DNA adduct levels and other somatic alterations, for example, loss of heterozygosity (LOH) that may occur at the very early stages of tobacco carcinogenesis [58, 59]. In addition, epigenetic changes such as abnormal promoter methylation of certain genes also occur more frequently in lung tumors from smokers than in never-smokers or may appear in healthy individuals who start to smoke [60, 61], highlighting the importance of both genetic and epigenetic changes in tobacco carcinogenesis. Tobacco carcinogens can also contribute to tumorigenesis by interacting with proteins, RNA, and lipids, in addition to DNA.

To avoid tobacco carcinogen-induced DNA damage, quitting smoking or avoiding exposure is the first and foremost important approach. However, once such damage is formed, DNA repair is the next major defense mechanism (Figure 1). Organisms from prokaryotes to mammals have evolved a number of repair pathways, including direct reversal, base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), and double-strand break (DSB) repair [62–64]. In model systems, such as cultured cells and genetically manipulated organisms, these pathways have been shown to operate on specific types of DNA lesions. NER is the major pathway for the repair of various duplex-distorting bulky DNA lesions such as those induced by PAHs. Small alkylated and oxidized lesions, including those arising from endogenous sources, are excised by the BER pathway which also repairs certain single-ring exocyclic DNA adducts. For certain alkylated bases and etheno adducts, they can be repaired through direct reversal carried out by specialized repair proteins. In general, the understanding of damage recognition and mechanism of repair is important to gain insight into the specific roles of tobacco DNA adducts in the development of cancer and other chronic diseases since, at the end, the overall cellular repair capacity in response to exposure is critically related to the levels of DNA adducts in the genome or mutations in genes. The role of individual variability in repair, for example, polymorphisms in repair genes, has been related to the increased cancer risk in smokers [22, 65, 66]. Ultimately, it is the impaired or poor repair of DNA adducts (e.g., those bulky adducts and oxidized bases with cytotoxicity and mutagenicity) that is expected to be most important in the etiology of smoking-related cancer and other disorders.

Studies in the last decade have revealed that if a DNA adduct is unrepaired or irreparable, cells may use translesion DNA synthesis (TLS) to bypass the adduct to ensure the continuum of DNA replication [67–69]. TLS is performed by various specialized DNA polymerases (polS), mostly from the Y-family, with the possibility of nucleotide misincorporation [70, 71]. These enzymes possess an open and preformed active site, enabling accommodation of a broad spectrum of DNA adducts with different structures [69]. In studies reported in literature, error-prone incorporations opposite an adducted nucleotide appear to occur commonly or co-exist with error-free incorporations [72–74]. However, in some other cases, TLS pols perform error-free bypass on damaged DNA templates such as the efficient and correct nucleotide incorporation at the acrolein adduct \( y\)-HO-PdG by pol \( \iota \) and subsequent extension of replication by pol \( \kappa \) [75]. The fidelity of TLS observed in these experiments tends to depend on the individual pol tested as well as the structure of the target adduct. In general, the primary roles of these pols...
and how they operate in the cell with regard to interacting with replication and repair machineries remain to be further understood.

This paper will focus on the formation and repair of bulky/exocyclic DNA adducts induced by the major tobacco carcinogens in relation to tobacco mutagenesis and carcinogenesis. For small base lesions induced by tobacco chemical carcinogens, see previous reviews by Singer [76] and Shrivastav et al. [77]. In general, the literature so far on the covered review topics has been extensive. Therefore, only selected published data are used to illustrate the relevant areas, ideas, and concepts. I regret that this review does not permit acknowledgment of the many researchers who made the original findings in these important areas.

2. Formation and Repair of Bulky DNA Adducts by Tobacco Carcinogens

2.1. Formation of Bulky DNA Adducts: An Overview. Since smokers are repetitively exposed to complex mixtures of genotoxic carcinogens, the collective formation of DNA adducts is very complex as reflected by their chemical types and cellular levels. Although this paper is focused on bulky DNA adducts, other types of DNA lesions by tobacco carcinogens may be equally or more important than bulky DNA adducts for a given carcinogen or cancer type. DNA adduct levels are normally analyzed in target tissues in order to elucidate the relationship between tobacco carcinogens and cancer development. These levels should reach steady state such that the number of newly formed adducts equals the number of adducts lost every day which are related to a number of factors including carcinogen reactivity, exposure doses, timing of exposure, metabolic processes, and DNA repair capacity [13]. Understanding of DNA adducts with regard to their formation, isolation, and identification can be critical in several ways: (1) to understand the mechanism of tobacco carcinogens; for example, the analysis of formation of DNA adduct at gene mutational hotspots [30, 78]; (2) to find biomarkers of tobacco carcinogens; for example, the analysis of formation of DNA adducts for a given carcinogen or cancer type. DNA adduct levels are normally analyzed in target tissues in order to elucidate the relationship between tobacco carcinogens and cancer development. These levels should reach steady state such that the number of newly formed adducts equals the number of adducts lost every day which are related to a number of factors including carcinogen reactivity, exposure doses, timing of exposure, metabolic processes, and DNA repair capacity [13]. Understanding of DNA adducts with regard to their formation, isolation, and identification can be critical in several ways: (1) to understand the mechanism of tobacco carcinogens; for example, the analysis of formation of DNA adduct at gene mutational hotspots [30, 78]; (2) to assess the biologically effective doses of tobacco carcinogens [25]; (3) to assess DNA repair capacity (DRC) towards the adducts [13, 79]; (4) to find biomarkers of tobacco genotoxicity and uptake/metabolism of specific carcinogens [12, 13].

Many types of DNA adducts, including those formed by benzo[a]pyrene (B[a]P), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), N'-nitrosodimethylamine (NNN), and 4-aminobiphenyl (4-ABP), have been detected from tissues of smokers as well as nonsmokers exposed to SHS [80, 84]. The common tobacco carcinogens and related metabolites that give rise to bulky DNA adducts are listed in Figure 2. In tissues, the typical adduct levels are at about 1 adduct in 10^6-10^7 normal bases [27, 83]. In general, DNA adduct levels as low as 1 in 10^6-10^12 normal bases can be significant with definite biological consequences [84]. Although 32P-postlabeling and immunoassay have been extensively utilized for adduct analysis, the detection and identification of DNA adducts at these or even lower levels have been greatly facilitated by the highly sensitive/specific and new types of techniques [11, 12, 25, 26] such as HPLC with fluorescence, mass spectrometry (MS), and electrochemical detection, particularly the coupling of liquid chromatography (LC) to MS and electrospray ionization (ESI), that is, LC-ESI-MS [25]. The number of compounds/DNA adducts in Figure 2 is expected to grow when more of such studies are carried out. It should be emphasized that SHS also contains all of the common carcinogenic compounds listed in Figure 2, albeit with varying concentrations. Some of the significantly existing chemical carcinogens in SHS are NNN, NNN, B[a]P, benz(a)anthracene, benzene, 1,3-butadiene, 4-ABP, and 2-naphthylamine [5].

In addition to being directly formed by tobacco carcinogens, DNA adducts can be generated through inflammation, particularly by ROS and reactive nitrogen species (RNS). Due to the direct surface exposure, cigarette smoking triggers an inflammatory response in human lung and causes chronic obstructive pulmonary disease (COPD) [85], which has been shown to possess significant abnormalities in inflammatory pathways [86–89]. Smokers are known to have elevated levels of oxidative stress [11, 47], which is increasingly linked to cancer and neurological diseases [90, 91]. Cigarette smoke may induce oxidative stress by several mechanisms [11, 47]: (1) it contains oxidizing compounds and ROS; (2) the ROS-generating redox cycling by quinone-hydroquinone complex as well as PAH quinones and their corresponding catechols; (3) smoking may weaken the antioxidant defense system. The elevated oxidative stress in smokers is accompanied by lipid peroxidation (LPO) [11] which results from reactions of reduced oxygen species with polyunsaturated fatty acids (PUFAs). It is well documented that LPO produces enals, including acrolein, crotonaldehyde, and trans-4-hydroxy-2-nonenal (HNE) [92–94]. These compounds react with DNA to produce etheno adducts as well as propano adducts [92–94]. This explains why chronic inflammation in humans is accomplished by increased levels of such adducts [94]. It should be noted that these adducts are also present in tissues of humans and untreated animals at very low levels as background lesions [92–94].

