OXFORD

SOT | Society of Toxicology
www.toxsci.oxfordjournals.org

TOXICOLOGICAL SCIENCES, 146(1), 2015, 170–182

doi: 10.1093/toxsci/kfv079
Advance Access Publication Date: April 21, 2015
Research Article

Formation, Accumulation, and Hydrolysis of Endogenous and Exogenous Formaldehyde-Induced DNA Damage

Rui Yu*,[1], Yongquan Lai*,[1], Hadley J. Hartwell*, Benjamin C. Moeller†, Melanie Doyle-Eisele†, Dean Kracko†, Wanda M. Bodnar*, Thomas B. Starr*,‡, and James A. Swenberg*,[2]

*Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, The University of North Carolina, Chapel Hill, North Carolina 27599, †Lovelace Respiratory Research Institute, Albuquerque, New Mexico 87108; and ‡TBS Associates, 7500 Rainwater Road, Raleigh, North Carolina 27615

[1]These authors contributed equally to this study.
[2]To whom correspondence should be addressed at Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, The University of North Carolina, Chapel Hill, North Carolina 27599. Fax: 919-966-6123. Email: jswenber@email.unc.edu.

ABSTRACT

Formaldehyde is not only a widely used chemical with well-known carcinogenicity but is also a normal metabolite of living cells. It thus poses unique challenges for understanding risks associated with exposure. N^2-hydroxymethyl-dG (N^2-HOMe-dG) is the main formaldehyde-induced DNA mono-adduct, which together with DNA-protein crosslinks (DPCs) and toxicity-induced cell proliferation, play important roles in a mutagenic mode of action for cancer. In this study, N^2-HOMe-dG was shown to be an excellent biomarker for direct adduction of formaldehyde to DNA and the hydrolysis of DPCs. The use of inhaled [$^{13}CD_2$]-formaldehyde exposures of rats and primates coupled with ultrasensitive nano ultra performance liquid chromatography-tandem mass spectrometry permitted accurate determinations of endogenous and exogenous formaldehyde DNA damage. The results show that inhaled formaldehyde only reached rat and monkey noses, but not tissues distant to the site of initial contact. The amounts of exogenous adducts were remarkably lower than those of endogenous adducts in exposed nasal epithelium. Moreover, exogenous adducts accumulated in rat nasal epithelium over the 28-days exposure to reach steady-state concentrations, followed by elimination with a half-life ($t_{1/2}$) of 7.1 days. Additionally, we examined artifact formation during DNA preparation to ensure the accuracy of nonlabeled N^2-HOMe-dG measurements. These novel findings provide critical new data for understanding major issues identified by the National Research Council Review of the 2010 Environmental Protection Agency's Draft Integrated Risk Information System Formaldehyde Risk Assessment. They support a data-driven need for reflection on whether risks have been overestimated for inhaled formaldehyde, whereas underappreciating endogenous formaldehyde as the primary source of exposure that results in bone marrow toxicity and leukemia in susceptible humans and rodents deficient in DNA repair.

Key words: formaldehyde; DNA-protein crosslinks; DNA monoadducts; artifacts; endogenous and exogenous; distribution; accumulation; steady state; half-life; nano liquid chromatography-tandem mass spectrometry

Formaldehyde was first shown to be a carcinogen in 1980, causing squamous cell carcinomas in nasal passages of exposed rats (Swenberg *et al.*, 1980). It was classified as a known human and animal carcinogen in 2006 (International Agency for Research on Cancer, 2006, 2012). Although widespread human exposure has raised longstanding public health concerns, formaldehyde is also an essential metabolic intermediate in all living cells. Formaldehyde is highly genotoxic due to its ability to covalently

bind and induce DNA monoadducts (McGhee and von Hippel, 1977), DNA-DNA crosslinks (Lu *et al.*, 2010a), DNA-protein crosslinks (DPCs) (Casanova-Schmitz and Heck, 1983; Magana-Schwencke and Ekert, 1978; Solomon and Varshavsky, 1985), and DNA-glutathione crosslinks (Lu *et al.*, 2009). The carcinogenic mode of action is based on the synergism among the formation of DNA mono-adducts, DPCs, and increased cell proliferation that lead to mutations (International Agency for Research on Cancer, 2006). Epidemiologic associations have been reported between formaldehyde exposure and the induction of leukemia (Beane Freeman *et al.*, 2009). However, whether or not inhaled formaldehyde exposure causes leukemia remains debatable. Previous experimental results have not supported the induction of leukemia (Lu *et al.*, 2010a), and epidemiological reports have been inconsistent across different studies (Beane Freeman *et al.*, 2009; Coggon *et al.*, 2014). K.J. Patel's laboratory has recently reported that mice deficient in both acetaldehyde catabolism ($ALDH2^{-/-}$) and the Fanconi anemia (FA) DNA repair pathway ($FANCD2^{-/-}$), spontaneously develop severe bone marrow toxicity and acute leukemia (Garaycoechea *et al.*, 2012; Langevin *et al.*, 2011). Moreover, they have shown endogenous formaldehyde to be more abundant and more genotoxic than acetaldehyde (Rosado *et al.*, 2011), demonstrating that endogenous aldehyde concentrations are sufficient to cause DNA damage and induce leukemia in FA mice (Garaycoechea *et al.*, 2012; Langevin *et al.*, 2011).

The U.S. EPA released a draft Integrated Risk Information System (IRIS) risk assessment in June 2010. The National Research Council (NRC) was asked to conduct an independent scientific review of the draft IRIS assessment, which raised 5 issues related to toxicokinetics and modes of action of formaldehyde (Committee to Review EPA's Draft IRIS Assessment of Formaldehyde & National Research Council, 2011). First, the normal presence of formaldehyde in all tissues, cells, and bodily fluids complicates any formaldehyde risk assessment. Thus, an improved understanding of when exogenous inhaled formaldehyde exposure alters normal endogenous or total formaldehyde tissue concentrations was identified as a critical need. Second, inhaled formaldehyde is found predominantly in the respiratory epithelium. The draft IRIS assessment presented divergent opinions regarding the systemic delivery of formaldehyde that need greater resolution. Third, the committee agreed with EPA that formaldehyde may act through a mutagenic mode of action. However, cytotoxicity and compensatory cell proliferation also play critical roles in formaldehyde-induced nasal tumors. The NRC committee recommended that EPA provide alternative calculations that factor in nonlinearities associated with the cytotoxicity-cell proliferation mode of action. Fourth, the NRC recommended that biologically based dose-response models be used. Fifth, the NRC recommended that EPA carefully consider the implications of recent analytical techniques that achieve superior sensitivity and can be highly informative in accurately distinguishing between and quantifying separately exogenous and endogenous formaldehyde-induced DNA mono-adducts and DPCs in tissues (Lu *et al.*, 2010a).

The research presented here is directly responsive to and provides critical information for understanding the major issues identified by the NRC. Of particular importance is the demonstration that exogenous DNA adducts from inhaled formaldehyde do not reach quasi-steady-state concentrations until approximately 28-days of exposure to 2 ppm [$^{13}CD_2$]-formaldehyde (6 h/day, 7 days/week). Thus, for the first time, it is possible to make accurate determinations of both exogenous and endogenous formaldehyde contributions to the total DNA adduct burden. With the support of ultrasensitive nano ultra performance liquid chromatography-tandem mass spectrometry (nano-UPLC-MS-MS), we demonstrate the accumulation of exogenous adducts, as well as the approximate time to steady state, and estimation of the $t_{1/2}$ for the repair/loss of DNA damage *in vivo*. We also provide data regarding the tissue-specific distribution patterns of the exogenous and endogenous DNA adducts, and show N^2-HOMe-dG to be the dominant degradation product of several DPCs, demonstrating that it is an excellent biomarker for both DPCs and direct adduction of formaldehyde to DNA. Finally, we carefully examined artifact formation during DNA isolation and digestion to ensure the accuracy of nonlabeled N^2-HOMe-dG measurements. These studies provide the amounts of endogenous DNA adducts that arise from both direct adduction and hydrolysis of DPCs for several tissues in control and exposed animals. Our data constitute critical information for understanding the roles of both endogenous and exogenous formaldehyde DNA damage in tissue-specific diseases associated with formaldehyde exposure.

MATERIALS AND METHODS

Chemicals and Materials

2′-Deoxyguanosine, reduced L-glutathione, Tris-EDTA (TE buffer), sodium phosphate monobasic (BioXtra), sodium phosphate dibasic (BioXtra), Trizma hydrochloride (BioXtra, BioUltra, BioPerformance), Trizma base (BioXtra), piperazine-N,N′-bis(2-ethanesulfonic acid) (PIPES), magnesium chloride solution, ammonium acetate, acetic acid, formic acid, sodium cyanoborohydride ($NaCNBH_3$), N^2-Me-dG, 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO), leucine aminopeptidase M, carboxypeptidase Y, alkaline phosphatase, phosphodiesterase, DNase I, and DNA single stranded oligo 5′-T_7GT_7-3′ (15-mer) were all purchased from Sigma (St Louis, Missouri). N-Terminal acetylated 12-mer human O^6-alkylguanine DNA methyltransferase (Acetyl-GNPVPILIPCHR, AGT) peptide was synthesized by Genscript Corporation (Piscataway, New Jersey). Methanol, acetonitrile, high-performance liquid chromatography (HPLC) grade water, and optima LC-MS grade water were all purchased from Thermo Fisher Scientific (Pittsburgh, Pennsylvania). High-purity LC-MS grade water and acetonitrile were purchased from Burdick and Jackson Honeywell (Muskegon, Michigan). A 20% solution of formaldehyde in H_2O was procured from Tousimis (Rockville, Maryland). [$^{13}C_{10}{}^{15}N_5$]-dG and a 20% solution of [$^{13}CD_2$]-formaldehyde in D_2O were obtained from Cambridge Isotope Laboratories, Inc (Cambridge, Massachusetts).

Chemical Hazards

Formaldehyde is a highly reactive compound and a known carcinogen and should be handled within a fume hood using appropriate personal protection equipment (PPE).

