

## ORIGINAL ARTICLE

# Adductomic signatures of benzene exposure provide insights into cancer induction

Hasmik Grigoryan<sup>1</sup>, William M.B. Edmands<sup>1</sup>, Qing Lan<sup>2</sup>, Henrik Carlsson<sup>1</sup>, Roel Vermeulen<sup>3</sup>, Luoping Zhang<sup>1</sup>, Song-Nian Yin<sup>4</sup>, Gui-Lan Li<sup>4</sup>, Martyn T. Smith<sup>1</sup>, Nathaniel Rothman<sup>2</sup> and Stephen M. Rappaport<sup>1,\*</sup>

<sup>1</sup>Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA 94720, USA,

<sup>2</sup>Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer

Institute, National Institutes of Health, Department of Health and Human Service, Rockville, MD 20850, USA, <sup>3</sup>Division of

Environmental Epidemiology, Institute for Risk Assessment Sciences, Utrecht University, 3508 TD Utrecht, The Netherlands and <sup>4</sup>National Institute of Occupational Health and Poison Control, Chinese Center for Disease Control and Prevention, 29

Nan Wei Road, Beijing 100050, China

\*To whom correspondence should be addressed: Tel: +1 510 642-4355; Fax: +1 510 642-8815; Email: [srappaport@berkeley.edu](mailto:srappaport@berkeley.edu)

## Abstract

Although benzene has long been recognized as a cause of human leukemia, the mechanism by which this simple molecule causes cancer has been problematic. A complicating factor is benzene metabolism, which produces many reactive intermediates, some specific to benzene and others derived from redox processes. Using archived serum from 20 nonsmoking Chinese workers, 10 with and 10 without occupational exposure to benzene (exposed: 3.2–88.9 ppm, controls: 0.002–0.020 ppm), we employed an adductomic pipeline to characterize protein modifications at Cys34 of human serum albumin, a nucleophilic hotspot in extracellular fluids. Of the 47 measured human serum albumin modifications, 39 were present at higher concentrations in benzene-exposed workers than in controls and many of the exposed-control differences were statistically significant. Correlation analysis identified three prominent clusters of adducts, namely putative modifications by benzene oxide and a benzene diolepoxide that grouped with other measures of benzene exposure, adducts of reactive oxygen and carbonyl species, and Cys34 disulfides of small thiols that are formed following oxidation of Cys34. Benzene diolepoxides are potent mutagens and carcinogens that have received little attention as potential causes of human leukemia. Reactive oxygen and carbonyl species—generated by redox processes involving polyphenolic benzene metabolites and by Cyp2E1 regulation following benzene exposure—can modify DNA and proteins in ways that contribute to cancer. The fact that these diverse human serum albumin modifications differed between benzene-exposed and control workers suggests that benzene can increase leukemia risks via multiple pathways involving a constellation of reactive molecules.

## Introduction

Benzene is an important industrial chemical, a prominent constituent of petroleum fuels, and a ubiquitous component of combustion products, including tobacco smoke (1). Although benzene is a well-known cause of hematotoxicity and myeloid leukemias in humans, the mechanism by which this simple molecule causes health effects has been elusive (2,3). A complicating factor is benzene metabolism, which gives rise to

numerous electrophilic species that can modify DNA and functional proteins (4). Some of these electrophiles are benzene-specific metabolites, including benzene oxide (BO), benzoquinones (BQs), muconaldehydes and benzene diolepoxides (BDEs) (5). In addition, reactive oxygen species (ROS), produced from oxidative metabolism of benzene via CYP2E1 (6) and futile cycling of diol and quinone forms of polyphenolic benzene metabolites (7),

Received: December 17, 2017; Revised: March 1, 2018; Accepted: March 8, 2018

© The Author(s) 2018. Published by Oxford University Press. All rights reserved. For Permissions, please email: [journals.permissions@oup.com](mailto:journals.permissions@oup.com).