To understand the chemistry between a carcinogen and DNA bases is instrumental in revealing the molecular mechanism of mutagenicity and carcinogenicity of the carcinogen [18]. A single carcinogen can cause several different types of DNA damage, mainly due to the process of metabolism that can yield several or many reactive metabolites. All the carcinogens listed in Figure 2 can form more than one type of DNA adducts. For example, NNK can form both simple methylated and bulky pyridyloxobutyl (POB) adducts [80, 84]. A single electrophilic carcinogen can form multiple adducts of the same nature by reacting with all four DNA bases. For example, benzene metabolites, hydroquinone (HQ) and para-benzoquinone (p-BQ), form exocyclic adducts on dA, dC, and dG [95–98]. BPDE, by way of another example, can generate both dG and dA adducts. Different carcinogens preferentially react with different sites on the bases [18]. For dG, PAHs predominantly bind to its 2-NH2 group, alkylating agents such as tobacco-specific nitrosamines (TSNAs) mainly react at the N-7 or O6 position, and aromatic amines tend to bind to the 8-carbon

<table>
<thead>
<tr>
<th>Carcinogen</th>
<th>Metabolite</th>
<th>Adduct Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>B[α]P</td>
<td>!(+)-anti-BPDE</td>
<td>!(+)-trans-BPDE-N(^2)-dG</td>
</tr>
<tr>
<td>TSNAs</td>
<td><img src="image" alt="TSNAs Metabolite" /></td>
<td><img src="image" alt="TSNAs Adduct" /></td>
</tr>
<tr>
<td>ABP</td>
<td><img src="image" alt="ABP Metabolite" /></td>
<td><img src="image" alt="dG-C8-ABP Adduct" /></td>
</tr>
<tr>
<td>Benzene</td>
<td><img src="image" alt="Benzene Metabolites" /></td>
<td><img src="image" alt="Benzene Adducts" /></td>
</tr>
<tr>
<td>Acrolein</td>
<td><img src="image" alt="Acrolein Metabolites" /></td>
<td><img src="image" alt="Acrolein Adducts" /></td>
</tr>
<tr>
<td>VC</td>
<td><img src="image" alt="VC Metabolites" /></td>
<td><img src="image" alt="VC Adducts" /></td>
</tr>
</tbody>
</table>

**Figure 2:** Structures of tobacco carcinogens, metabolites, and bulky dG adducts. These are only the partial list of DNA reactive carcinogens in cigarette smoke. The dG adducts listed are representatives of the adducts formed by the corresponding metabolites. Note that the stereochemistry of BPDE-DNA adducts is complex (see text). Acrolein can form two propano dG isomers, \(\gamma\)-HO-PdG (major), and \(\alpha\)-HO-PdG (minor). CAA forms angular \(N^2,3\)-εG, an isomer for \(1,N^2\)-εG.
position [25]. All oxygen and nitrogen sites on DNA bases are actually reactive with alkylating agents in vitro under physiological conditions [18].

Of the DNA adducts formed by tobacco carcinogens, exocyclic adducts have been extensively studied for their chemistry of formation [34, 99]. Bifunctional electrophilic compounds such as HQ and p-BQ, acrolein, and VC metabolites, are all able to form exocyclic adducts [44]. The common sites for forming an exocyclic ring are N-1 and N-6 of dA, N-3 and N-4 of dC, N-1 and N-2 of dG as well as N-2 and N-3 of dG (superscript indicates exocyclic oxygen or nitrogen) [100]. Adducts may be promutagenic if formed at coding sites of the bases, including O-6, N-1, and N-2 of dG, N-1 and N-6 of dA, O-2, N-3, and N-4 of dC, and O-4 and N-3 of T [101]. Structurally, exocyclic adducts are analogous but can differ in ring structure such as size (e.g., 5- versus 6-membered), number (e.g., one ring versus two rings), angularity (e.g., linear versus angular), substituents' nature (e.g., –OH versus –CH2OH), and location (e.g., α-HO-PdG versus γ-HO-PdG) [34, 102]. The structural features of specific adducts may define the specificity and efficiency of their repair, as discussed below, as well as their mutagenicity.

2.2 Repair of Bulky DNA Adducts: An Overview. In the last two decades or so, considerable progress has been made in understanding the specificity, mechanism of action, and in vivo importance of several repair enzymes and pathways. This has been greatly facilitated by major advances in discovery of new enzymes or novel activities, synthesis of site-directed damage-containing oligonucleotides, construction of damage-containing shuttle vectors and viral genomes for in vivo studies, determination of high-resolution structures of repair enzymes and damaged DNA, development of gene mutant models, identification of protein interaction networks, gene analyses such as mutation spectrum mapping and single nucleotide polymorphisms (SNPs), and by the latest studies using omic profiling technology. There are many excellent reviews specifically related to the complete process as well as specific repair pathways that restore DNA to its normal state [62–64, 103–108].

Several major mechanisms have been shown to be involved in the repair of bulky DNA adducts that can be induced by tobacco carcinogens, as discussed below in detail. It is important to determine which adducts are removed efficiently or poorly, as those adducts that persist may cause a greater long-term mutagenic potential. Excision repair, whether it is base (BER) or nucleotide excision (NER), has to be at least a two-step process in which the damage recognition and excision is followed by DNA replication, whereas direct reversal, catalyzed by O-6-alkylguanine DNA alkyltransferase (AGT, also known as MGMT) or AlkB homologs (ABHs), restores the normal base without excision [64]. Multiple repair mechanisms could be involved in the removal of various DNA adducts produced by a single compound. As will be described below, the benzetheno adducts of HQ/p-BQ are substrates for AP endonuclease, and the hydroxyphenyl dG adduct formed by the same compounds is repaired by NER. In some cases, more than one enzyme or mechanism can act on the same adduct, which may serve as backups or operate with different functions in the cell. For example, 3, N4-ethenocytosine (eC) is excised by three different DNA glycosylases and repaired by two different repair pathways, BER and ABH2. Although the mismatch repair (MMR) pathway appears not to be directly implicated yet as significantly as the above pathways in response to the bulky adducts, MutS protein has been shown to bind to the propano dG and M1G adducts [109], suggesting that MutS can bind to exocyclic adducts and may trigger a MMR-mediated response.

Although certain repair data come from research using prokaryotic enzymes, this paper will concentrate on mammalian/human repair enzymes whenever possible. In principle, the analogous enzymes and general mechanisms exist in both prokaryotes and eukaryotes, and such conservation has provided a solid foundation for our understanding of mammalian repair. It should also be pointed out that much of the repair data concerning tobacco carcinogen-derived adducts was not directly obtained from tobacco-related studies, but rather based on reports focusing on chemical carcinogens per se.

2.3 Repair of Specific DNA Adducts by Different Pathways

2.3.1 Nucleotide Excision Repair (NER). NER is the most versatile repair pathway in the cell and the primary mechanism for the removal of chemical carcinogen-induced bulky DNA adducts that significantly distort the DNA helix structure [64, 107, 110, 111]. The molecular mechanism of NER is now well understood. Its pathway in eukaryotes consists of at least 30 gene products [112] and can be reconstituted with purified key proteins in vitro [113, 114]. Mutations in some of these NER genes may lead to xeroderma pigmentosum (XP), a genetic disorder with seven complementation groups (from XPA to XPG), and a higher incidence of skin cancer [64]. The steps in NER consist of sequential assembly of proteins that perform different functions: damage recognition by XPC-HR23B, opening of a denaturation bubble by TFIIH, incision of the damaged strand by XPG and ERCC1-XPF, displacement and excision of the lesion-containing oligonucleotide (24–32 base long), repair synthesis by DNA polymerase δ/ε, and DNA ligation by ligase III. There are two subsets of the pathway: global genomic repair (GGR) and transcription-coupled repair (TCR) that differ in the mode of damage recognition and are regulated by differential cellular mechanisms [105, 107, 110, 111]. GGR is involved in repair of DNA lesions from the transcriptionally silent regions of the genome and the nontranscribed strand of the active genes. GGR probes for DNA lesions that cause structural distortion or chemical alteration. TCR preferentially repairs the distorting lesions on the transcribed strand in active genes in order to avoid a stalled RNA polymerase II. The mammalian NER activity appears to be mostly modulated by posttranslational modifications and by protein-protein interactions.

NER activity can be measured in cell-free extracts by the cleavage of site-directed oligonucleotide containing an adduct or by the extent of DNA repair synthesis in damaged plasmid DNA. Figure 3 shows the structures of important
toxic and mutagenic bulky adducts as NER substrates which are formed by those major carcinogens in cigarette smoke, including PAHs, acrolein, 4-ABP, and benzene. NER also processes endogenous bulky DNA adducts formed by aldehydes from LPO [115]. In addition, intra- and interstrand crosslinks such as those generated by UV light and cisplatin are usually repaired by NER [64]. These crosslinks could also be formed by bifunctional tobacco chemicals such as acrolein and crotonaldehyde [116]. Therefore, NER is a critical repair pathway for protecting against the tobacco carcinogen-induced mutagenesis and carcinogenesis.