Synthesis of dG-Me-GSH Crosslink

The synthesis of dG-Me-GSH was carried out according to a previous report (Lu *et al.*, 2010b). Briefly, glutathione (85 mM) was incubated with formaldehyde (100 mM) in 1.75 ml sodium phosphate buffer (100 mM, pH 7.2) at 37°C for 4 h. dG was then added to a final concentration of 16 mM. The reaction mixture was further incubated at 37°C for 6 h. The purification of dG-Me-GSH was performed on an Agilent 1200 series UV HPLC system (Agilent Technologies, Santa Clara, California). The analyte was

separated on a C_{18} reverse phase column (Waters Atlantis T3, 3 μm, 150 × 4.6 mm, Waters Corporation, Milford, Massachusetts). The mobile phases were 0.05% acetic acid in water (A) and pure acetonitrile (B). The column temperature was kept at 15°C. The flow rate was 0.45 ml/min with a starting condition of 2% B, which was held for 3 min, followed by a linear gradient up to 15% B at 20 min, and held for 8 min, followed by 7 min at 80% B, then re-equilibrated to the starting conditions for 10 min. The UV detector was set at 254 nm and dG-Me-GSH eluted at the retention time (RT) of 24.5 min. The fractions containing dG-Me-GSH were combined, followed by concentration using a speed vac.

Synthesis of dG-Me-Cys Crosslink

dG-Me-Cys was synthesized by digestion of dG-Me-GSH using leucine aminopeptidase M and carboxypeptidase Y. Initially, dG-Me-GSH was prepared and purified using the same methods as described earlier. dG-Me-GSH was digested in 1.2 ml sodium phosphate buffer (40 mM, pH 6.0) in the presence of 33 μg/ml carboxypeptidase Y, 133 μg/ml leucine aminopeptidase M, 10 mM $MgCl_2$, and 10 mM $CaCl_2$. The mixture was incubated at room temperature for 16 h and transferred to a centrifugal filter. Enzymes were removed by centrifugation at 11 000 × g for 30 min, and the resultant filtrate was separated by the same HPLC method as described earlier with a different elution gradient. The flow was initiated with a starting condition of 2% B, which was held for 3 min, followed by a linear gradient up to 4.1% B at 42 min, followed by 13 min at 90% B, then re-equilibrated to the starting conditions for 10 min. dG-Me-Cys eluted at RT of 38.5 min and the fractions containing dG-Me-Cys were collected and concentrated in a speed vac.

Synthesis of dG-Me-AGT Crosslink

12-Mer AGT peptide (0.8 mM) was incubated with formaldehyde (20 mM) in 98 μl sodium phosphate buffer (10 mM, pH 7.2) at 37°C for 3 h. Then, 2 μl dG was added to a final concentration of 2 mM, followed by additional incubation at 37°C for 7 h. dG-Me-AGT was purified by the same HPLC method for dG-Me-GSH crosslink with a different elution gradient. The flow was initiated with a starting condition of 6% B, which was held for 18 min, followed by a linear gradient up to 20% B at 30 min, and kept for 18 min, then re-equilibrated to the starting conditions for 12 min. dG-Me-AGT eluted at 44.0 min and the HPLC fractions containing dG-Me-AGT were collected and concentrated in a speed vac.

Degradation of Crosslinks

To study the kinetics of crosslink degradation, crosslinks in sodium phosphate buffer (20 mM, pH 7.2) were incubated at 37°C. A 25 μl aliquot was removed to determine the concentration of dG, N^2-HOMe-dG, and crosslinks at different time points. For those samples that were not readily analyzed, they were acidified with 1.5 μl of 3% acetic acid solution to stop the degradation of crosslinks and then kept at 4°C until analysis. The analysis of the reaction mixture at different time points was performed on an Agilent 1200 series UV HPLC system as described earlier. Isocratic elution with 6% solvent B at flow rate of 0.45 ml/min was used for dG-Me-Cys and dG-Me-GSH degradation studies. For dG-Me-AGT degradation analysis, the flow rate was 0.45 ml/min with a starting condition of 8% B, which was held for 14 min, followed by a linear gradient up to 30% B at 20 min, and kept for 10 min, then re-equilibrated to the starting conditions for 10 min. UV detector was set at 254 nm, because dG, N^2-HOMe-dG, dG-Me-Cys, dG-Me-GSH, and dG-Me-AGT shared similar UV absorption at 254 nm (Supplementary Fig. SI-1). The concentrations of dG, N^2-HOMe-dG, and crosslinks were calculated according to a dG standard calibration curve ranging from 0.02 to 0.8 μM (Supplementary Fig. SI-2).

Structural Identification of Synthesized Crosslinks

The synthesized crosslinks were characterized using an Agilent 1200 series diode array detector (DAD) HPLC system coupled with Agilent quadrupole-time-of-flight (QTOF)-MS (Agilent Technologies, Santa Clara, California). HPLC separation was carried out on a C_{18} reverse phase column (Waters Atlantis T3, 3 μm, 150 × 2.1 mm) with a flow rate at 0.2 ml/min and mobile phase A (0.05% acetic acid in water) and B (acetonitrile). For the separation of dG, N^2-HOMe-dG, dG-Me-Cys, and dG-Me-GSH, the flow was initiated with a starting condition of 2% B, which was held for 5 min, followed by a linear gradient up to 10% B at 10 min, and kept for 8 min, then re-equilibrated to the starting conditions for 10 min. dG, N^2-HOMe-dG, dG-Me-Cys, and dG-Me-GSH were eluted with retention times of 12.5, 13.5, 12.3, and 14.3 min, respectively. For the separation of dG-Me-AGT, the flow was initiated with a starting condition of 2% B, which was held for 3 min, followed by a linear gradient up to 30% B at 20 min, and kept for 5 min, then re-equilibrated to the starting conditions for 10 min. dG-Me-AGT was eluted with RT of 20.2 min. The DAD detector was set at 254 nm and the UV lamp wavelength ranged from 190 to 300 nm. The electrospray ion source in positive mode with the following conditions were used: gas temperature, 200°C; drying gas flow, 12 l/min; nebulizer, 35 psi; Vcap, 4000 V; fragmentor, 175 V; skimmer, 67 V. Electrospray ionization (ESI)-MS-MS spectrum of dG-Me-Cys and dG-Me-GSH was obtained by in source fragmentation, whereas collision energy was set at 17 V for the fragmentation of dG-Me-AGT.

Animal Exposures

Test atmospheres of formaldehyde were generated by the thermal depolymerization of [$^{13}CD_2$]-paraformaldehyde (CAS No. 30525-89-4) obtained from Cambridge Isotopes Laboratories, Inc. Confirmation of the [$^{13}CD_2$]-paraformaldehyde purity and identity was accomplished using gas chromatography-mass spectrometry. The test atmospheres of [$^{13}CD_2$]-formaldehyde were generated from an 80 l Tedlar bag using a peristaltic pump to deliver vaporized [$^{13}CD_2$]-formaldehyde into the nose-only chamber supply airflow. The concentration of the exposure chamber was monitored by collection of Waters XpoSure Aldehyde Sampler cartridges every 5 min continuously throughout the exposure. Cartridges were extracted with acetonitrile and extracts were analyzed by HPLC.

Animal use in this study was approved by the Institutional Animal Use and Care Committee of The Lovelace Respiratory Research Institute and was conducted in accordance with the National Institutes of Health guidelines for the care and use of laboratory animals. Animals were housed in fully accredited American Association for Accreditation of Laboratory Animal Care facilities. Male F344 rats were exposed to 2 ppm [$^{13}CD_2$]-formaldehyde atmospheres for 7, 14, 21, or 28 consecutive days (6 h/day) with postexposure at 6, 24, 72, and 168 h using a single nose-only unit (Lab Products, Seaford, Delaware). Control rats were exposed to filtered air (6 h/day, 28 days) using nose-only exposure chambers. Monkeys (cynomolgus macaques) were whole body exposed to 6 ppm [$^{13}CD_2$]-formaldehyde on 2 consecutive days (6 h/day). The air supply was maintained at approximately 23 l/min, which was controlled with a mass flow

controller. Control monkeys were whole body exposed to filtered air (6 h/day, 2 days).

Animals were anesthetized with Nembutal (50–60 mg/kg) IP within 2 h of the end of exposure; 3–5 ml of blood was collected by cardiac puncture for lymphocyte isolation. Nasal respiratory epithelium from the right and left sides of the nose and from the septum were collected, as were entire tissues of white blood cells, spleen, thymus, tracheal bronchial lymph nodes (TBLN), mediastinal lymph nodes, trachea, lung, kidney, liver, and brain. Bone marrow was collected from both femurs by saline extrusion with a large bore needle. For monkeys, bone marrow was collected by both saline extrusion and direct scraping. Tissue samples were collected and immediately frozen on dry ice followed by storage at −80°C. The sacrifice and tissue collection of air control animals were the same as the exposed animals.

DNA Isolation, Reduction, and Digestion

DNA was isolated from the tissues using a NucleoBond DNA isolation kit (Macherey-Nagel, Bethlehem, Pennsylvania), as instructed by the manufacturer with small modifications. The resultant DNA was quantified and stored at −80°C for further analysis. Two μl of DNA was diluted in 998 μl of 1X TE buffer and its amount was quantitated using a Thermo Spectronic BioMate 5 UV-visible spectrophotometer. DNA amounts from different tissues ranged from 1.1 to 1396.9 μg. To reduce endogenous and exogenous N^2-HOMe-dG to N^2-Me-dG, DNA was thawed and incubated with $NaCNBH_3$ (50 mM) and sodium phosphate buffer (100 mM, pH 7.2) for 6 h at 37°C. Following reduction, DNA was frozen at −80°C until digested. Reduced DNA was thawed, and 200 μl of sodium phosphate/$MgCl_2$ buffer (50/20 mM final concentration, pH 7.2) along with 20 fMol (for samples measured on Thermo Quantum) or 2 fMol (for samples measured on AB SCIEX) of the internal standard, [$^{13}C_{10}{}^{15}N_5$]-N^2-Me-dG, was added. DNA was digested with DNAse I (200 units), alkaline phosphatase (5 units), and phosphodiesterase (0.005 units) for 1 h at 37°C. Following digestion, hydrolyzed DNA was filtered with a Pall NanoSep 3 kDa filter (Port Washington, New York) at 8000 rpm for 50 min.