**Abbreviations**

aBen	benzene in breathing-zone air
BO	benzene oxide
BO-ALB	HSA modified by benzene oxide
BDE	benzene diolepoxides
BQ	benzoquinones
BQ-ALB	HSA modified by benzoquinones
EDTA	ethylenediamine-tetraacetic acid
GC-MS	gas chromatography-mass spectrometry
Hb	hemoglobin
HSA	human serum albumin
nLC-HRMS	nano-liquid chromatography-high-resolution mass spectrometry
PAR	peak-area ratio
RCS	reactive carbonyl species
ROS	reactive oxygen species
TCEP	tris(2-carboxyethyl)phosphine.

can react with macromolecules and also oxidize microsomal lipids to produce reactive carbonyl species (RCS) that modify proteins and DNA (8).

Because reactive electrophiles cannot be measured directly in blood, investigators have studied their dispositions *in vivo* by tracking modifications to abundant proteins, particularly hemoglobin (Hb) and human serum albumin (HSA). Although most assays have targeted particular adducts of Hb and HSA (9), recent work has explored untargeted avenues for characterizing adductomes (10–12). Our laboratory has developed an adductome pipeline to investigate modifications at the highly nucleophilic Cys34 residue of HSA (11), which is ubiquitous in extracellular fluids. We focused on Cys34, not only because its sulfhydryl group efficiently scavenges small reactive electrophiles in the extracellular space (13–15) but also because its oxidation generates a host of covalent modifications to circulating thiols that act as redox switches in homeostatic processes (16–20). After tryptic digestion of HSA, Cys34 adducts are present as modifications of the third largest peptide (T3-peptide, <sup>21</sup>ALVLI AFAQYLQQC<sup>34</sup>PFEDHVK<sup>41</sup>) that are detected by nano-liquid chromatography-high-resolution mass spectrometry (nLC-HRMS) (11).

We performed Cys34 adductomics with archived HSA from 10 nonsmoking workers exposed to benzene and 10 nonsmoking local controls from a large study of benzene biomarkers conducted in China (21). Three prominent clusters of correlated adducts were observed: (i) products of benzene metabolites that grouped with other measures of benzene exposure, (ii) adducts of ROS and RCS and (iii) Cys34 disulfides of small thiols. Interestingly, concentrations of 16 modifications were demonstrably elevated in benzene-exposed workers, including Cys34 oxidation products and putative adducts of BO, BDE and RCS. BDEs, particularly the *anti*-isomers, are mutagens (22) and carcinogens (23) that have received scant attention as causes of human cancers, and ROS and RCS modify DNA and proteins in ways that can contribute to tumor formation (24–26). This set of diverse adducts suggests that benzene exposure can increase cancer risks via multiple pathways involving a constellation of reactive intermediates.

**Materials and methods****Benzene-exposed and control subjects**

Archived HSA was obtained from a study of Cys34 adducts of BO and 1,2-BQ and 1,4-BQ in workers from Chinese factories where benzene was

used and was not used (27). Serum had been collected with informed written consent under protocols approved by Human Subjects Institutional Review Boards from all participating institutions (21). Exposures to benzene had been carefully documented by repeated personal air sampling over 1 year and exposures to additional hydrocarbons, other than toluene, were found to be trivial (28).