(1) Formation and Repair of PAH-Derived DNA Adducts. PAHs are thought to be the major contributors to the etiology of smoke-induced cancers, particularly lung cancer. B[a]P was the first carcinogen to be found in cigarette smoke [117] and has been extensively studied as a surrogate for PAHs. It is present in cigarette smoke at low levels (10–50 ng/cigarette) but is higher than other PAHs [11]. In one well studied metabolic pathway mediated by successive P450/epoxide hydrolase/P450, B[a]P yields active carcinogenic epoxides, mainly B[a]P-7,8-diol-9,10-epoxide (BPDE) [18, 118]. The diol epoxide exists as two diastereoisomers, anti- and syn-BPDE, each of which can be resolved into (+)- and (−)-enantiomers. The isomer (+)-anti-BPDE has the greatest tumorigenicity in vivo [119]. The latter reacts at the N2 of dG to yield either trans- or cis-ring opening of the epoxide ring, forming (+)-trans- and (+)-cis-BPDE-N2-dG adducts [120, 121]. Similarly, (−)-anti-BPDE forms (−)-trans- and (−)-cis-BPDE-N2-dG adducts. Of these four stereoisomeric adducts, (+)-trans-BPDE-N2-dG is the most abundant, which is also the major adduct identified in vivo [122, 123] and was detected in 45% of smokers’ lung [124]. A second path for the activation of B[a]P involves P450-mediated activation to yield free cations [125] that can induce unstable adducts leading to AP sites. A third metabolic pathway is through aldo-keto reductase superfamily-mediated oxidation of B[a]P-7,8-diol to catechol that enters into a redox cycle to form a reactive B[a]P-7,8-quinone (BPQ) [126, 127]. Although a recent study did not support that BPQ forms stable DNA adducts in mice [128], there is evidence that this pathway operates in human lungs leading to ROS-mediated genotoxicity such as causing G to T transversions that inactivate p53 [127, 129]. So far, the relative importance of these pathways in cancer development remains to be determined.

BPDE-DNA adducts are recognized and repaired by E. coli NER complex UvrABC nuclease [130] and human NER [131–133]. Taking advantage of the stereochemistry involved in the formation of these bulky adducts, a number of studies addressed the effects of adduct conformation, base paring, and sequence context on DNA repair. For example, using an in vitro repair system with oligonucleotides containing one of the four BPDE-N2-dG adducts described above Hess et al. showed that the rates of human NER repair of these adducts are dependent on their different stereochemical configurations [131]. The rates of excision were found to vary over 100-fold among these dG adducts, and the cis-adducts of dG are repaired more rapidly than the trans-adducts [131]. It was later found that different conformations of these adducts
are recognized differentially by the NER lesion recognition complex XPC-HR23B, which can be correlated with the relatively low repair of (+)-trans-BPDE-N²-dG [121]. Similar correlations were observed with UvrABC nuclease [134]. To further show the importance of local DNA conformation, the nature of the base opposite a BPDE adduct is found to be critical in modulating the repair rates [38]. As will be discussed below is Section 2.5, the processing of BPDE-DNA adducts by both UvrABC and human NER is also sequence dependent. BPDE forms N⁶-dA adducts in native DNA as well, although relatively inefficiently [135, 136], which exhibit differential conformation and perturbation of DNA duplex than the BPDE-dG adducts [35]. Using cell extracts, human NER activity has been shown for the (+)- or (-)-trans-anti-BPDE-dA adducts [38, 131].

Several early studies showed that repair of BPDE-DNA adducts occurs much faster in the transcribed strand than in the nontranscribed strand of HPRT or p53 genes, indicating that these adducts are subject to TCR [137, 138]. These adducts block human RNA pol II elongation on the transcribed strand, which could be a signal for initiating TCR, also in a stereochemistry- and sequence-dependent manner [139]. A later work shows that common genetic variations in Cockayne syndrome A (CSA) and B (CSB) proteins are associated with NER repair capacity of BPDE-induced DNA damage in smokers [140]. Mutations with CSA and CSB result in Cockayne syndrome with impaired TCR [64]. Taken together, this strand preference in repair may contribute to the mutational property of the human lung cancer p53 gene in response to BPDE exposure: repair of BPDE adducts along the nontranscribed strand of p53 is consistently slower than repair in the transcribed strand, and repair at the major damage hotspots in the nontranscribed strand is 2–4 times slower than repair at other damage sites [138].

Dibenzo[a,l]pyrene (DB[a,l]P) is another PAH that has been found to be present in tobacco smoke particulates and is the most potent carcinogen of the PAHs tested to date in rodent systems. Similar to the B[a]P-derived adducts, the bulky adducts formed by (±)-anti-DBPDE possess different structures and adopt different conformations [141]. They are differentially repaired by NER in human cells with some being poorly removed, as shown by a recent study [39]. The repair of DBPDE-DNA adducts by NER has been shown to be slower than the repair of BPDE-DNA adducts [142]. In general, the poor repair by NER of DBPDE-DNA adducts, at least some of them, may account for the high carcinogenicity of the parent compound.

(2) Formation and Repair of DNA Adducts of Aromatic Amine 4-ABP. Chemicals in this class such as 4-ABP bind to DNA bases mainly at C-8 position. Adducts can also be formed at N²- and O⁶- of dG and N⁶- of da [30]. 4-ABP has been established as a major human bladder carcinogen [143]. 4-ABP forms DNA adducts after N-hydroxylation by P450 to the mutagenic metabolite N-hydroxy-4-aminobiphenyl (N-OH-4-ABP). N-(deoxyguanosin-8-yl)-4-aminobiphenyl (dG-C8-ABP) (Figure 2) is the major adduct of 4-ABP, and the minor adduct is N-(deoxyadenosin-8-yl)-4-aminobiphenyl (dA-C8-ABP) [144]. The major adduct has been detected in the human cells after exposure to N-OH-4-ABP [145]. This adduct was also identified from DNA of the bladder biopsy samples from smokers and is quantitatively related to smoking status [146]. dG-C8-ABP adducts have been identified from human bladder cancer tissues [147, 148]. Moreover, higher levels of DNA adducts correlated with more invasive tumors (higher tumor grades) [147]. The unique binding pattern of 4-ABP in the p53 gene, that is, the p53 mutational hotspots in bladder cancer at several codons are also the preferential sites for 4-ABP adduct formation, links 4-ABP to the etiology of bladder cancer [149].

Although the detailed molecular mechanism of the repair of 4-ABP-DNA adducts is not clear, DNA fragments modified with N-OH-4-ABP were shown to be incised by E. coli NER complex, the UvrABC nuclease [150]. An early study investigated the rate of disappearance of dG-C8-ABP in human transitional cell carcinomas of the bladder and showed that the majority of the adducts can be removed within 48 hours after treatment with 4-ABP [151]. Another study showed that dG-ABP was repaired rapidly while dA-ABP persisted in human uroepithelial cells [144]. There is evidence of the human NER pathway involvement in the repair of these adducts, as the host cell reactivation (HCR) assays performed in NER-deficient cells showed reduced repair of DNA lesions from plasmid treated with 4-ABP [152]. In addition, it was shown that loss of function of the p53 gene in human bladder epithelial cancer cells reduces the efficiency of repair of dG-C8-ABP, suggesting that p53 may modulate its repair in target cells [151, 153]. The relationship between deficient DNA repair of 4-ABP-DNA adducts and increased bladder cancer risk was supported by the findings that such repair capacity was significantly lower in bladder cancer cases than in controls, and ever-smokers with low DNA repair capacity exhibited a 6-fold increased risk compared with never smokers with normal repair capacity [152].

(3) Formation and Repair of Propano Adducts of α,β-Unsaturated Aldehydes (Enals). Enals can arise from both cigarette smoking and endogenous LPO [154]. Cigarette smoke contains relatively high concentrations of acrolein and crotonaldehyde. HNE is a unique product of ω-6 of PUFAs [92]. Acrolein is the simplest enal and is a model chemical for this class of carcinogens. Acrolein is one of the most abundant compounds in MSS (60–100 μg/cigarette) and is also present in SSS at high concentrations [5]. It is highly reactive without metabolic activation. Acrolein forms exocyclic adducts on DNA bases, predominantly 1,N²-dG adducts [155, 156]. The principal adduct is γ-hydroxypropano-2'-deoxyguanosine (γ-OH-PdG) that exists as a mixture of C8-OH epimers (Figure 3) [157], and the other adduct is a-hydroxypropano-2'-deoxyguanosine (α-OH-PdG). The mutagenicity of α-OH-PdG is well established, while the mutagenicity of γ-OH-PdG has been reported with mixed results [158, 159]. Both adducts were recently found in human lungs using LC-ESI-MS/MS [160]. Acrolein-DNA adducts have been detected in the tissues of cigarette smokers with significantly higher levels than those of nonsmokers.
[161, 162]. Likewise, both crotonaldehyde and HNE also form stereoisomeric propano dG adducts, and increased crotonaldehyde-dG adduct levels were observed in smokers [163]. HNE adducts have also been detected in rodent and human tissues [163, 164]. The mutagenic potential of these dG adducts were recently summarized by Minko et al. [157].