HPLC Purification and Fractionation

Hydrolyzed DNA was injected onto an Agilent 1200 HPLC fraction collection system equipped with a DAD (Agilent Technologies, Santa Clara, California). Analytes were separated by reversed-phase liquid chromatography using an Atlantis C_{18} T3 (150 × 4.6 mm, 3 μm) column. The column temperature was kept at 30°C. The mobile phases were water with 0.1% acetic acid (A) and acetonitrile with 0.1% acetic acid (B). The flow rate was 1.0 ml/min with a starting condition of 2% B, which was held for 5 min, followed by a linear gradient of 4% B at 20 min, 10% B at 30 min, followed by 6 min at 80% B, then re-equilibrated to the starting conditions for 20 min. dG and N^2-Me-dG eluted with retention times of 13.1 and 25.5 min, respectively. The amount of dG in the samples was quantitated by the UV peak area ($\lambda = 254$ nm) at the corresponding RT using a calibration curve ranging from 0.02 to 100 nMol dG on column.

Liquid Chromatography-Tandem Mass Spectrometry

LC-MS-MS analyses of N^2-HOMe-dG were performed on 2 instruments. First, a TSQ Quantum Ultra triple-stage quadrupole mass spectrometer (Thermo Scientific, San Jose, California) was operated in selected reaction monitoring (SRM) mode to detect and quantify N^2-Me-dG. The mass spectrometer was interfaced with a nano-acquity UPLC system from Waters Corporation (Milford, Massachusetts). A 20 × 0.18 mm Symmetry C_{18} trap column (5 μm particle size) and a 100 × 0.1 mm HSS T3 analytical column (1.8 μm particle size) from Waters were used. The trap column was kept at room temperature, and the analytical column temperature was kept at 35°C during each run. Mobile phases were comprised of water with 0.1% acetic acid (A) or acetonitrile with 0.1% acetic acid (B). Analytes were first retained on a trap column with a flow rate of 5 μl/min of 2% mobile phase B, followed by transfer to the analytical column with an initial starting condition of 2% B at 0.6 μl/ min for 5 min followed by a linear gradient to 30% B over 12.5 min and to 80% B over 1.5 min. The flow was then held at 80% B for 1 min followed by re-equilibration for an additional 5 min. The analytes were introduced to the mass spectrometer using positive-mode electrospray ionization with a source voltage of 2300 V and no additional gases. The ion transfer tube was held at 300°C and skimmer offset set to zero. Scan speed was set at 75 ms, scan width at 0.1 mass to charge ratio (*m/z*), and a peak width at 0.7 *m/z* for Q1 and Q3. Collision energy was set at 17 eV with Argon as the collision gas set at 1.5 arbitrary units. N^2-HOMe-dG was quantified as N^2-Me-dG after reduction using the transition of *m/z* 282.2 to *m/z* 166.1, and [$^{13}CD_2$]-N^2-HOMe-dG was quantified as [$^{13}CD_2$]-N^2-Me-dG with the transition of *m/z* 285.2 to *m/z* 169.1. Two additional transitions, including *m/z* 284.2 to *m/z* 168.1 and *m/z* 283.2 to *m/z* 167.1, were also monitored in case hydrogen-deuterium (H-D) exchange occurred. Similarly, the transition of *m/z* 297.2 to *m/z* 176.1 was chosen for the internal standard, [$^{13}C_{10}{}^{15}N_5$]-N^2-Me-dG. The calibration curves were obtained using the integrated peak area and amount of injected analytical standard and internal standard. The injection volume was 6 μl.

The most sensitive nano-LC-MS-MS instrument was an AB SCIEX Triple Quad 6500 mass spectrometer (Foster City, California) operated in SRM mode to detect and quantify N^2-Me-dG. The mass spectrometer was interfaced with an Eksigent nanoLC Ultra 2D system (Dublin, California). A 150 × 0.075 mm Eksigent ChromXP $3C_{18}$-CL-120 analytical column (3 μm particle size) was used. The column was kept at room temperature during each run. Mobile phases were comprised of water with 0.1% formic acid (A) or acetonitrile with 0.1% formic acid (B). Analytes were injected into the analytical column with an initial starting condition of 5% B at 0.3 μl/ min for 10 min followed by a linear gradient to 35% B over 12 min and to 60% B over 1 min. The flow was then held at 60% B for 6 min followed by re-equilibration for an additional 10 min. The analytes were introduced to the mass spectrometer using positive-mode nanospray ionization with a source voltage of 2400 V, curtain gas of 20 psi, ion source gas of 10 psi, and collision gas setting of 10 arbitrary units. The interface heater temperature was held at 120°C. Declustering potential was set to 25 V, entrance potential was 10 V, collision energy was 15 eV, and collision cell exit potential was 12 V. The transitions were the same as those used on Thermo TSQ Quantum. The calibration curves were obtained using the integrated peak area and amount of injected analytical standard and internal standard. The injection volume was 1 μl.

Statistical Analysis and Pharmacokinetic Modeling

In vitro *hydrolytic degradation of DPCs*. For the *in vitro* hydrolysis of DPCs, half-life values were calculated using GraphPad Prism 5.0 (GraphPad Software, San Diego, California) based on a 1 or 2 phase exponential decay model. Data represent mean ± standard deviation (SD). Degradation data obtained from both dG-Me-Cys and dG-Me-GSH were best fit with a single phase model. Degradation data of dG-Me-AGT were best fit with

a 2 phase exponential decay model. Samples were measured in triplicate unless otherwise indicated.

Endogenous N^2-HOMe-dG adducts. The statistical significance of adduct concentration differences were assessed using statistical tests appropriate to the different experimental designs. Two-sided unpaired student's *t*-tests were used to compare tissue-specific endogenous adduct concentrations in exposed and air control animals. The 2-sided Dunnett's test (Dunnett, 1964) for comparing multiple treatment groups with a common control group was utilized in assessing differences by treatment duration in relation to the single air control group employed in the 28-day exposure study. Concentration differences were considered to be statistically significant if $P < .05$ for all statistical tests.

Estimated half-life ($t_{1/2}$) and time for exogenous [$^{13}CD_2$]-N^2-HOMe-dG adducts to reach a steady-state concentration. Samples were measured in triplicate unless otherwise indicated; tabulated data represent mean ± SD. Modeling of the exogenous adduct data from 28-days study data required a single exponential decay formulation modified to account for the 28 consecutive daily 6-h exposure periods. This pharmacokinetic model consists of a single linear compartment with constant, but intermittent, forcing. The model has 2 parameters, A, the asymptotic adduct concentration with continuous forcing, and *T*, the mean adduct lifetime. The contribution to adduct burden from the *i*th exposure period ($i = 1, 28$) for time *t* (expressed in days), with $0 \le t < (i-1) + 0.25$, is given by

$$A * (1 - \exp(-(t - (i-1))/T)),$$

whereas for $t \ge (i-1) + 0.25$, the contribution from the *i*th exposure period is given by:

$$A * (1 - \exp(-0.25/T)) * \exp(-(t - (i-1) - 0.25)/T)).$$

The total adduct concentration at time *t* is obtained by summing the individual contributions from prior and, when appropriate, concurrent 6-h exposure periods.

A maximum likelihood estimate of the adduct half-life was obtained by using the Microsoft Excel 2007 Solver nonlinear optimization routine and inverse variance weighting of squared residuals. A likelihood ratio test-based confidence interval (90%) for the half-life estimate was also generated.

RESULTS

Characterization and In Vitro *Degradation of Formaldehyde-Induced DPCs*

Based on previous work in our lab (Lu *et al.*, 2009, 2010b), dG-Me-GSH and dG-Me-AGT are formed *in vitro* and contain specific dG-Me-Cysteine linkages (Fig. 1A). These DPCs were synthesized by reaction of formaldehyde with GSH or AGT in solution prior to dG addition. dG-Me-Cys was obtained by digestion of dG-Me-GSH with aminopeptidase M and carboxypeptidase Y. These crosslinks showed similar UV spectra to that of dG in wavelength ranges from 240 to 300 nm (Supplementary Fig. SI-1). The chemical structures of these crosslinks were characterized using ESI-QTOF-MS-MS. Two major product ions were observed from MS-MS spectra of the protonated precursor ion of dG-Me-Cys, resulting from the loss of the deoxyribosyl group followed by the additional loss of cysteine (Supplementary Fig. SI-3A). The MS-MS spectrum of dG-Me-GSH is consistent with results from our previous study (Lu *et al.*, 2009) showing 2 additional fragment ions caused by the cleavage of the dG-Me and Me-GSH bonds (Supplementary Fig. SI-3B). As shown in Supplementary Fig. SI-3C, 2 major product ions having (+2) charges were generated from the doubly charged precursor ion of dG-Me-AGT via loss of the deoxyribosyl group and dG. The accurate masses for parent crosslinks and their major fragment ions by ESI-QTOF-MS-MS are summarized in Supplementary Table SI-1, with mass error <1.69 ppm.

The linkage N^2-dG-Me-Cys is chemically labile due to the active methylene group. Under physiological pH and temperature, the estimated half-life for dG-Me-Cys and dG-Me-GSH was 11.6 min ($R^2 = 0.9823$) and 79.6 min ($R^2 = 0.9964$), respectively (Fig. 1B and C). The degradation kinetics of dG-Me-AGT consisted of a rapid initial fall with half-life of 7.4 min followed by a much slower decay phase characterized by a half-life of 90.2 min (PercentFast = 36.5%, $R^2 = 0.9987$) (Fig. 1D). During the hydrolysis of DPCs, dG could be directly formed if cleavage occurs between dG and methylene (N^2-dG-Me bond). However, dG was not detected in the first 3, 10, or 20 min incubations for dG-Me-Cys, dG-Me-AGT, and dG-Me-GSH, respectively (Fig. 1B–D). On the other hand, N^2-HOMe-dG was readily formed and increased rapidly after initiating incubation of all crosslinks, suggesting cleavage of the Me-Cys bond, but not N^2-dG-Me bond. dG was only detected after long periods of incubation, because N^2-HOMe-dG is also unstable and degrades to dG. Thus, the mechanism of DPC hydrolysis involves breakdown to N^2-HOMe-dG, followed by further degradation to dG (Fig. 1A). These results suggest that DPCs may be important sources of formaldehyde-induced DNA mono-adducts, supporting N^2-HOMe-dG as a major biomarker of formaldehyde exposure.