In the original study, serum had been purified by dialysis followed by addition of ammonium sulfate to precipitate immunoglobulins, and 5 mg portions of HSA had been derivatized to volatile fluorinated derivatives of the targeted Cys34 adducts of BO and the BQs (i.e. BO-Alb, 1,2-BQ-Alb and 1,4-BQ-Alb) that were measured by gas chromatography-mass spectrometry (GC-MS) in negative-ion-chemical-ionization mode (27). The remaining HSA was dissolved in water at 25 mg/ml and stored at –80°C. For the current investigation, samples of archived HSA were selected from 10 nonsmoking workers exposed to benzene in shoe factories (where benzene was a component of glues) and 10 nonsmoking local control workers who worked in a clothing factory where solvents were not used. Groups of both exposed and control workers consisted of eight females and two males of comparable ages (34.2 ± 9.84 years and 30.2 ± 9.54 years for mean ± SD of exposed and control subjects, respectively). To maximize contrast, exposed subjects were selected from those having air benzene exposures greater than 3 ppm and full sets of covariates. Benzene exposures were estimated as geometric mean air concentrations from repeated personal air samples in exposed workers (range: 3.2–88.9 ppm, median = 12.3 ppm) and as predicted from repeated measurements of urinary benzene in control workers (range: 0.002–0.020 ppm, median = 0.003 ppm) (29).

**Chemicals and reagents**

Triethylammonium bicarbonate buffer, ethylenediamine-tetraacetic acid (EDTA, anhydrous), porcine trypsin, acetonitrile (Ultra Chromasolv, LC-MS grade), acetic acid (LCMS grade), dimethyl sulfoxide, acrolein (≥ 99%, anhydrous) and tiglic aldehyde (≥ 96%) were obtained from Sigma-Aldrich (St. Louis, MO, USA). Methanol (Optima, LCMS grade), tris(2-carboxyethyl)phosphine (TCEP), iodoacetamide, and formic acid (Optima, LCMS grade) were purchased from Fisher Scientific (Pittsburgh, PA, USA). Crotonaldehyde (≥ 99%), ethyl vinyl ketone (>98%), and propionyl bromide (95%) were obtained from Acros Organics (Morris Plains, Morris, NJ, USA). Methyl vinyl ketone (90%) was from Alfa Aesar (Haverhill, MA, USA). The following chemicals were custom synthesized: the T3 peptide with sequence ALVLI AFAQYLQQCPFEDHVK (>97%, Biomatik, Wilmington, DE, USA), isotopically modified T3 with sequence AL-[<sup>15</sup>N, <sup>13</sup>C-Val]-LIAFAQYLQQCPFEDH-[<sup>15</sup>N, <sup>13</sup>C-Val]-K. Purified water (18.2 mΩ cm resistivity at 25°C) was prepared with a PureLab purification system (Elga LabWater, Woodridge, IL, USA).

**Analysis of adducts**

Aliquots containing 0.1 mg of archived HSA from the previous investigation were analyzed after storage at –80°C for 15 years. Samples were digested with trypsin via pressure cycling and analyzed by nLC-HRMS as described previously (11). Duplicate assays were performed for half of the samples and duplicate injections were made for all. Modifications to the T3-peptide, which contains Cys34, <sup>21</sup>ALVLI AFAQYLQQC<sup>34</sup>PFEDHVK<sup>41</sup>, were located in MS2 spectra of triply charged precursor ions with  $m/z = 811.7594 \pm \Delta m/z$ , where  $\Delta m/z$  represents the difference in mass from the T3-thiolate ion. **Supplementary Figures 1 and 2**, available at *Carcinogenesis* Online display MS2 spectra and selected ion chromatograms for the 13 adducts that had not been previously reported (11,30). For quantitation and to adjust each sample for the amount of digested HSA, the tryptic peptide adjacent to T3, i.e. <sup>42</sup>LVNEVTEFAK<sup>51</sup> (doubly charged precursor ion at  $m/z = 575.3111$ ) was used as a housekeeping peptide. The ratio of peak areas of each modified T3 peptide to the housekeeping peptide (PAR) was shown to be a robust linear measure of the adduct concentration over a wide dynamic range (11).