γ-HO-PdG is a substrate for E. coli UvrABC nuclease [43, 165]. In humans, NER of this adduct has been reported [158]. There is also biochemical evidence that the HNE-DNA adducts are repaired by UvrABC [92] and mammalian NER in cell-free extracts [92, 166]. A recent study revealed that NER and recombination, but not MMR, are involved in repair of HNE-treated phage DNA replicating in E. coli [167]. Moreover, the repair rates were shown to be affected by the adduct stereochemistry when four HNE-dG isomers were tested [166]. Interestingly, although BER is able to excise the 1,2-epi-dG adduct, it appears to have no in vitro activity towards the structurally analogous adducts γ-OH-PdG, α-OH-PdG, and PdG [168] and no in vivo protective role in a mutagenesis assay based on the vector containing a γ-OH-PdG [169].

Recent findings also pointed to the role of highly accurate TLS in protecting cells from the potential genotoxicity of the acrolein-DNA adducts [158, 165]. Previous in vivo site-specific mutagenicity studies have shown an efficient error-free bypass of the γ-HO-PdG, and PdG [165, 170]. Work from E. coli indicated that NER, recombination repair, and error-free TLS are all involved in the cellular response to this major acrolein-dG adduct [165].

(4) Formation and Repair of In Vivo HQ-/p-BQ-Induced Hydroxyphenyl Adducts. Benzene is a well-established human carcinogen and is associated with an increased risk of leukemia [171]. It is a significant volatile compound in the vapor phase (12–48 μg/cigarette) [5]. In one major metabolic pathway, benzene is converted by P450 to benzene oxide which is further converted to phenol, catechol (CAT) and various derivatives [20, 172] (Figure 4). One biologically important stable metabolite is p-BQ, an oxidation product of HQ [20, 172]. A number of bulky DNA adducts have been detected in vitro and in vivo when exposed to HQ or p-BQ [95–98, 173–177]. Reaction of p-BQ or HQ with DNA in vitro has been shown to result in the formation of two ring exocyclic benzetheno adducts on dC, dA, and dG [95–98]. These adducts are highly mutagenic as tested in vitro with human pols involved in TLS and in yeast by site-directed mutagenesis [178]. The Bodell group has found that the DNA adducts formed in animals after benzene administration are identical to those produced in cells treated with HQ, suggesting that HQ is the main benzene metabolite causing adduct formation in vivo [177]. By 32P-postlabeling, the principal DNA adduct caused by HQ or p-BQ corresponds to N2-(4-hydroxyphenyl)-2′-dG(N2-4-HOPh-dG) [173, 177]. Exocyclic adducts were also detected in vitro from reactions of trans, trans-muconaldehyde (MUC), a reactive ring-opened diene dialdehyde formed from a minor metabolic route [179, 180]. It is still unclear as to what role the above covalent DNA adducts may play in benzene-induced carcinogenesis, since benzene also induces other types of DNA damage as well as chromosomal damage. For example, oxidized bases such as 8-oxoG can be caused through the quinone/hydroquinone redox cycling [11] (also see Section 2.1). Benzene also generates DNA strand breaks [181, 182] through direct attack by ROS or unstable DNA adducts. As shown in Figure 4, catechol quinones can react with DNA by 1,4-Michael addition to yield major N3A and N7G adducts which are unstable and generate AP sites [183].

We recently reported the repair of N2-4-HOPh-dG E. coli UvrABC nuclease [184]. The specificity of such repair was also compared with those of DNA glycosylases and damage-specific endonucleases of E. coli both of which were found to have no detectable activity toward this adduct. We also showed that p-BQ-modified plasmid is efficiently cleaved by UvrABC, indicating the involvement of NER in repair of benzene-derived DNA damage [184]. The role of NER in the repair of HQ/p-BQ-induced DNA damage was also suggested in another mutagenesis study using HQ- or p-BQ-treated plasmid containing the supF reporter gene in NER-deficient (XPA) human cells [185]. Note that HQ- and p-BQ-derived exocyclic adducts are repaired by a different mechanism called nucleotide incision repair (NIR), as discussed below. In general, although benzene metabolites show relatively low DNA binding activity in vitro, their induced DNA damage and repair seem to be complex [186].

2.3.2. Base Excision Repair (BER). BER is the primary repair mechanism for the removal of small DNA lesions such as alkylated, oxidized, and deaminated bases from endogenous sources or environmental carcinogens [64, 187–191]. The steps of the BER pathway have been well characterized [64]: it is initiated by a damage-specific DNA glycosylase that recognizes a modified base and cleaves the N-glycosyl bond between the base and the sugar moiety. Glycosylases can be divided into monofunctional, for example, alkylpurine-DNA glycosylase (AAG, also MPG, APNG, and ANPG) and thymine-DNA glycosylase (TDG), and bifunctional (with an AP lyase activity), for example, OGG1, endonuclease III homolog 1 (NTH1), and endonuclease VIII-like glycosylases (NEILs). Each DNA glycosylase has its unique specificity, but overlapping activities are common among various DNA glycosylases that may have different structures and/or catalytic mechanisms [191]. After glycosylase, the resulting AP site is processed by 5′ AP endonuclease, AP lyase, and DNA polymerase activities to cleave the AP site, trim strand break intermediates, and catalyze repair synthesis. A DNA ligase finally completes the process by sealing the remaining nick. The basic BER mechanism described above is complicated by the identification of subpathways (i.e., the short-patch and long-patch BER) in mammalian systems [192]. There is also a network of protein-protein interactions involving numerous proteins inside and outside of BER, which is thought to play a key role in coordination of BER components as well as in regulation of cellular BER functions [193–195]. Evidence has emerged to support that BER deficiency is an important contributing factor of cancer susceptibility, as shown in both animal models and human studies [196].
Figure 4: Multiple benzene metabolites, different types of DNA lesions, and proposed biological effects. Benzene has a complex metabolism and the listed metabolites are not a complete list. The adducts formed include those identified in vitro and in vivo, stable and unstable, bulky and oxidized adducts. DSBs are one of the most severe DNA lesions caused directly and indirectly by benzene metabolites. Repair of benzene-DNA adducts may include multiple mechanisms such as BER, NER, and NIR. Only those adducts that finally escape all the defense mechanisms such as repair, or are misrepaired, may lead to mutations. Persistence or coexistence of different types of lesions could form a broad-based attack on the genomic stability. It is also known that a number of benzene metabolites can inhibit topoisomerase II (topo II) activity, which may represent a potential mechanism for benzene's clastogenic effects [326].

Surveying the activities of known glycosylases indicates that many of them are able to excise tobacco carcinogen-induced DNA adducts, including the common alkylated and oxidized bases and some exocyclic adducts. The tobacco carcinogen-derived exocyclic DNA adducts listed in Figure 5 are known substrates for respective glycosylases as described below. It should be noted that a crucial role of BER is to repair an AP site which is mutagenic because of its noncoding nature [197, 198]. As stated above, certain tobacco carcinogens generate unstable DNA adducts that are an important source of the AP sites in the genome.

(1) Formation and Repair of Etheno DNA Adducts. Etheno (ε) adducts are the most extensively studied exocyclic adducts [34, 199] which are formed by the attack of bifunctional aldehydes or epoxides at a nitrogen of the base, followed
Adducts may be responsible for double-stranded [205, 206] and single-stranded DNA [207]. Cigarette smoke is a significant source for these adducts as shown by the urinary levels of εC [200] and 1,N2-εG [201] in smokers. These adducts could be formed by VC in cigarette smoke (5–30 ng/cigarette) as well as LPO products [94]. VC is processed by P450 yield unstable chloroethylene oxide (CEO), which quickly converts to chloroacetaldehyde (CAA) [202, 203]. Both CEO and CAA can form ε-adducts. In experimental animals and in humans exposed to VC, liver angiosarcomas are the most common type of tumors. Both CEO and CAA can form ε-adducts. NER is not involved in its repair [213].