Method Validation and Artifact Determination

Details for method validation and artifact determination can be found in Supplementary Results and Supplementary Discussion. Briefly, the limit of detection values of [$^{13}C_{10}{}^{15}N_5$]-N^2-Me-dG in DNA matrix on the Thermo TSQ Quantum and the AB SCIEX Triple Quad 6500 [defined as signal to noise (S/N) > 3] were 10 and 0.5 aMol, respectively. The sodium phosphate buffer, as opposed to Tris•HCl buffer or PIPES buffer, did not produce significant matrix effects during sample reduction and digestion, such as N^2-Me-dG signal suppression or enhancement. Moreover, minimal artifacts were found to be generated from lipid oxidation in the presence of Tris•HCl buffer in the NucleoBond DNA isolation kit (Macherey-Nagel, Bethlehem, Pennsylvania), or from our routine DNA isolation method. We demonstrated that less than 1.1% artifact was associated with measurements of endogenous formaldehyde DNA adducts.

Formation of N^2-Hydroxymethyl-dG DNA Adducts In Vivo *from Endogenous and Exogenous Sources (28-Day Rat Inhalation Exposure Study)*

Following completion of exposures, tissue samples were collected, immediately frozen, and stored at −80°C until DNA isolation, reduction, digestion, HPLC fractionation, and DNA adduct quantitation by nano-LC-MS-MS. Figure 2A top panel shows the chromatogram of N^2-Me-dG in the nasal DNA from a rat exposed to [$^{13}CD_2$]-formaldehyde for 7 days. The peak corresponding to the specific transition of *m/z* 282.2 to *m/z* 166.1 represents the endogenous N^2-Me-dG, which is the reduced form

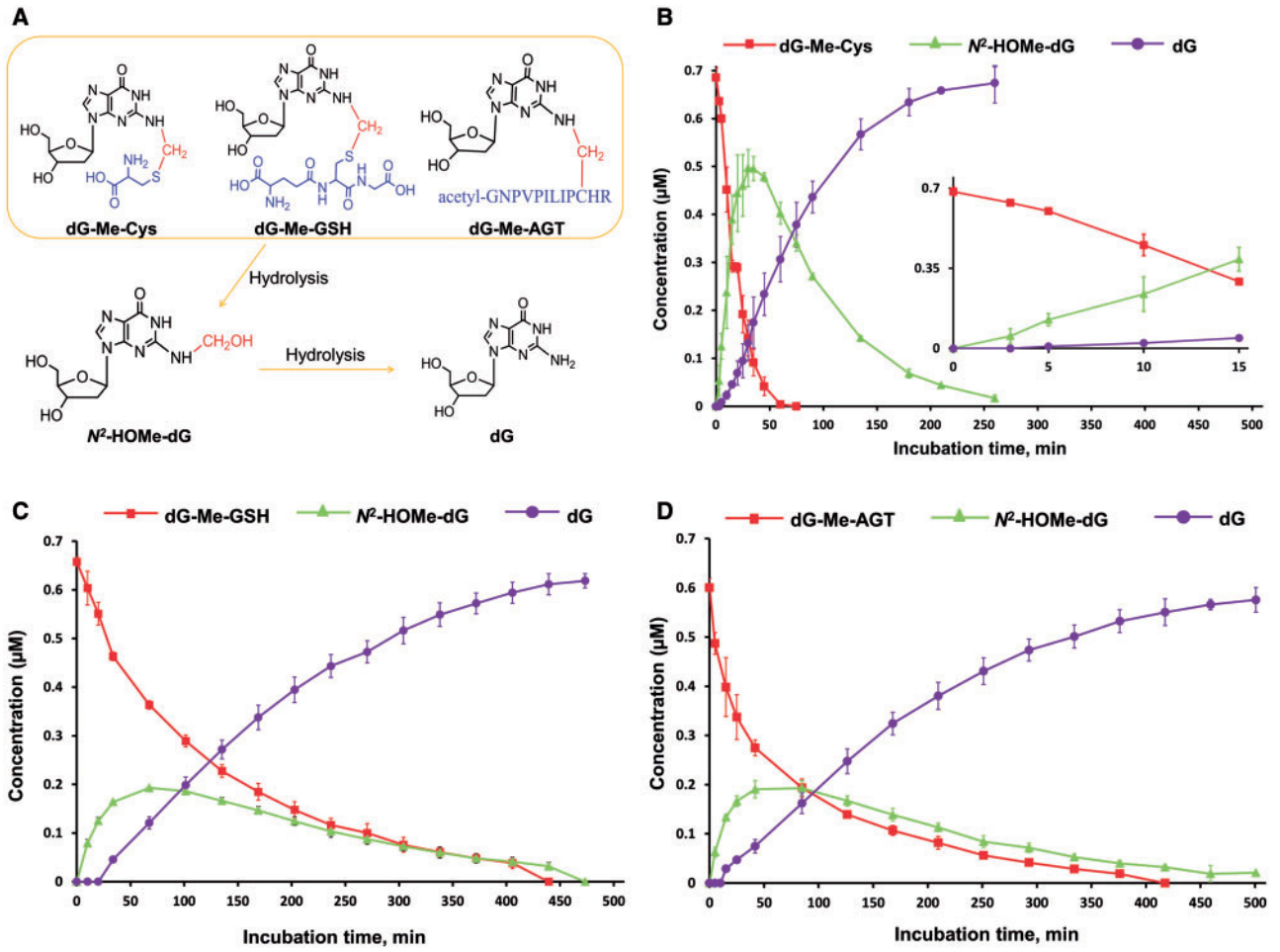

FIG. 1. Hydrolytic degradation pathway and kinetics of DNA-protein crosslinks (DPCs). A, Degradation pathway; B–D, *In vitro* time dependence of the breakdown of DPCs, the formation and degradation of N^2-HOMe-dG, and the formation of dG. Inserted graph in B shows the time course reaction of dG-Me-Cys from 0 to 15 min. The *in vitro* degradation study of crosslinks was performed at 37°C in sodium phosphate buffer (20 mM, pH 7.2). Data represent mean ± SD ($n = 3$).

of N^2-HOMe-dG derived from endogenous formaldehyde. The bottom panel represents the internal standard, [$^{13}C_{10}{}^{15}N_5$]-N^2-Me-dG, in the transition of *m/z* 297.2 to *m/z* 176.1. The peak in the middle panel corresponding to the transition of *m/z* 285.2 to *m/z* 169.1 represents the exogenous [$^{13}CD_2$]-N^2-Me-dG, which is the reduced form of [$^{13}CD_2$]-N^2-HOMe-dG derived from exogenous [$^{13}CD_2$]-formaldehyde. Similarly, both endogenous and exogenous DNA adducts can be clearly identified and quantified in nasal epithelial DNA after 28 days of exposure, as shown in Figure 2B and C. Compared with Figure 2A, the amount of exogenous [$^{13}CD_2$]-N^2-Me-dG in nasal DNA increased with extended exposure time, indicating that DNA adducts in the form of N^2-HOMe-dG accumulated during the 28-day exposure period *in vivo*. In control rats, equivalent amounts of endogenous adducts and no exogenous adducts were observed (Fig. 2D and E).

In all other tissues, only endogenous formaldehyde-induced N^2-HOMe-dG could be observed except for 1 bone marrow sample (Table 1) from a 28-day exposure rat. No signals corresponding to the transition of *m/z* 285.2 to *m/z* 169.1 of exogenous adducts were present following 28 days of labeled formaldehyde exposure or in control rats. Furthermore, 2 additional transitions monitoring the precursor masses of *m/z* 284.2 and 283.2 did not detect any exogenous DNA adducts arising from possible H-D exchange. Thus, with the exception of the 1 bone marrow sample mentioned earlier, there were no detectable amounts of [$^{13}CD_2$]-N^2-HOMe-dG in sites remote to the portal of entry after 28 days of exposure.

The amounts of endogenous N^2-Me-dG and exogenous [$^{13}CD_2$]-N^2-Me-dG in nasal respiratory epithelium, bone marrow, and white blood cells were determined from replicate samples ($n = 4$–12/time point) and are shown in Table 1. Amounts of endogenous and exogenous DNA adducts in thymus, TBLN, mediastinal lymph nodes, trachea, lung, spleen, kidney, liver, and brain were determined in 28-day exposures and control samples and are shown in Table 2.