**Adduct annotations and comparisons with reference standards**

Identifications of adducts relied on reference standards of T3 modifications from our previous studies (11,30) as well as new standards generated here by reactions of the following electrophiles with fresh human serum: acrolein, crotonaldehyde, ethylvinylketone, methylvinylketone,

propionyl bromide and tiglic aldehyde. These new adducts were generated by incubation overnight with serum from a healthy volunteer subject (pH ~7.4 and 37°C) at 1:1, 1:10 and 1:100 molar ratios with continuous mixing. The excess electrophiles were removed using 30K MWCO spin columns (Millipore Sigma, MA, USA). The modified serum samples were precipitated with 60% methanol followed by dilution (1:5) with digestion buffer, digestion with trypsin and nLC-HRMS as described above. The monoisotopic masses for all T3 modifications were extracted from total ion chromatograms using a mass tolerance of 5 ppm. Because the acrolein/propionyl bromide, crotonaldehyde/methylvinylketone and tiglic aldehyde/ethylvinylketone pairs share the same added masses, corresponding to empirical formulae  $C_3H_5O$ ,  $C_4H_7O$  and  $C_5H_9O$ , respectively, the retention times were used for confirmation. This led to unambiguous identification of crotonaldehyde and tiglic aldehyde as the added masses representing formulae  $C_4H_7O$  and  $C_5H_9O$ , respectively. [We had previously putatively annotated the  $C_4H_7O$  adduct as methylvinylketone (11). Here, we used retention times to confirm that 835.11 was, in fact, an adduct of crotonaldehyde.] However, identification of the electrophile that generated the *in vivo* adduct with empirical formula  $C_3H_5O$  (830.44) did not match those of T3 modifications by acrolein and propionyl bromide.

To provide additional evidence concerning putative annotations of Cys34 mixed disulfides, we compared T3 adducts in protein digests with and without addition of TCEP, which reduces all disulfide bonds. Seven validation samples from benzene exposed and control subjects that were heavily populated with T3 adducts were treated with the standard protocol except that TCEP was added to a final concentration of 2 mM prior to digestion. (Two minor adducts, i.e. 820.09 and 830.41, were not present in these samples.) Virtually all of the adducts that contained sulfur in their elemental compositions disappeared, adding further confirmation that they were Cys34 disulfides. Several additional adducts for which elemental compositions were not available also disappeared and are, therefore likely to be disulfides.

## Statistical analysis

Statistical analyses were performed with SAS for Windows (v.9.4, SAS Systems, Cary, NC, USA). Linear mixed models were used to adjust log-transformed PARs from replicate assays and injections for batch effects and to generate best linear unbiased predictors of subject-specific adduct levels that were used for statistical analyses (11,30). Non-detected samples were imputed a PAR of  $4 \times 10^{-6}$  based on the smallest values detected across all adducts. Wilcoxon rank-sum (exact) tests were performed to determine whether PARs for each adduct differed between benzene-exposed and control workers, with Bonferroni correction for multiple comparisons ( $n = 46$ ). One adduct (859.41) was not tested because PARs did not vary across subjects (intra-class correlation coefficient = 0). Spearman correlation coefficients of PARs were estimated pairwise across adducts and covariates. Correlation networks were visualized with Cytoscape (31).

## Results

### Adducts detected in archived HSA

As shown in Table 1, a total of 47 modifications to the T3 peptide were detected in nonsmoking workers from Chinese workplaces where benzene was used and was not used. Accurate masses led to reasonable elemental compositions within 5 ppm for 39 adducts (added masses between -46 and 489 Da relative to the T3 peptide with the Cys34 thiolate ion). We had previously detected 34 of these T3 modifications in serum/plasma from two other studies, composed of healthy smokers and nonsmokers in the USA (11) and nonsmoking women exposed to indoor combustion products and local controls in China (30) (footnoted in Table 1). This points to a pool of modifications to the T3-peptide that arise from common precursor molecules, including ROS, RCS and small thiols, in healthy humans. The remaining 13 adducts were unique to the current study, and two putative modifications, theoretically corresponding to adducts of BO [837.10, representing S-phenylation (32)] and BDE [(854.44 (5)), were only

observed in benzene-exposed workers. Overall, putative annotations were gleaned for 37 adducts, of which 14 were confirmed by comparisons with reference standards, including those of crotonaldehyde and tiglic aldehyde. Although we had suspected that 830.44 was the product of reaction between Cys34 and acrolein, a known product of lipid peroxidation, the retention time of this *in vivo* adduct did not match that of our synthetic acrolein modification. We refer to 830.44 as an unknown adduct with elemental composition of  $C_3H_5O$ , consistent with modification of Cys34 by an aldehyde.