Excision of εC. We and Saparbaev et al. independently found that the εC activity resides in both human TDG and E. coli Mug proteins [214, 215]. The main biological role of TDG appears to remove thymine from a T:G mismatch resulting from the deamination of 5-methylcytosine (5-mC) in a CpG site, which could be involved in active DNA demethylation when in combination with a deaminase that converts 5-mC to T leading to a T:G mismatch [216]. In vitro assays have shown that the activity of TDG is most efficient when T:G or εC:G is in a CpG sequence context [217, 218]. Recently, the crystal structure of TDG (the catalytic domain) complexed with DNA containing an AP site was reported [219]. Other studies have also shown low excision of εC by human single-strand-selective monofunctional uracil-DNA glycosylase (SMUG1) [220] and methyl-CpG binding domain protein (MBD4 or MED1) [221].

Excision of εG adducts. In rodents, N2,3-εG represents the predominant ε-adduct and is readily induced in hepatic nonparenchymal cells by VC, the target cells for this compound [222]. There is a correlation between the levels of this adduct and the incidence of VC-induced angiosarcoma [222]. E. coli AlkA excises N2,3-εG from CAA-treated DNA [223]. Both in vitro [224] and animal studies [222] showed that the human removal of N2,3-εG is slow. It was also shown that repair capacity would be different in various

Studies on repair of exocyclic adducts have largely been focused on BER, except for propano-dG adducts that are repaired by NER. ε-Adducts are repaired by BER, initiated mainly by two human DNA glycosylases: AAG and TDG [34]. In E. coli, they are repaired by functional homologs of AAG and TDG: AlkA (m3A-DNA glycosylase II) and mismatch uracil-DNA glycosylase (Mug), respectively [34].

Excision of εA. Human AAG excises this adduct from double-stranded [205, 206] and single-stranded DNA [207]. The crystal structure of human AAG bound to DNA containing an εA has been solved [208, 209]. AAG is also the major activity against εA in vivo, as shown in Aag−/− knockout mice [210, 211]. Increased mutations were observed in the hprt gene, and levels of εA were significantly higher and persisted longer in DNA from Aag−/− mice than those from wild-type mice when treated with vinyl carbamate [210, 211]. Moreover εA and εC accumulated to higher levels in Aag−/− mice following stimulation of colonic inflammation, indicating that the repair of such adducts formed by LPO is important for protection against chronic inflammation-induced ROS and carcinogenesis [212]. As described below, the AlkB/ABH pathway is also involved in the repair of εA and εC adducts. NER is not involved in its repair [213].

Figure 5: Structures of smoking-induced bulky DNA adducts that are substrates for BER.
cell types in liver in that the expression of AAG mRNA was induced in the hepatocytes of rat exposed to VC, while the nonparenchymal cells had only 20% of the AAG mRNA of hepatocytes, indicating that the target cells for VC had much lower expression of this glycosylase [225]. It should be noted that N²,3-ε-G is also an endogenous adduct arising from LPO [225]. 1,N²-ε-G, an isomer of N²,3-ε-G, is a substrate for both E. coli Mug and human AAG as tested in vitro [207, 226].

Excision of hydroxymethyl ε-adducts. We recently studied in vitro repair of two exocyclic adducts formed by acrolein metabolite glycidaldehyde (GDA), a potent mutagen and animal carcinogen. 7-(Hydroxymethyl)-1,N²-ethenoadenine (7-hm-εA), the main adduct, can be found in skin cells of mice treated typically with GDA [227]. Minor adducts with guanosine and deoxyguanosine were also found [228]. The 8-hm-εC adduct has only been identified in vitro [229]. These ε analogs are expected to be as promutagenic as the corresponding ε-adducts, and 8-hm-εC has been shown to be miscible when tested with mammalian DNA pols [230]. Biochemical assays have shown that 7-hm-εA is primarily repaired by AAG [231]. While 8-hm-εC is excised by E. coli Mug and human TDG [232], these excision activities were from half-to-a few-fold lower than the corresponding ε activities, which could be attributed to the extra −CH₂OH group on the ε-ring [231].

2.3.3. Direct Reversal of DNA Damage

(1) Formation and Repair of Pyridyloxobutyl (POB)-DNA Adducts of TSNA. Tobacco-specific nitrosamines (TSNAs) are exclusively found in cigarette smoke and are formed through N-nitrosation of nicotine during tobacco curing and processing [233]. Common TSNAs found in cigarette smoke particles include NNK, NNN, and N-nitrosoanatabine (NAB). NNK has a strong affinity for the lung and is a systematic lung carcinogen [234]. NNK is among the most potent lung cancer carcinogens in tobacco smoke [27]. TSNAs require activation by the P450 system [80]. In one pathway, NNK is activated to form mutagenic O⁶-mG [235]. In addition, NNK- and NNN-generated reactive intermediates form bulky POB-DNA adducts, including the 7- and O⁶-positions of dG and the O²-position of dC and T [84]. Four of them have been recently characterized and detected in NNK- or NNN-treated animals [236]. One of them, O⁶-[4-(3-pyridyl)-4-oxobut-1-yl]-2'-dG (O⁶-POB-dG) [237–239] (Figure 2), will be discussed here, which has been shown to be mutagenic in both E. coli and human cells using a site-specific mutagenesis assay and is considered a critical lesion in NNK/NNN carcinogenesis [240]. Higher levels of adducts are found in lung and tracheobronchial tissues of smokers than in nonsmokers, by the detection of 4-hydroxy-1-(3-pyridyl)-1-butanone (HPB), a product from acid hydrolysis of the POB-DNA adducts [80]. Most recently, 1-(N-methyl-N-nitrosamino)-1-(3-pyridyl)-4-butanal (NNA) was identified from thirdhand smoke (THS) as the major product resulting from the reaction of nicotine with nitrous acid (HONO), along with NNK and NNN [241] (see Section 3.2).

Both O⁶-mG and O⁶-POB-dG adducts have been shown to be substrates for AGT [106, 237, 242]. However, the repair of the bulky POB adducts has been much less studied compared to O⁶-mG. AGT primarily repairs O⁶-alkylguanine adducts and protects against mutagenicity of respective alkylating agents [106]. It is not an enzyme but an alkyl group acceptor. Repair occurs by transfer of the alkyl group at the O⁶ position of G to a cysteine residue at its active site of the protein, which results in a protein conformational change that signals for its degradation [243]. AGT reaction is stoichiometric with O⁶-mG acting as a suicide substrate; therefore, the cellular repair capacity is limited by the constitutive levels of AGT that can be also depleted under overdose of alkylating agents [106]. O⁶-POB-dG has been shown to be repaired by AGT both in vitro [237] and in vivo [244]. This adduct is efficiently repaired by mammalian AGTs but poorly repaired by bacterial counterparts, AdaC and OgT [235, 244]. Since both O⁶-mG and O⁶-POB-dG may have implications in NNK-induced carcinogenesis, the relative repair rates of these two adducts by AGT should be an important factor in determining the levels and biological importance of these two lesions. Studies by Mijal et al. [235] demonstrated that human AGT showed an ~2-fold preference for the removal of O⁶-mG over O⁶-POB-dG, rodent AGTs exhibited the same rate, and the bacterial proteins reacted poorly with O⁶-POB-dG. These data indicate the high importance of protein structure with respect to substrate efficiency. In conclusion, AGT is expected to be critical in the repair of O⁶-alkylguanine adducts formed by tobacco-derived N-nitrosamines. It should be noted that cytotoxicity and mutagenesis studies suggest an NER involvement in the removal of NNK-derived DNA damage [236, 245]. A very weak but time-dependent in vitro NER activity was also detected using oligonucleotide containing an O⁶-POB-dG and reconstituted human excision nuclease [236].

(2) Repair of Etheno Adducts by AlkB Homologs. It was reported in 2005 that E. coli AlkB protein and its human homolog, ABH3, repair εA and εC in vitro [246, 247]. Later, ABH2 was shown to exhibit robust activity for εA and is the principal dioxygenase for removal of εA in vivo as shown by Abh2+/− mouse studies [248]. Further experiments showed that ABH2, but not ABH3, is able to complement the E. coli alkb mutant which is defective in the repair of ε-adducts [248]. AlkB is a member of the superfamily of iron-/α-ketoglutarate-dependent dioxygenases. Using bioinformatics, eight mammalian homologs of AlkB, ABH1 to ABH8, have been identified [249]. The direct reversal mechanism for their action involves a unique iron-mediated reaction with cofactor α-ketoglutarate that could epoxidize the exocyclic double bond of the ε-adducts [246, 247]. The epoxide generated can be hydrolyzed to form the lesion-free base and glyoxal. In addition to εA and εC, these proteins also repair other methylated/ethylated bases [250, 251]. They can act on both single- and double-stranded DNA substrates and may play different/complementary roles to the glycosylases as mentioned above. In general, single-strand
specificity suggests repair of lesions in single-stranded DNA regions that are transiently generated during replication and transcription.