Steady-State and Half-Life Modeling of Exogenous N^2-Hydroxymethyl-dG DNA Adducts In Vivo *(28-Day Inhalation Exposure Study)*

As exogenous DNA adducts only existed in nasal respiratory epithelium, the steady state and $t_{1/2}$ estimations were based on the amounts of [$^{13}CD_2$]-N^2-HOMe-dG that had accumulated at various time points over the course of the 28-day exposure and 7-day postexposure periods (data from a single 6 h exposure to 2 ppm were included from a previous study) (Lu *et al.*, 2011). As

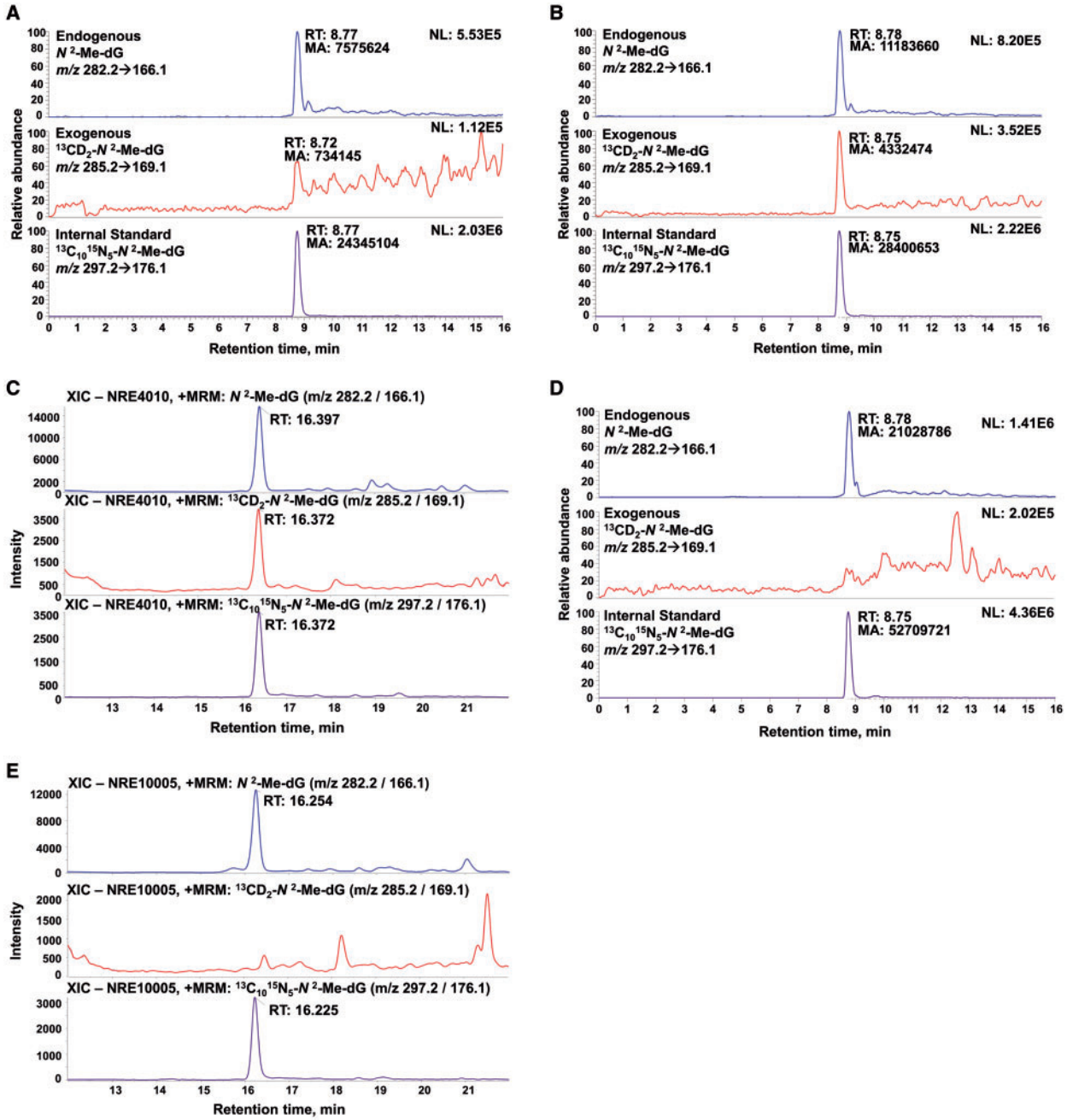

FIG. 2. Typical nano-LC-MS-MS SRM chromatograms of endogenous and exogenous N^2-Me-dG adducts in rat nasal respiratory epithelium. A, Seven-day exposed tissue analyzed on Thermo TSQ Quantum; B, Twenty-eight-day exposed tissue analyzed on Thermo TSQ Quantum; C, Twenty-eight-day exposed tissue analyzed on AB SCIEX Triple Quad 6500; D, Control rat tissue analyzed on Thermo TSQ Quantum; E, Control rat tissue analyzed on AB SCIEX Triple Quad 6500. Abbreviations: RT, retention time; MA, manual quantitated peak area; NL, normalized spectrum to largest peak in particular chromatogram; XIC, extracted-ion-chromatogram; NRE, nasal epithelium; MRM, multiple reaction monitoring.

shown in Figure 3, the exogenous DNA adducts, [$^{13}CD_2$]-N^2-HOMe-dG, approached a steady-state concentration during the 28-day exposure phase. After the last exposure, there was a rapid initial loss of nearly 20% of the DNA adducts during the 6 h immediately following the last exposure. This rapid initial loss was then followed by a much slower decrease in adducts (Fig. 3). The $t_{1/2}$ for the formation and repair/loss of [$^{13}CD_2$]-N^2-HOMe-dG adducts in nasal respiratory epithelium was estimated to be 7.1 days, 90% confidence interval (CI) = [6.0, 8.7] days using all of the data.

Formation of N^2-Hydroxymethyl-dG DNA Adducts In Vivo from Endogenous and Exogenous Sources (Monkey Inhalation Exposure Study)

As shown in Table 3, combining our first monkey [Nonhuman primate 1 (NHP1)] (Moeller *et al.*, 2011) and second monkey (NHP2) studies, both endogenous and exogenous N^2-HOMe-dG adducts were found in all regions of nasal passages studied. Moreover, the amounts of exogenous adducts in different sections of the exposed monkey nasal epithelium were 5- to 11-fold lower than endogenous adducts in the

TABLE 1. Formation of N^2-HOMe-dG Mono-Adducts (mean ± SD) in Rat Nasal Epithelium, Bone Marrow, and White Blood Cells Exposed to 2 ppm Labeled Formaldehyde for 28 Days

Exposure Period	Rat Nasal Epithelium			Rat Bone Marrow			Rat White Blood Cells		
	N^2-HOMe-dG (adducts/10^7 dG)			N^2-HOMe-dG (adducts/10^7 dG)			N^2-HOMe-dG (adducts/10^7 dG)		
	Endogenous[a]	Exogenous	*n*	Endogenous[a]	Exogenous	*n*	Endogenous[a]	Exogenous	*n*
7 days	2.51 ± 0.63	0.35 ± 0.17	5	3.37 ± 1.56	n.d.	6	2.62 ± 1.12	n.d.	4
14 days	3.09 ± 0.98	0.84 ± 0.17	5	2.72 ± 1.36	n.d.	6	2.26 ± 0.46	n.d.	4
21 days	3.34 ± 1.06	0.95 ± 0.11	5	2.44 ± 0.96	n.d.	6	2.40 ± 0.47	n.d.	4
28 days	2.82 ± 0.76	1.05 ± 0.16	6	3.43 ± 2.20	0.34[b]	12	2.49 ± 0.50	n.d.	4
28 days + 6 h postexpo	2.80 ± 0.58	0.83 ± 0.33	9	2.41 ± 1.14	n.d.	6	2.97 ± 0.58	n.d.	4
28 days + 24 h postexpo	2.98 ± 0.70	0.80 ± 0.46	9	4.67 ± 1.84	n.d.	5	2.57 ± 0.58	n.d.	4
28 days + 72 h postexpo	2.99 ± 0.63	0.63 ± 0.12	9	5.55 ± 0.76	n.d.	6	1.75 ± 0.26	n.d.	4
28 days + 168 h postexpo	2.78 ± 0.48	0.67 ± 0.20	10	2.78 ± 1.94	n.d.	4	2.61 ± 1.22	n.d.	4
Air control	2.84 ± 0.54	n.d.	8	3.58 ± 0.99	n.d.	6	2.76 ± 0.66	n.d.	6

[a]No statistically significant difference was found using the 2-sided Dunnett's test (multiple comparisons with a control) (Dunnett, 1964).
[b]The amount of exogenous N^2-HOMe-dG adducts that was found in only 1 bone marrow sample analyzed by AB SCIEX Triple Quad 6500.
n.d., not detected.

TABLE 2. Formation of N^2-HOMe-dG Mono-Adducts (Mean ± SD) in Rat Thymus, Tracheal Bronchial Lymph Nodes, Mediastinal Lymph Nodes, Trachea, Lung, Spleen, Kidney, Liver, and Brain Exposed to 2 ppm Labeled Formaldehyde for 28 Days

Rat tissues	Exposure Period	N^2-HOMe-dG (adducts/10^7 dG)			Exposure Period	N^2-HOMe-dG (adducts/10^7 dG)		
		Endogenous	Exogenous	*n*		Endogenous	Exogenous	*n*
Thymus[a]	28 days	0.63 ± 0.06	n.d.	4	Air control	0.78 ± 0.04	n.d.	4
TBLN		3.01 ± 0.71	n.d.	4		3.46 ± 1.24	n.d.	4
Lymph nodes		2.80 ± 1.38	n.d.	4		2.99 ± 0.85	n.d.	4
Trachea		2.63 ± 0.92	n.d.	4		3.18 ± 0.72	n.d.	4
Lung		2.13 ± 0.26	n.d.	4		2.29 ± 0.24	n.d.	4
Spleen		1.83 ± 0.25	n.d.	4		2.18 ± 0.19	n.d.	4
Kidney		1.99 ± 0.09	n.d.	4		2.17 ± 0.60	n.d.	4
Liver		1.80 ± 0.02	n.d.	4		1.97 ± 0.38	n.d.	4
Brain		2.35 ± 1.00	n.d.	4		2.13 ± 0.17	n.d.	4

[a]Statistically significant difference was found using the 2-sided unpaired Student's *t*-tests.
TBLN, tracheal bronchial lymph nodes.
n.d., not detected.

corresponding sections in both exposed and air control monkey samples. In monkey bone marrow, white blood cells, and trachea, only endogenous adducts were found; however, the amount of endogenous N^2-HOMe-dG adducts present in direct scraped bone marrow was at least 2 times higher than in collection using saline extrusion, indicating distinct differences between these 2 bone marrow collection procedures.

DISCUSSION

The formation of formaldehyde-induced DPCs has long been thought to be a major form of DNA damage (Magana-Schwencke and Ekert, 1978). Due to its significant biological consequences, there have been investigations on formaldehyde-induced DPCs in animals and humans. Increased DPCs were found in nasal samples of rats exposed to formaldehyde at concentrations ≥2 ppm (Casanova *et al.*, 1984). Similarly, DPCs were formed in the respiratory tract of rhesus monkeys exposed to formaldehyde at concentrations as low as 0.7 ppm (Casanova *et al.*, 1991). In those studies, no DPCs were detected in the tissues distant to the contact site in rats and monkeys exposed to formaldehyde at concentrations higher than 6 ppm, such as lung and bone marrow. In contrast, some studies have reported an apparent increase in DPCs in circulating lymphocytes of humans exposed to ambient concentrations in the workplace (Shaham *et al.*, 1996, 2003). Increased DPCs were also found in several remote tissues, such as bone marrow, liver, kidney, and testes of mice exposed to inhaled formaldehyde (Ye *et al.*, 2013).