Regarding quantitation, Table 1 shows the estimated mean values of PARs for the 46 adducts measured in this study and the corresponding approximate concentrations in pmol adduct/mg HSA (11). A 79000-fold range of adduct concentrations was observed (0.018–1426 pmol/mg HSA). Because these concentrations are derived from PARs of untargeted analytes, rather than from internal standards for known targets, we emphasize their approximate nature.

### Adducts that discriminate for benzene exposure

Subject-specific PARs were compared between benzene-exposed and control subjects with Wilcoxon rank-sum (exact) tests. Figure 1 depicts a volcano plot of the 46 quantitated adducts (Table 1), and Supplementary Table 1, available at Carcinogenesis Online provides summary statistics and boxplots of PARs stratified by exposure status. Interestingly, 39 of these 46 adducts were present at higher concentrations in the benzene-exposed workers than in the controls. And despite the modest numbers of subjects, 17 of these differences were significant at a  $P$ -value  $< 0.05$ , and five differences remained significant after Bonferroni correction for multiple comparisons ( $P < 0.0011$ ,  $n = 46$ ).

The discriminating adducts represent Cys34 reactions with electrophiles that were generated both directly and indirectly from benzene metabolism. Of these 17 modifications, all but one (808.73) were present at higher concentrations in benzene-exposed subjects. The most highly associated modifications ( $P < 0.0011$ ) included a putative benzene-specific adduct of BDE (854.44), three products of Cys34 oxidation (816.42, 822.42 and 827.76) and a labile T3-modification (811.76<sub>1</sub>) that is probably a product of electrospray dissociation in the HRMS (11,30). [Note that the monooxidation product (816.42) is detected as a sulfenamide adduct that results from dehydration and intrapeptide cyclization of the Cys34 sulfenic acid (-SOH) (33)]. The other adducts for which statistical evidence points to differences between exposed and unexposed subjects ( $P < 0.05$ ) include: the putative S-phenylation product (837.10, from reaction with BO), two products of oxidized sulfur species (833.08 and 849.07), four likely products of RCS (826.10, 830.44, 835.11 and 839.78), a rearrangement representing conversion of Cys34 to dehydroalanine (Dha, 800.43), and Cys34 disulfides of mercaptoacetic acid (841.75) and homocysteine (hCys, 850.10 with loss of  $H_2O$  and 856.43 with  $NH_2 \rightarrow OH$ ). Two additional adducts associated with benzene exposure are unknown, i.e. 808.73 (a T3 modification at a locus other than Cys34) and 965.49 (probably a disulfide).

### Correlation network

Figure 2 depicts a correlation network for all pairs of adducts and covariates having Spearman coefficients  $\geq 0.72$ . The covariates represent independent measures of benzene exposure from the initial investigation, i.e. benzene in breathing-zone air (aBen) and targeted Cys34 adducts of BO (BO-Alb) and 1,2-BQ (1,2-BQ-Alb) and 1,4-BQ (1,4-BQ-Alb) (27). Three clusters of correlated adducts are apparent. Cluster 1 represents putative