Given that DNA glycosylases also act on ε-adducts (see Section 2.3.2(1)), at least two repair pathways may act on these adducts in vivo. This may explain why the incidences of carcinomas were similar between wild-type and Aag-knockout mice treated with vinyl carbamate [252]. Genetic studies using AlkA-proficient and -deficient cells show that AlkB is important for counteracting the mutagenicity of the ε-adducts [246]. A study comparing the repair efficiency of AlkB versus AlkA in E. coli shows that AlkA seems to be the more important enzyme in response to exposure to CAA [246]. Similar data were obtained from knockout mice [248], which showed that the Abh2 activity is not sufficient for the removal of spontaneously produced εA adducts in Aag^−/− mouse liver, whereas mouse Aag activity is sufficient to repair spontaneously produced εA lesions in Abh2^−/− mouse liver. These results suggest that both AAG and ABH2/3 proteins can play a role in the cellular response to the exposure of tobacco carcinogens that generate these ε-adducts.

2.3.4. Nucleotide Incision Repair (NIR). The exocyclic benzetheno p-BQ adducts are bulkier than the ε-adducts, with an additional five-membered ring and a hydroxy group. As might be expected, such bulky adducts hinder replication in vitro and in vivo and cause frameshift deletions and base mispairing [178]. In the past years, we have studied repair of three major in vitro adducts formed by HQ and p-BQ (designated as pBQ adducts), 1,N^6-pBQ-dA, 3,N^4-pBQ-dC, and 1,N^2-pBQ-dG (Figure 6). Our initial study discovered that these adducts are recognized by the major human AP endonuclease (APE1, also known as HAP1, APEX, and Ref-1) as well as E. coli exonuclease III and endonuclease IV [253, 254]. Mechanistic studies showed that human APE1 hydrolyzes the phosphodiester bond 5' next to the adduct, leaving the p-BQ derivative on the 5'-terminal of the 3' fragment as a “dangling base” [253, 255, 256]. This mode of incision was later named nucleotide incision pathway [257, 258], which also acts on several oxidized DNA bases. While the AP site is the preferred substrate for APE1, cleavage of the pBQ-dC adduct requires the same catalytic center as the AP site as shown from mutant APE1 proteins [256]. Molecular dynamics simulations [36] suggest that APE1 utilizes a reaction mechanism for phosphodiester bond cleavage of DNA containing pBQ-dC similar to that reported for the AP site [259]. Given that these adducts have not been reported to be present in vivo, the biological role of these adducts as well their repair by NIR awaits further investigation.

2.3.5. Summary. The repair mechanisms for representative bulky DNA adducts discussed above are summarized in Figure 7. It should be emphasized that most of the repair studies in the past have applied a single compound for modification or exposure, and results from such studies cannot be simply extrapolated to real exposures. However, the information on repair specificity and efficiency from such studies provides the framework for further evaluation of potential relationships among repair deficiencies, carcinogen mutagenicity, and human susceptibility to cigarette smoke. Also, as stated above, tobacco carcinogens are able to generate other specific DNA lesions in addition to bulky DNA adducts. The importance of bulky DNA adducts relative to other types of DNA lesions needs further investigation, and in any case, a combined action from different types of DNA/chromosomal damage, as demonstrated in Figure 4 for benzene’s biological effects, is expected to be the basis of genotoxicity conferred by many tobacco carcinogens.

2.4. Molecular Structure of DNA Adducts and DNA Repair. A crucial question in repair is how repair proteins recognize DNA adducts, since repair specificity has both biochemical and biological implications. A related question is “what are the factors responsible for good and poor repair?” To date, we have learned a great deal with regard to what structural factors of adducts and repair proteins determine the specificity and rate of repair, mainly based on biochemical data atomic resolution structures of adducted DNA and repair proteins [64, 102, 260, 261].

The study of how a DNA adduct affects the structure of DNA and how it interacts with its repair protein is essential in order to develop a theory for both why only some adducts are repaired and the specific mode of repair. When surveying the multiplicity of DNA substrates for NER, the primary repair
pathway for many bulky DNA adducts, one basic question is how NER recognizes these chemically diverse substrates through the common structural features of the recognition unit XPC-HR23B. Moreover, what is the basis for the fact that the repair rates of different bulky adducts by NER can vary by several orders of magnitude [111, 262]? It should be recognized that, as described above, the complexity of these adducts is enormous, including those adducts that are formed by compounds with stereochromically properties (e.g., BPDE, acrolein, and HNE). Recognition and repair rates of such stereoisomeric adducts generally reflect or depend on adduct conformation. In the case of BPDE-N2-dG adducts, both their removal by human NER [131] and patterns of helix opening by XPC-HR23B [121] are stereochemistry dependent. Recently, crystal structures of NER proteins, that is, bacterial UvrB [263] and yeast Rad4 (human XPC homolog) [264], complexed with damaged DNA, have been described, both of which suggest a mode of action involving strand separation and nucleotide flipping for the bulky DNA lesions.

As for BER, a number of high-resolution structures of glycosylases, including those complexed with DNA lesions, have provided a valuable insight into adduct selection as well as mechanisms of base flipping and catalysis [260, 261, 265]. In addition, molecular modeling studies such as molecular dynamics simulations have offered additional information about the structural features of DNA substrates and their interactions with repair enzymes [36, 266]. Some common features of recognition by DNA glycosylases can be summarized based on the reported structural studies: (1) adduct shape, hydrogen-bonding potential, and electric charge distribution are key for recognition; (2) base unstacking is present at the lesion site; (3) the target nucleotide has to be flipped out of the DNA duplex and fit in the active site of a glycosylase [261]. Similarly, AP endonucleases flip an AP site out into their active sites, as shown by the co-crystal structures [259, 267]. Early on, we had some puzzling questions for this pathway and its glycosylases. For example, why are structurally related adducts repaired differentially? for instance, BER excises the 5-membered unsaturated exocyclic εG but not the 6-membered propano-dG [34, 157]. Also, why can lesions with largely diverse structures be processed by the same protein, as seen by human APE1 acting on both AP site and pBQ-dC [253, 255, 256]? In general, the key to the specificity of recognition has been known to be not only the primary adduct structure, but the localized effect of each adduct on DNA structure as well as on thermodynamics, and, moreover, the structure and function of the repair protein. Since most repair enzymes act on adducts in double-stranded DNA, each adduct may cause differing distortion and flexibility as a result of factors such as being in the major or minor groove, syn or anti, planar or angular, adjacent base pair tilting, propeller twist, and helical twist [102]. Ultimately, it is hoped that we can predict repair specificity/efficiency as well as identify structural hallmarks of mutagenic lesions in the genome, using appropriate computational and/or screening approaches. This can only be done after a large amount of structural and theoretical data have been acquired, which relate adduct structural features with outcomes of repair and mutagenesis.

2.5. Nucleic Acid Sequence, Mutational Hotspots, and DNA Repair. The sequence-dependent repair infers that local DNA structures adjacent to an adduced nucleotide are important determinants of repair specificity and efficiency. The study of the role of sequence in base modification was started mainly in the 1980s in terms of sequence selectivity for mutational events which were generally induced by environmental agents [268]. Extensive work on relating sequence-dependent adduct formation and mutation has
been done using chemical modification (e.g., BPDE) of genomic DNA, followed by determination of the mutation pattern and spectra [268]. These data were among the first used to substantiate the concept of “hotspots”; that is, for a given reagent, there was site specificity for DNA modification. An example is that in vivo, only the second guanine residue in codon 12 (GGA) of the \(H\)-\(ras\) \(A\) : \(T\) mutations [269], which is consistent with other studies demonstrating the sequence-dependent formation of \(O^\beta\)-alkylguanine in DNA [270, 271].

In addition to preferential DNA adduct formation at specific sites, poor repair is another major determinant of mutational hotspots [30]. Many studies have highlighted the importance of sequence context in influencing the rate and extent of repair [268]. Examples among the tobacco carcinogen-induced bulky DNA adducts are BPDE-DNA adducts [134, 138, 272–274], POB-DNA adducts [242], eA [275], and eC [232], whose repair efficiencies could vary over manyfold when present in different neighbor sequences. These examples involve repair systems including at least NER, BER, NIR, and AGT [268]. A review by Singer and Hang commented on many enzymes in these pathways with regard to the role of adduct, neighbor bases, and repair rate [268]. Also, Donigan and Sweasy recently summarized the known sequence context-specific activities of several glycosylases and polymerase \(\beta\) in BER [276]. It can be concluded now that sequence-dependent repair tends to be predominant, instead of being a random phenomenon.