The debate on the formation of formaldehyde-induced DPC in the animal tissues distant to initial contact may be due to the use of nonspecific DPC assays, such as potassium-SDS precipitation and chloroform/iso-amyl alcohol/phenol extraction. The methods lack specificity for the chemical composition of DPCs (Stingele *et al.*, 2014; Zhitkovich and Costa, 1992; Shaham et al., 1996, 2003). Moreover, they cannot differentiate between exogenous and endogenous formaldehyde-induced DPCs. Other studies have shown a rapid decline of the formaldehyde-induced DPCs in cell cultures, and no accumulation of DPCs was observed in rats after repeated exposure (Casanova *et al.*, 1994; Grafstrom *et al.*, 1984; Shoulkamy *et al.*, 2012). Further evidence indicates that the major portions of formaldehyde-induced DPCs are lost from lymphocytes through spontaneous

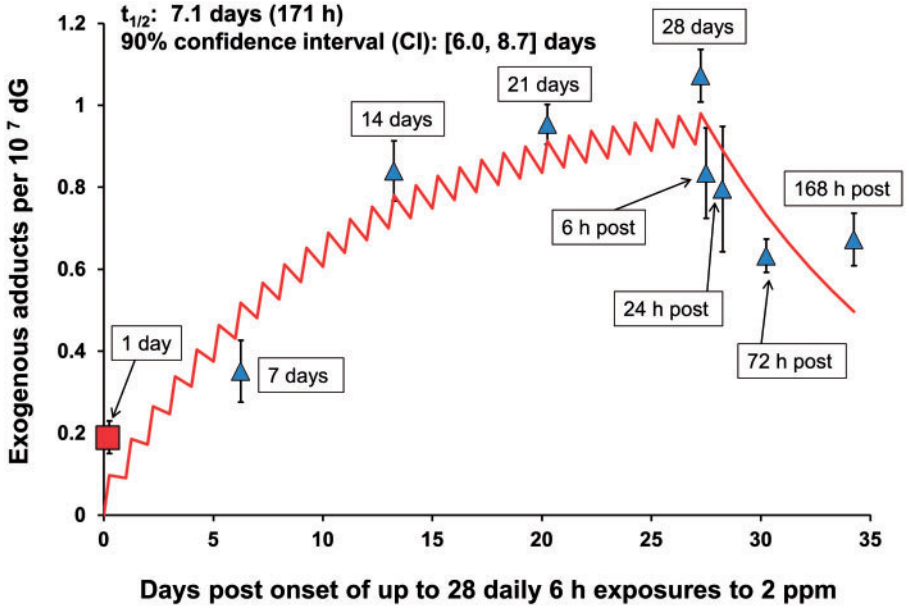

FIG. 3. Estimated time for exogenous [$^{13}CD_2$]-N^2-HOMe-dG adducts to reach the steady-state concentration and $t_{1/2}$ of exogenous [$^{13}CD_2$]-N^2-HOMe-dG adducts following a 2 ppm (6 h/day) exposure for 28 days (observed [mean ± SD] and predicted [solid line]) in rat nasal epithelium. Data from a single 6 h exposure to 2 ppm (cubic dot) were included from a previous study (Lu *et al.*, 2011), $n = 5$. The numbers of other data points (n) are indicated in Table 1. Triangle dots represent real data with error bars, and the zigzag line represents predicted trends of the accumulation and decay during the exposure and postexposure period.

TABLE 3. Formation of N^2-HOMe-dG Mono-Adducts (Mean ± SD) in Monkey Nasal Epithelium, Bone Marrow, White Blood Cells, and Trachea Exposed to 2 and 6 ppm Labeled Formaldehyde for 2 Days

Project	Monkey Tissues	Exposure Concentration	N^2-HOMe-dG (adducts/10^7 dG)		
			Endogenous	Exogenous	*n*
NHP1[a]	Nasal maxillo-turbinate	1.9 ppm	2.50 ± 0.40	0.26 ± 0.04	3
		6.1 ppm	2.05 ± 0.54	0.41 ± 0.05	3
	Scraped bone marrow	1.9 ppm	17.50 ± 2.60	n.d.	3
		6.1 ppm	12.40 ± 3.60	n.d.	3
NHP2	Nasal dorsal mucosa	6 ppm	3.62 ± 1.28	0.40 ± 0.07	4
		Air control	3.81 ± 1.19	n.d.	4
	Nasal nasopharynx	6 ppm	3.62 ± 1.34	0.33 ± 0.10	4
		Air control	3.48 ± 0.53	n.d.	4
	Nasal septum	6 ppm	3.56 ± 0.69	0.39 ± 0.15	4
		Air control	3.75 ± 0.32	n.d.	4
	Nasal anterior maxillary	6 ppm	3.80 ± 0.91	0.34 ± 0.12	4
		Air control	4.21 ± 0.53	n.d.	4
	Nasal posterior maxillary	6 ppm	3.46 ± 1.05	0.36 ± 0.16	4
		Air control	3.95 ± 0.74	n.d.	4
	Scraped bone marrow	6 ppm	11.00 ± 2.01	n.d.	10
		Air control	10.18 ± 1.35	n.d.	10
	Saline extrusion bone marrow	6 ppm	4.41 ± 1.00	n.d.	10
		Air control	5.65 ± 2.12	n.d.	4
	White blood cells	6 ppm	3.79 ± 1.19	n.d.	4
		Air control	3.64 ± 1.09	n.d.	4
	Trachea carina	6 ppm	2.33 ± 1.12	n.d.	4
		Air control	2.69 ± 0.95	n.d.	4
	Trachea proximal	6 ppm	2.50 ± 1.10	n.d.	4
		Air control	2.35 ± 1.05	n.d.	4

[a]NHP, Nonhuman primate.
n.d., not detected.

hydrolysis, rather than being actively repaired (Quievryn and Zhitkovich, 2000). This raises the important question of whether another type of DNA damage is generated through DPC hydrolysis.

Previous *in vitro* studies demonstrated DPCs were induced by formaldehyde with formation of N-CH_2-N and N-CH_2-S linkages between DNA and protein (Heck *et al.*, 1990). To study the chemical identity of formaldehyde-induced DPCs, we investigated the *in vitro* crosslinking reaction between nucleosides, nucleotides, and amino acids and peptides in the presence of formaldehyde (Lu *et al.*, 2010b). Our results demonstrate that the most abundant crosslink was formed with N-CH_2-N linkage between dG and lysine. Unfortunately, the lability of N-CH_2-N linkage in crosslinks has prevented the isolation of these compounds and characterization of their chemical structures by spectroscopic methods. On the other hand, the second most abundant crosslink was formed with N-CH_2-S linkage between dG and cysteine, and was stable enough to be characterized. It should be also noted that formaldehyde exposure is associated with oxidative stress in animals and humans, resulting in the production of reactive oxygen species that can cause DNA adducts, such as malondialdehyde-deoxyguanosine (Bono *et al.*, 2010). Although there is lack of evidence on the formation of indirect DPCs *in vivo* arising from oxidative stress after formaldehyde exposure, we cannot rule out the possibility that reactive oxygen species can interact with DNA and protein to form indirect DPCs based on their high reactivity. However, it has not been possible to know the chemical identity of formaldehyde-induced DPCs formed *in vivo* due to the lack of chemical specific and accurate measurement. In this study, dG-Me-Cys containing N-CH_2-S linkage was chosen for the degradation experiments based on its relatively high stability and abundance (Lu *et al.*, 2010b). This study also included degradation of other crosslinks, dG-Me-GSH, and dG-Me-AGT, both containing N-CH_2-S linkage. Formaldehyde has been reported to induce dG-Me-GSH following incubation of formaldehyde with reduced GSH and calf thymus DNA (Lu *et al.*, 2009). This finding raises the possibility that dG-Me-GSH can form endogenously, because both formaldehyde and GSH are present in reasonably high concentrations innately within cells (Swenberg *et al.*, 1980). The results of our investigations on the hydrolytic degradation of formaldehyde-induced DPCs demonstrate clearly that DPC hydrolysis involves breakdown to the DNA mono-adduct, N^2-HOMe-dG, followed by further degradation to dG. These results suggest that DPCs are likely to be an important source of formaldehyde-induced DNA mono-adducts and such mono-adducts represent a biomarker of both direct attack of DNA and a breakdown product of DPCs.

With the support of powerful analytical instruments and improved methodology, endogenous formaldehyde-induced N^2-HOMe-dG adducts were observed in all rat and monkey tissues analyzed. However, exogenous formaldehyde-induced [$^{13}CD_2$]-N^2-HOMe-dG adducts were only detected in the nasal respiratory epithelial DNA (with the exception of the 1/95 rat bone marrow samples) (Fig. 2; Tables 1–3). These results were in accord with our previous studies (Lu *et al.*, 2010a, 2011; Moeller *et al.*, 2011), providing compelling evidence that inhaled formaldehyde does not reach tissues distant to the site of initial contact. Thus, the plausibility of speculative hypotheses that inhaled formaldehyde causes leukemia via direct DNA damage must be seriously questioned.

The numbers of exogenous N^2-HOMe-dG adducts found in rat and monkey nasal epithelium were all substantially smaller than the corresponding numbers of endogenous adducts (Tables 1 and 3). Surprisingly, 1 bone marrow sample from the 95 exposed rats contained 0.34 exogenous N^2-HOMe-dG adducts/10^7 dG (=0.19 fMol N^2-Me-dG on column) (Tables 1 and 4). The question remains why exogenous N^2-HOMe-dG adducts were not found in the other 94 exposed rats. Considering all the bone marrow samples we have previously analyzed (Lu *et al.*, 2010a, 2011; Moeller *et al.*, 2011), this was the only instance out of 121 rats and primates where exogenous adducts were detected in bone marrow samples. Moreover, the amounts of endogenous adducts in air control rat bone marrow match very nicely with the corresponding amounts measured in the 28-day-exposed rat bone marrow.