**Table 1.** Adducts detected in HSA from subjects with and without occupational exposure to benzene

Adduct	Ret. time (min)	MIM observed (m/z, +3)	MIM theoretical (m/z, +3)	$\Delta$ Mass (ppm)	Added mass (Da)	Elemental composition of added mass	Annotation	PAR $\times$ 1000	Concentration (pmol/mg HSA)
796.43 <sup>a,b</sup>	26.93	796.4312	796.4301	-1.31	-45.9913	-CH <sub>2</sub> S	Cys34→Gly	0.520	2.300
800.43 <sup>a,b</sup>	27.60	800.4317	800.4301	-1.97	-33.9873	-SH <sub>2</sub>	Cys34→Dehydroalanine	0.084	0.370
805.76 <sup>a,b,c</sup>	26.58	805.7632	805.7618	-1.75	-17.9965	-SH <sub>2</sub> , + O	Cys34→Oxalanine	0.164	0.727
808.73 <sup>a,b</sup>	27.15	808.7299			-9.0923		Not Cys34 adduct	0.036	0.160
811.76_1 <sup>a,b</sup>	27.07	811.7605	811.7594	-1.44	1.0072	+H	T3 labile adduct	2.553	11.295
811.76_2 <sup>a,b,c</sup>	27.43	811.7614	811.7594	-2.48	1.0097	+H	Unmodified T3 <sup>d</sup>	7.772	34.381
811.43 <sup>a,b</sup>	28.92	811.4257	811.4234	-2.87	2431.2480	+C <sub>114</sub> H <sub>172</sub> N <sub>27</sub> O <sub>30</sub> S	T3 dimer <sup>d</sup>	0.928	4.103
816.42 <sup>a,b,c</sup>	26.74	816.4201	816.4191	-1.18	13.9786	-H <sub>2</sub> , +O	S-Monooxidation <sup>d</sup>	2.656	11.751
816.43 <sup>a,b,c</sup>	27.64	816.4324	816.4312	-1.42	15.0228	+CH <sub>3</sub>	Methylation (not at Cys34)	0.351	1.553
820.09 <sup>a,e</sup>	27.59	820.0919	820.0911	-0.96	26.0013	+CN	S-Cyanylation	0.043	0.189
822.42 <sup>a,b,c</sup>	26.52	822.4238	822.4226	-1.36	32.9969	+HO <sub>2</sub>	S-Dioxidation <sup>d</sup>	2.864	12.672
825.76	26.74	825.7644	825.7629	-1.84	43.0188	+C <sub>2</sub> H <sub>3</sub> O	S-Acetylation	0.043	0.191
826.10 <sup>c</sup>	26.74	826.099	826.0988	-0.22	44.0226	+C <sub>2</sub> H <sub>4</sub> O	Unknown (likely S-addition of an aldehyde)	0.198	0.877
827.10	27.46	827.0958	827.0945	-1.51	47.0129	+CH <sub>3</sub> O <sub>2</sub>	S-(O)-O-CH <sub>3</sub>	0.185	0.817
827.09 <sup>b</sup>	28.02	827.0901	827.0886	-1.82	46.9959	+CH <sub>3</sub> S	S-methylthiolation	0.257	1.136
827.76 <sup>a,b,c</sup>	26.87	827.7552	827.7543	-1.09	48.9911	+HO <sub>3</sub>	S-trioxidation <sup>d</sup>	1.100	4.867
830.41 <sup>e</sup>	26.45	830.4059			56.9433		Unknown	0.018	0.082
830.44 <sup>c</sup>	27.33	830.4359	830.4348	-1.38	57.0333	+C <sub>3</sub> H <sub>5</sub> O	Unknown (likely S-addition of an aldehyde)	0.373	1.648
833.08	27.83	833.0813	833.0800	-1.60	64.9696	+HO <sub>2</sub> S	S-Addition of SO <sub>2</sub>	0.121	0.533
835.11 <sup>a</sup>	27.55	835.1082	835.1066	-1.82	71.0501	+C <sub>4</sub> H <sub>7</sub> O	S-Addition of crotonaldehyde <sup>d</sup>	0.428	1.893