Mechanistically, the structural factors that modulate sequence-dependent repair have been well studied with certain adducts such as BPDE-DNA adducts [274, 277]. An NMR study using a sequence containing a natural GG mutational hotspot showed that the presence of the major BPDE-derived dG adduct at one of the two neighboring G positions resulted in significantly different local structural distortions, especially bending or kinking at the adduct position and destabilization of Watson-Crick hydrogen bonding of the flanking base pairs [35]. Using the same sequences, Kropachev et al. demonstrated that such hydrogen bonding destabilization elicits the most significant NER response, while the flexible kink is less important in such interaction [274]. It is also apparent that the chemical nature of a DNA adduct itself can affect the effect of neighbor sequences. For example, the repair of a POB adduct by AGT is more strongly influenced by its neighbor bases than that of the smaller methylated base substrate [242].

Regardless of the nature of the specific structural differences discussed above, current evidence also supports that thermodynamic stability of lesion-containing oligonucleotides plays an important role in sequence-dependent repair [268, 274, 278]. In the case of BPDE-DNA adducts, the degree of local thermodynamic destabilization was related to the degree of recognition of duplex sequences containing a bulky adduct by the NER machinery [262, 274, 279]. Both of our studies on sequence-dependent repair of eA [275] and AP site [280] also demonstrated a role of thermodynamic properties in influencing double-strandedness of the substrates and repair their efficiency.

Recent studies have discovered a strong coincidence of mutational hotspots in human lung cancers and the sites of preferential binding of BPDE [42, 281] and acrolein [43] in the \(p53\) gene. The overall prevalence of \(p53\) mutations is higher in cigarette smokers than in nonsmokers [51, 52]. A number of hotspots have been found along \(p53\) in lung cancer which are generally G to T transversions [30]. It has been reported that in PAH- or acrolein-treated cells the same positions of their mutation hotspots are also major hotspots for mutations observed in human lung cancers from smokers, strongly suggesting a role of DNA adducts in etiology of these cancers [30]. As discussed above, poor/slow repair of DNA adducts at these sites may be a major factor for their occurrence and persistence at these mutational hotspots. To further support this nation several groups in the mid-1990s examined the in vivo repair rates along a gene fragment using the ligation-mediated PCR technique [282]. As for BPDE adducts in human \(HPRT\) and \(p53\) genes, Wei et al. [272] found that repair rates can differ markedly from site to site over a time period, as measured by the percentage of adduct remaining. Moreover, very slow repair was observed at certain positions that are frequently mutated after BPDE treatment [272]. These studies clearly indicate a correlation between inefficient DNA repair and the occurrence of mutation hotspots. Finally, in addition to site-specific preferential formation of DNA adducts and sequence context of DNA repair, the biological selection of induced mutations is also considered important for the hotspot phenomenon, which gives cells with specific mutation(s) a growth advantage and results in dominant mutations in cancer cells [283, 284].

### 2.6. Interindividual Variations in Response to Tobacco Carcinogens and Cancer Risk

It has long been recognized that only a small percent of cigarette smokers develop cancer, for example, \( \sim 11\%–24\% \) of smokers develop lung cancer [285], which suggested that interindividual variability in key cellular processes is crucial in response to tobacco carcinogens. Many molecular and epidemiological studies have been revealing a multifactorial nature of such variability [286–289]. One top aim of the target cancer prevention programs is to identify smokers and nonsmokers exposed to SHS with higher susceptibility. The interindividual differences discussed below will be focused on those genotypes and phenotypes involving DNA repair capacity (DRC) in relation to the mutagenicity and carcinogenicity of tobacco-induced bulky adducts.

Considerable progress has been made towards a better understanding of the association between individual tobacco carcinogens and tumor development in specific organs/tissues [12], as exemplified by the following cases: NNK and PAH are potent lung carcinogens; aromatic amines such as 4-ABP are the main cause of bladder cancer; benzene induces acute myelogenous leukemia (AML). Whether reduced or deficient DRC for tobacco carcinogen-derived DNA damage is associated with somatic mutation and susceptibility to cancer has been a subject of investigation. A commonly used approach is to measure the repair or levels of specific DNA adducts which serve as
an intermediate end-point of genotoxicity [14]. Decreased repair activities for bulky DNA adducts have been observed in cells/tissues of cancer patients. For example, epidemiologic studies using the HCR assay showed that low cellular DRC in response to BPDE-induced DNA damage is associated with increased risk of lung, head, and neck cancers [65, 286, 290]. Biochemical studies also showed that reduced repair of εA and εC adducts was present in lung adenocarcinomas [291]. In principle, the phenotypes of DNA repair must be characterized for mutagenic adducts and any newly identified adducts in smokers and nonsmokers exposed to SHS.

Deficient repair towards tobacco smoking-related DNA adducts may occur under various mechanisms. A common one is polymorphisms in relevant DNA repair genes, such as those identified in the NER, BER, and AGT pathways [289, 292, 293]. It has been shown that cellular levels of DNA adducts such as those arising from BPDE exposure can be affected in some of those genetic variants [132, 294]. In the last decade, although mixed or discrepant results have been reported, positive results on numerous polymorphisms in DNA repair genes, along with those in metabolic genes, have been revealed and are continuously being found in the context of cigarette smoking and cancer [22, 295, 296]. It seems that polymorphisms in tobacco metabolism and/or repair should lead to differences in both local carcinogen levels and/or DNA adduct levels in vivo [14]. Another mechanism that can cause the loss of DNA repair capacity is LOH; for instance, LOH at the human 8-oxoG-DNA glycosylase (OGG1) gene locus is a frequent event in lung cancer, which would increase the mutational load from 8-oxoG due to ROS in smokers [297].

DRC could also be influenced by nongenetic factors that cause a phenotypic reduction/ablation of repair activities. Examples of such factors include those that are disease related, for example,NER deficiency in XP and Cockayne syndrome patients [64, 140], and those that are repair-inhibition based. For the latter, carcinogen-mediated effects on proteins play an important role. Two types of tobacco chemicals can cause protein damage and inhibit repair. There are at least 30 metals in cigarette smoke, including arsenic, cadmium, nickel, and chromium, which have been shown to inhibit various DNA repair enzymes and pathways [298–302]. For instance, both arsenic and nickel compounds interfere with the repair of BPDE-DNA adducts [298, 299]. Cadmium and chromium (VI) are also known as effective DNA repair inhibitors and play a similar role in influencing adduct levels in smokers [303, 304]. Therefore, coexposure to these heavy metals by smokers may enhance the mutagenic potential of genotoxic tobacco carcinogens [305]. In addition to metals, certain tobacco carcinogens themselves have been found to inhibit DNA repair, as exemplified by acrolein inhibiting NER repair of BPDE-DNA adducts [43]. Similar to metals, the coexistence of these chemicals in cigarette smoke is also expected to lead to more persistent or severe DNA damage as a result of suppressed DNA repair.

In addition to tissue differences in bioactivation of tobacco carcinogens [306, 307], recent studies in rodents suggest that the DRC in a given organ/tissue/cell type may play a role in organoselectivity of tobacco carcinogens. For example, both AGT [308] and NER [309] activities are related to the interorgan differences in response to NNK treatment, which could be an important factor in determining organ-specific susceptibility to NNK-induced carcinogenesis. Another example is benzene which targets the bone marrow. DNA adduct levels in the bone marrow of benzene-treated mice are significantly higher than those in liver, as shown by tissue distribution studies [310]. Interestingly, although no repair capacity has been tested for a particular benzene adduct in vivo, when treated with alkylating agents, the DRC of primary human hematopoietic CD34+ cells from bone marrow was significantly lower than more differentiated CD34− cells of the same donor [311]. In general, the tissue- or cell type-specific responses to tobacco exposure and their mechanisms are not well understood and await more extensive investigation.