TABLE 4. Formation of Endogenous N^2-HOMe-dG Mono-Adducts (Mean ± SD) in Various Tissues in All Rat [$^{13}CD_2$]-Formaldehyde Exposure Studies.

Rat Tissues	Endogenous N^2-HOMe-dG (adducts/10^7 dG)	*n*
Nasal epithelium	3.26 ± 0.97	147
Bone marrow[a]	3.41 ± 0.99	95
White blood cells	1.90 ± 0.50	71
Lung	2.37 ± 0.25	18
Liver	2.48 ± 0.35	18
Spleen	2.20 ± 0.35	18
Thymus	1.47 ± 0.11	18
TBLN	3.24 ± 0.98	8
Lymph nodes	2.90 ± 1.12	8
Trachea	2.91 ± 0.82	8
Kidney	2.08 ± 0.35	8
Brain	2.24 ± 0.59	8

[a]Exogenous N^2-HOMe-dG was only detectable in 1/95 rat bone marrow sample.

A statistically significant difference was found between the amounts of endogenous adducts present in the thymus of exposed and air control rats (Table 2). However, this difference (0.15 adducts/10^7 dG) is the smallest of any of the differences that were obtained in the various tissues examined, with none of the other differences being either biologically or statistically significant. We therefore consider the statistical significance of the small thymus difference to simply be a chance finding that arose as a result of the surprisingly tight SD estimates for the thymus data.

Similar findings have been demonstrated for N^6-formyllysine, a chemical homolog of biologically important N^6-acetyllysine that has recently emerged as a widespread modification of proteins. Studies conducted at Massachusetts Institute of Technology that utilized the same animal tissues have independently demonstrated that [$^{13}CD_2$]-N^6-formyllysine was only present in the rat nasal epithelium. Likewise, tissues distant to the site of contact only had endogenous N^6-formyllysine. Furthermore, accumulation of [$^{13}CD_2$]-N^6-formyllysine adducts was only detected in the nasal epithelium (Edrissi *et al.*, 2014). Similar tissue distributions were also reported in a rat dosimetry study (Edrissi *et al.*, 2013).

The 28-day rat study demonstrates both accumulation and attainment of quasi-steady-state concentrations of exogenous formaldehyde-induced [$^{13}CD_2$]-N^2-HOMe-dG adducts in nasal epithelium over a 28-day exposure period (Fig. 3). Due to the instability of N^2-HOMe-dG, its $t_{1/2}$ was modeled pharmacokinetically and estimated with the model parameters. Similar to the previous 10 ppm exposure study (Swenberg *et al.*, 2013), there was a rapid initial loss of nearly 20% of the adducts in the first 6 h postexposure. This rapid initial loss was followed by a much slower decrease in adducts with a $t_{1/2}$ estimated to be 7.1 days

(Fig. 3), which is believed to reflect DNA repair and/or spontaneous hydrolysis. In the 10 ppm exposure study, the rapid loss during the first 6-h postexposure is thought to be the result of cell death, but not DNA repair. In the current 28-day study, 2 ppm inhalation exposures to formaldehyde should have not resulted in prominent decreases in DNA mono-adducts due to cytotoxicity. Thus, the mechanism responsible for the apparent rapid initial loss of exogenous N^2-HOMe-dG remains unknown.

In the first monkey study (Moeller *et al.*, 2011), the amount of endogenous N^2-HOMe-dG was found to be extremely high in scraped bone marrow (Table 3). Thus, 2 different collection methods were used in the second monkey study, traditional saline extrusion, and direct scraping. The amount of endogenous N^2-HOMe-dG adduct present in direct scraped bone marrow was found to be at least 2-fold higher than in saline extrusion (Table 3). Because the saline extruded bone marrow is likely to contain predominantly blood, the number of DNA adducts in these samples would be diluted by the relatively low amount of DNA adducts present in blood. In contrast, the large number of DNA adducts found in direct bone marrow scraping reflects the true value of endogenous DNA damage. These data support our hypothesis that endogenous formaldehyde induced DNA damage in bone marrow could result in "spontaneous" mutations and thus may represent an important source of leukemia and bone marrow failure (BMF).

Of immediate interest to both EPA and our lab, we also investigated whether or not inhaled formaldehyde could cause an increase in the amounts of endogenous formaldehyde-induced DNA adducts. As background information, formaldehyde clearance is predominantly mediated by a saturable pathway, glutathione conjugation. Thus, additional exposures of exogenous formaldehyde might interfere with the clearance of endogenous formaldehyde. As shown in Tables 1–3, the average endogenous amount of N^2-HOMe-dG in each exposed tissue was not significantly different from the corresponding values in control tissues. We conclude, therefore, that neither a 28-day 2 ppm nor a 2-day 6 ppm [$^{13}CD_2$]-formaldehyde exposure of rats or monkeys resulted in any significant change in the number of endogenous N^2-HOMe-dG, indicating that inhaled exogenous formaldehyde at these concentrations does not induce detectable changes in the amounts of endogenous formaldehyde-induced DNA adducts. By pooling data from all rat exposure studies we have conducted (Table 4, Supplementary Table SI-4), there was no significant difference ($P = .1$ and .2) in the amount of endogenous adducts between exposed groups and non-exposed groups. It should be pointed out, however, that inhalation exposure to approximately 15 ppm formaldehyde for 28 days or longer has not been tested. Such high exposures might conceivably induce increases in endogenous DNA adducts due to depletion of glutathione.

With high amounts of endogenous DNA adducts always present, we expect that the endogenous DNA damage results in "spontaneous" mutations and represents an important source and initiator of cancer, leukemia, and other diseases (Ames, 1989; Swenberg *et al.*, 2008), especially when compared with low environmental concentrations. Many reactive aldehydes are endogenously produced within humans (Nakamura *et al.*, 2014; Voulgaridou *et al.*, 2011). Most of them cause DNA damage, and some of them accumulate due to inherited genetic mutations and detoxification deficiencies. ALDH2 is mutated in approximately 1 billion people, most frequently observed in Southeast Asians, and is commonly known as the Asian flushing syndrome. FA is a genetic disease with an incidence of 1 per 350 000 births, with a higher frequency in Ashkenazi Jews and Afrikaners in South Africa. It is the result of a genetic defect in a cluster of proteins responsible for DNA repair (Hira *et al.*, 2013) have recently found that ALDH2 deficiency dramatically accelerates BMF in Japanese FA patients. Most strikingly, those patients entirely deficient for ALDH2 developed BMF within the first 7 months of life (Hira *et al.*, 2013). Patel's laboratory has demonstrated that blocking ADH5 (formaldehyde dehydrogenase) or ALDH2 triggered DNA damage and was indeed genotoxic to hematopoietic stem cells, causing FA mice to spontaneously develop acute leukemia and profound BMF (Garaycoechea *et al.*, 2012; Garaycoechea and Patel, 2014; Langevin *et al.*, 2011). These reports confirm unequivocally that endogenous aldehydes are sufficient to cause DNA damage and induce leukemia and BMF in both humans and mice deficient in aldehyde detoxification and DNA repair.

Risk assessment practices have progressed from minimal incorporation of mechanistic data and reliance on linearly extrapolated models for low-dose risk assessment, to greater emphasis on Mode of Action and the incorporation of scientific data into the risk assessment process. Yet, some scientists and risk assessors continue to support the use of linear low-dose risk assessment over the incorporation of scientific data into the decision-making process (National Research Council, 2009). There are important questions that need to be addressed. For example, which exposure, exogenous or endogenous, is driving mutagenesis or carcinogenesis at low exposures? What is the safe exposure level of formaldehyde when substantial amounts of endogenous formaldehyde are always present *in vivo*? Results from the present study, as well as previous studies (Lu *et al.*, 2010a, 2011; Moeller *et al.*, 2011), provide critical data that improve our ability to develop science-based cancer risk assessments utilizing known biology and toxicology related to formaldehyde and cancer, rather than relying on default approaches like linear low-dose extrapolation. The NRC Review of the EPA's Draft IRIS Assessment of Formaldehyde in 2011 strongly suggested that such data should be incorporated explicitly into a revised risk assessment.

CONCLUSIONS

In summary, N^2-HOMe-dG was shown to be the only dominant product of several formaldehyde-induced DPCs investigated. These DPCs are likely to be an important source of formaldehyde-induced DNA mono-adducts that represent a biomarker of both direct attack of DNA and a breakdown product of DPCs. Possible origins of interfering artifacts were identified, minimized, or eliminated throughout the analyses. Plus, with the application of highly sensitive instruments and accurate assays, inhaled formaldehyde was found to reach nasal respiratory epithelium, but not other tissues distant to the site of initial contact in the 28-day 2 ppm exposure rat study as well as in the 2-day 6 ppm exposure monkey study. In contrast, endogenous N^2-HOMe-dG adducts were readily detected in all tissues examined with remarkably higher amounts present. Therefore, the plausibility of speculative hypotheses that inhaled formaldehyde causes leukemia must be seriously questioned. Moreover, the amounts of exogenous formaldehyde-induced N^2-HOMe-dG adducts were 3- to 8-fold and 5- to 11-fold lower than the average amounts of endogenous formaldehyde-induced N^2-HOMe-dG adducts in rat and monkey nasal respiratory epithelium, respectively. Thus, our improved understanding that ever-present endogenous formaldehyde leads to the formation of severe DNA lesions (both DPCs and DNA mono-adducts)

warrants reflection. Historically, the risk to humans from inhaled formaldehyde may have been over-stated and the role of endogenous formaldehyde in the development and progression of bone marrow toxicity and leukemia may have been under-appreciated. To this end, we have developed a conservative approach to estimating risks to the general population from endogenous and low concentrations of exogenous formaldehyde (Starr and Swenberg, 2013), and have highlighted the importance of considering endogenous sources of DPCs and DNA mono-adducts during the risk assessment process.

SUPPLEMENTARY DATA

Supplementary data are available online at http://toxsci.oxfordjournals.org/.

FUNDING

National Institutes of Health (P42 ES005948, P30 ES010126 to J.A.S.); Texas Commission for Environmental Quality (582-12-21861 to J.A.S); The Research Foundation for Health and Environmental Effects (RFHEE) and Formacare provided the support to Lovelace Respiratory Research Institute for the animal exposures. RFHEE also covered the cost of mass spectrometry analyses at UNC. TCEQ provided partial support for supplies and laboratory personnel. None of the funding organizations see any draft or submitted manuscripts from the Swenberg Laboratory until the article has been accepted for publication and is available online.

ACKNOWLEDGMENTS

The authors acknowledge the support of Valeriy Afonin for his assistance with the isolation of DNA from tissues, Leonard Collins for his assistance with HPLC purification and nano-UPLC-MS-MS, and Dr Zhengfa Zhang and Dr Avran Gold for their assistance with the synthesis of standards.

REFERENCES

Ames, B. N. (1989). Endogenous DNA damage as related to cancer and aging. *Mutat. Res.* **214,** 41–46.

Beane Freeman, L. E. B., Blair, A., Lubin, J. H., Stewart, P. A., Hayes, R. B., Hoover, R. N., and Hauptmann, M. (2009). Mortality from lymphohematopoietic malignancies among workers in formaldehyde industries: The National Cancer Institute Cohort. *J. Natl. Cancer Inst.* **101,** 751–761.

Bono, R., Romanazzi, V., Munnia, A., Piro, S., Allione, A., Ricceri, F., Guarrera, S., Pignata, C., Matullo, G., Wang, P., *et al.* (2010). Malondialdehyde-deoxyguanosine adduct formation in workers of pathology wards: The role of air formaldehyde exposure. *Chem. Res. Toxicol.* **23,** 1342–1348.

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

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

Casanova, M., Starr, T. B., and Heck, H. D. (1984). Differentiation between metabolic incorporation and covalent binding in the labeling of macromolecules in the rat nasal mucosa and bone marrow by inhaled [^{14}C]- and [^{3}H]-formaldehyde. *Toxicol. Appl. Pharmacol.* **76,** 26–44.

Casanova-Schmitz, M., and Heck, H. D. (1983). Effects of formaldehyde exposure on the extractability of DNA from proteins in the rat nasal mucosa. *Toxicol. Appl. Pharmacol.* **70,** 121–132.

Coggon, D., Ntani, G., Harris, E. C., and Palmer, K. T. (2014). Upper airway cancer, myeloid leukemia, and other cancers in a cohort of British chemical workers exposed to formaldehyde. *Am. J. Epidemiol.* **179,** 1301–1311.

Committee to Review EPA's Draft IRIS Assessment of Formaldehyde & National Research Council. (2011). *Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde*. National Academies Press, Washington, D. C.

Dunnett, C. W. (1964). New tables for multiple comparisons with a control. *Biometrics* **20,** 482–491.

Edrissi, B., Taghizadeh, K., Moeller, B. C., Kracko, D., Doyle-Eisele, M., Swenberg, J. A., and Dedon, P. C. (2013). Dosimetry of N^6-formyllysine adducts following [$^{13}C^2H_2$]-formaldehyde exposures in rats. *Chem. Res. Toxicol.* **26,** 1421–1423.

Edrissi, B., Taghizadeh, K., Moeller, B. C., Yu, R., Kracko, D., Doyle-Eisele, M., Swenberg, J. A., and Dedon, P. C. (2014). N^6-Formyllysine as a biomarker of formaldehyde exposure: Inhalation studies in rats reveal formation and accumulation of N^6-Formyllysine adducts in nasal epithelium. *The Toxicologist* **138,** 398.

Garaycoechea, J. I., Crossan, G. P., Langevin, F., Daly, M., Arends, M. J., and Patel, K. J. (2012). Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature* **489,** 571–575.

Garaycoechea, J. I., and Patel, K. J. (2014). Why does the bone marrow fail in Fanconi anemia? *Blood* **123,** 26–34.

Grafstrom, R. C., Fornace, A., and Harris, C. C. (1984). Repair of DNA damage caused by formaldehyde in human cells. *Cancer Res.* **44,** 4323–4327.

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

Hira, A., Yabe, H., Yoshida, K., Okuno, Y., Shiraishi, Y., Chiba, K., Tanaka, H., Miyano, S., Nakamura, J., Kojima, S., *et al.* (2013). Variant ALDH2 is associated with accelerated progression of bone marrow failure in Japanese Fanconi anemia patients. *Blood* **122,** 3206–3209.

International Agency for Research on Cancer. (2006). Formaldehyde, 2-Butoxyethanol and 1-tert-Butoxypropan-2-ol. In: *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* **Vol. 88,** 1–287.

International Agency for Research on Cancer. (2012). Formaldehyde, A review of human carcinogens. Part F: *Chemical Agents and Related Occupations*, **Vol. 100**, 401–435.

Langevin, F., Crossan, G. P., Rosado, I. V., Arends, M. J., and Patel, K. J. (2011). Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* **475,** 53–58.

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

Lu, K., Moeller, B. C., Doyle-Eisele, M., McDonald, J., and Swenberg, J. A. (2011). Molecular dosimetry of N^2-hydroxymethyl-dG DNA adducts in rats exposed to formaldehyde. *Chem. Res. Toxicol.* **24,** 159–161.

Lu, K., Ye, W., Gold, A., Ball, L. M., and Swenberg, J. A. (2009). Formation of S-[1-(N^2-deoxyguanosinyl) methyl]glutatione between glutathione and DNA induced by formaldehyde. *J. Am. Chem. Soc.* **131,** 3414–3415.

Lu, K., Ye, W., Zhou, L., Collins, L. B., Chen, X., Gold, A., Ball, L. M., and Swenberg, J. A. (2010b). Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers. *J. Am. Chem. Soc.* **132,** 3388–3399.

Magana-Schwencke, N., and Ekert, B. (1978). Biochemical analysis of damage induced in yeast by formaldehyde. II. Induction of cross-links between DNA and protein. *Mutat. Res.* **51,** 11–19.

McGhee, J. D., and von Hippel, P. H. (1977). Formaldehyde as a probe of DNA structure. r. Mechanism of the initial reaction of formaldehyde with DNA. *Biochemistry* **16,** 3276–3293.

Moeller, B. C., Lu, K., Doyle-Eisele, M., McDonald, J., Gigliotti, A., and Swenberg, J. A. (2011). Determination of N^2-hydroxymethyl-dG adducts in nasal epithelium and bone marrow of non-human primates following $^{13}CD_2$-formaldehyde inhalation exposure. *Chem. Res. Toxicol.* **24,** 162–164.

Nakamura, J., Mutlu, E., Sharma, V., Collins, L. B., Bodnar, W., Yu, R., Lai, Y., Moeller, B. C., Lu, K., and Swenberg, J. A. (2014). The endogenous exposome. *DNA Repair* **19,** 3–13.

National Research Council. (2009). *Science and decisions: Advancing risk assessment*. National Academies Press, Washington, D. C.

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

Rosado, I. V., Langevin, F., Crossan, G. P., Takata, M., and Patel, K. J. (2011). Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nat. Struct. Mol. Biol.* **18,** 1432–1434.

Shaham, J., Bomstein, Y., Gurvich, R., Rashkovsky, M., and Kaufman, Z. (2003). DNA-protein crosslinks and p53 protein expression in relation to occupational exposure to formaldehyde. *Occup. Environ. Med.* **60,** 403–409.

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

Shoulkamy, M. I., Nakano, T., Ohshima, M., Hirayama, R., Uzawa, A., Furusawa, Y., and Ide, H. (2012). Detection of DNA-protein crosslinks (DPCs) by novel direct fluorescence labeling methods: Distinct stabilities of aldehyde and radiation-induced DPCs. *Nucleic Acids Res.* **40,** e143.

Solomon, M. J., and Varshavsky, A. (1985). Formaldehyde-mediated DNA-protein crosslinking: A probe for in vivo chromatin structures. *Proc. Natl. Acad. Sci. U.S.A.* **82,** 6470–6474.

Starr, T. B., and Swenberg, J. A. (2013). A novel bottum-up approach to bounding low-dose human cancer risks from chemical exposures. *Regul. Toxicol. Pharmacol.* **65,** 311–315.

Stingele, J., Schwarz, M. S., Bloemeke, N., Wolf, P. G., and Jentsch, S. (2014). A DNA-depedent protease involved in DNA-protein crosslink repair. *Cell* **158,** 327–338.

Swenberg, J. A., Fryar-Tita, E., Jeong, Y. C., Boysen, G., Starr, T. B., Walker, V. E., and Albertini, R. J. (2008). Biomarkers in toxicology and risk assessment: Informing critical dose-response relationships. *Chem. Res. Toxicol.* **21,** 253–265.

Swenberg, J. A., Kerns, W. D., Mitchell, R. I., Gralla, E. J., and Pavkov, K. L. (1980). Induction of squamous cell carcinomas of the rat nasal cavity by inhalation exposure to formaldehyde vapor. *Cancer Res.* **40,** 3398–3402.

Swenberg, J. A., Moeller, B. C., Lu, K., Rager, J. E., Fry, R. C., and Starr, T. B. (2013). Formaldehyde carcinogenicity research: 30 years and counting for mode of action, epidemiology, and cancer risk assessment. *Toxicol. Pathol.* **41,** 181–189.

Voulgaridou, G. P., Anestopoulos, I., Franco, R., Panayiotidis, M. I., and Pappa, A. (2011). DNA damage induced by endogenous aldehydes: Current state of knowledge. *Mutat. Res.* **711,** 13–27.

Ye, X., Ji, Z., Wei, C., McHale, C. M., Ding, S., Thomas, R., Yang, X., and Zhang, L. (2013). Inhaled formaldehyde induces DNA-protein crosslinks and oxidative stress in bone marrow and other distant organs of exposed mice. *Environ. Mol. Mutagen.* **54,** 705–718.

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