3. From the Known to the Unknown: Future and Perspective

3.1. Uncharacterized Adducts and Repair. The bulky DNA adducts discussed here are only a small list of representative derivatives from tobacco carcinogens. We need not only more information on these important bulky adducts, but also more knowledge of those other adducts that are still not understood. (1) There are many other chemicals in cigarette smoke that have not yet been evaluated for DNA adduct-forming ability. For instance, there are 18 N-nitrosamines and 106 aldehydes present in cigarette smoke [10]. (2) Chemicals may be neglected only because of their lower concentrations in cigarette smoke. For example, some PAHs exist in cigarette smoke at low amounts relative to B[a]P. Some of these compounds may even have stronger carcinogenic effect [30]. (3) The research attention tends to be more focused on the effect of the major adducts of a chemical carcinogen. However, in some cases, the minor adducts are substantially more mutagenic than the major ones. For example, the biological response of N-nitrosamines could be correlated with minor forms of DNA damage [312]. However, the repair specificity and kinetics of these DNA lesions are still unknown. (4) Many bulky adducts have been identified in vitro under physiological conditions but have not been detected yet in vivo. Examples include a number of exocyclic adducts formed by benzene metabolites (Figure 4). Although many in vitro studies have been focused on the HQ- and p-BQ-derived benzenethio adducts [313] and MUC-DNA adducts [179], no conclusion can yet be drawn on whether they are formed in cells/tissues. (5) In some cases, bulky DNA adducts are detected from the tissues of exposed humans or animals using 32P-postlabeling, but their chemical structures have not been elucidated. These adducts were generally described in the literature as “aromatic or hydrophobic adducts”. (6) Endogenous DNA adducts can be formed as a result of cigarette smoke, which is currently attributed to the formation of LPO products [92–94]. Further research is needed to learn the dose-response relationship with regard to the formation of these adducts as well as their relative importance in smoking-induced carcinogenesis. The complexity of such analysis
is that the same DNA adducts, for example, the ε-DNA adducts, are generated by both tobacco carcinogens and endogenously formed compounds, even though they can be chemically different [49]. (7) Although the majority of studies on adducts and cancer have focused on stable ones, many adducts formed by tobacco carcinogens are chemically unstable, for example, the benzene-derived CAT-4-N3A and CAT-4-N7G [183], and the potential effects of such adducts versus stable adducts are largely unknown. (8) Many DNA adducts have been detected, some in cells/tissues, but not yet characterized with respect to their repair. An example is that the epoxides of 1,3-butadiene [314, 315], a tobacco carcinogen in both MSS and SSS, cause the formation of N-1-(2,3,4-trihydroxybutyl) adenine adduct in human samples [316], but its repair is unknown. (9) Sometimes cellular repair has been detected using repair-deficient cells/animals in combination with mutagenesis assays, but detailed biochemical properties of such repair have not been elucidated. In other cases, repair studies have only been performed in vitro, such as the p-BQ-derived benzetheno adducts [253, 255]. Although animal studies have limitations with significant differences from humans, many important experiments on the formation and repair of carcinogenic adducts can only be performed in appropriate animals. The major difference is that in vivo, enzymes may be inducible and become saturated at the carcinogen dose used, and in the genome, there can be preferential repair as well as organ and cell variations and there may be cooperative repair mechanisms which have not been well understood. This is only a partial list of the reasons why repair specificity and efficiency discovered in vitro have to be validated in vivo.

3.2. SHS and THS Carcinogens and DNA Adducts. To investigate the relationship between cigarette smoking and DNA damage, understanding of the chemical and biophysical properties of various forms of tobacco smoke is also critical. For instance, although a great deal is known about the chemistry and toxicity of MSS and SHS, little is known about the identity and molecular toxicology of toxicants produced de novo in THS [6, 7]. However, it has been well established that indoor surfaces significantly adsorb semi- and nonvolatile SHS compounds, for example, nicotine, 3-ethenylpyridine, naphthalene, cresols, and phenol which are slowly reemitted into the air [317–320]. Therefore, all of these compounds can be significant components in THS. Moreover, compounds sorbed onto a surface can undergo chemical transformations by reacting with common reactive atmospheric species. Recent indoor chemistry studies have elegantly revealed that nicotine reacts with ozone (O₃) to yield aldehydes and possibly myosmine [317] and with nitrous acid (HONO) to form NNA, NNN, and NNK [241] (Figure 8). NNA was identified as the major product, which is absent in freshly emitted tobacco smoke but found in in vitro reaction of nicotine with NaNO₂ [321]. NNA has a mutagenic activity similar to that of NNN [322], but its tumorigenic activity in animals was not conclusive [323, 324]. So far, nothing has been tested for its potential to form DNA adducts, but it is expected to be reactive with DNA due to its aldehyde group. Therefore, it would be important to assess its intake and biological properties. Based on their chemical structure, all of the by-products shown in Figure 8 are anticipated or have already been known to form DNA adducts.
As stated earlier, SSS and SHS, the precursors of THS, contain thousands of chemicals, including nicotine and well-defined carcinogens, partitioned between the gas and particulate phases. The secondary analysis by Schick’s group of animal experimental data from the past documents of Philip Morris available at University of California, San Francisco (UCSF) concluded that the mature and “aged” SSS is several-folds more toxic than the fresh SSS [325]. Such evidence constitutes compelling and sufficient rationale to determine the chemical composition of unique toxicants produced in THS de novo and to detect new/different types of DNA adducts formed by “aged” SSS. In conclusion, this is a new and important research area with regard to developing strategies and methods to identify the chemical components SHS aging and THS and to assess currently unknown genotoxic potential and biomarker availability through DNA adduct studies.

3.3. Future Directions. Understanding the mechanisms of formation and repair of bulky DNA adducts is critical for analysis and management of tobacco-induced mutagenesis and carcinogenesis. It is hoped that continued studies will provide more information about the structural and biological implications of specific bulky DNA adducts and the broader range of their effects during the pathogenesis. Future studies should be made to identify and characterize novel compounds and adducts, such as those produced in aged SHS and THS, to identify those highly mutagenic adducts that are refractory to DNA repair, to find adducts as reliable biomarkers for measuring exposure, especially for SHS and THS, to explore new potential biological functions of adducts, such as their interactions with cellular signaling networks, impact on stem cells of target tissues, and roles in epigenetic changes, and to find effective ways to inhibit DNA adduct formation in target organs. Innovative study designs along with more comprehensive approaches (e.g., systems biology) and new technology development (e.g., high-throughput analysis) will be important for achieving these goals. Ultimately, a better knowledge of the mechanisms by which the chemical carcinogen exposure increases cancer risk in smokers and individuals exposed to SHS/THS could lead to new strategies for cancer prevention and could serve as the experimental evidence for framing and enforcing tobacco control policies in order to minimize health hazards and protect vulnerable populations.

Abbreviations

MSS: Mainstream smoke
SSS: Sidestream smoke
SHS: Secondhand smoke
ETS: Environmental tobacco smoke
THS: Thirdhand smoke
PAH: Polycyclic aromatic hydrocarbon
B[a]P: Benz[a]pyrene
BPDE: Benzo[a]pyrene diol epoxide
TSNAs: Tobacco-specific nitrosamines
NNK: 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone
NNN: N’-Nitrosodinitrosonicotine
NNA: 1-(N-methyl-N-nitrosamino)-1-(3-pyridinyl)-4-butanal
POB: Pyridylloxobutyl
O6-POB-dG: O6-[4-(3-Pyridyl)-4-oxobut-1-yl]-2’-dG
4-ABP: 4-Aminobiphenyl
dG-C8-ABP: N-(Deoxyguanosin-8-yl)-4-aminobiphenyl
enals: α,β-Unsaturated aldehydes
PdG: 1,N2-Propanoguanine
p-BQ: para-Benzoxquinone
HQ: Hydroquinone
N2-4-HOPh-dG: N2-(4-Hydroxyphenyl)-2’-dG
MUC: trans,trans-Muconaldehyde
VC: Vinyl chloride
CEO: Chloroethylen oxide
CAA: Chloroacetdehyde
HNE: trans-4-Hydroxy-2-nonenal
e: Etheno
hm: Hydroxymethyl
M1G: Pyrimido[1,2-α]-purin-10(3H)-one
O6-mG: O6-Methylguanine
8-oxoG: 7,8-Dihydro-8-oxoguanine
DSB: Double-strand break
ROS: Reactive oxygen species
LPO: Lipid peroxidation
AGT: O6-Alkylguanine-DNA alkyltransferase
ABH: AlkB homolog
NER: Nucleotide excision repair
GGR: Global genomic repair
TCR: Transcription-coupled repair
BER: Base excision repair
NIR: Nucleotide incision repair
MMR: Mismatch repair
AP: Apurinic/apyrimidinic
Mug: Mismatch-specific uracil-DNA glycosylase
TDG: Thymine-DNA glycosylase
AlkA: E. coli 3-methyladenine DNA glycosylase
II
AAG: Alkylpurine DNA glycosylase
OGG1: 8-Oxoguanine DNA glycosylase
APE1: Major human AP endonuclease
pol: Polymerase
HCR: Host-cell reactivation
DRC: DNA repair capacity
SNP: Single nucleotide polymorphism.

Acknowledgments

The work was supported by the Grant 19XT-0070 (B.H.) from the Tobacco-Related Disease Research Program (TRD RP) of the State of California. The research on chemistry and biochemistry of DNA adducts in our laboratory was previously supported by the NIH/NCI Grant R01CA72079 (to B.H.) and by the Laboratory Directed Research and Development Program (LDRD) of Lawrence Berkeley National Laboratory under U.S. Department of Energy Contract no. DE-AC02-05CH11231.
References