837.10 <sup>c</sup>	27.36	837.1041	837.1031	-1.21	77.0380	+C <sub>6</sub> H <sub>5</sub>	S-Phenylation	0.047	0.209
839.78 <sup>c</sup>	27.70	839.7797	839.7785	-1.39	85.0647	+C <sub>5</sub> H <sub>9</sub> O	S-Addition of tiglic aldehyde <sup>d</sup>	0.173	0.767
841.75 <sup>a,b</sup>	27.28	841.7531	841.7519	-1.45	90.9849	+C <sub>2</sub> H <sub>3</sub> O <sub>2</sub> S	S-Addition of mercaptoacetic acid	0.511	2.262
845.42 <sup>a,b</sup>	26.44	845.4249	845.4239	-1.26	102.0004	+C <sub>3</sub> H <sub>4</sub> NOS	S-Addition of Cys (-H <sub>2</sub> O)	1.443	6.383
847.10	26.66	847.0963	847.0956	-0.78	107.0145	+C <sub>2</sub> H <sub>5</sub> O <sub>2</sub> S	S-methylethylsulfonylation	0.551	2.440
849.07 <sup>a,b</sup>	27.36	849.0707	849.0690	-2.02	112.9377	+HO <sub>3</sub> S <sub>2</sub>	S-Addition of S <sub>2</sub> O <sub>3</sub> H	0.116	0.515
850.10 <sup>a,b</sup>	27.18	850.0973	850.0958	-1.81	116.0175	+C <sub>4</sub> H <sub>6</sub> NOS	S-Addition of hCys (-H <sub>2</sub> O)	0.134	0.591
851.43 <sup>a,b</sup>	25.96	851.4289	851.4274	-1.77	120.0123	+C <sub>3</sub> H <sub>5</sub> NO <sub>2</sub> S	S-Addition of Cys <sup>d</sup>	322.3	1426
851.76 <sup>a,b</sup>	26.89	851.7574	851.7554	-2.38	120.9979	+C <sub>3</sub> H <sub>5</sub> O <sub>3</sub> S	S-Addition of Cys (NH <sub>2</sub> →OH)	1.120	4.955
854.44 <sup>c</sup>	25.91	854.4428	854.4418	-1.13	129.0539	+C <sub>6</sub> H <sub>9</sub> O <sub>3</sub>	S-Addition of BDE	0.305	1.350
856.10_1 <sup>a,b</sup>	25.77	856.0997	856.0993	-0.50	134.0247	+C <sub>3</sub> H <sub>5</sub> NO <sub>2</sub> S	S-Addition of hCys <sup>d</sup>	15.67	69.32
856.10_2 <sup>a,b</sup>	26.14	856.1001	856.0993	-0.85	134.0256	+C <sub>4</sub> H <sub>8</sub> NO <sub>2</sub> S	S-Addition of hCys <sup>d</sup>	2.198	9.724
856.43	27.15	856.4287	856.4273	-1.66	135.0117	+C <sub>4</sub> H <sub>7</sub> O <sub>3</sub> S	S-Addition of hCys (NH <sub>2</sub> →OH)	0.073	0.322
857.1 <sup>a,b</sup>	26.78	857.1003	857.0992		137.0264		Unknown	1.850	8.184
858.75 <sup>a,b</sup>	25.60	858.7538	858.7547	1.03	141.9871	+C <sub>3</sub> H <sub>5</sub> NO <sub>2</sub> SNa	Na Adduct of 851.43	1.937	8.570
859.41	25.54	859.4085			143.9513		Unknown	NQ	NQ
862.09 <sup>a</sup>	26.08	862.0859			151.9833		Not Cys34 adduct	0.452	1.999
864.08 <sup>a,b</sup>	25.58	864.0767			157.9558		Not Cys34 adduct	1.804	7.982
864.43 <sup>a,b</sup>	26.22	864.4318	864.4310	-0.93	159.0211	+C <sub>5</sub> H <sub>7</sub> N <sub>2</sub> O <sub>2</sub> S	S-Addition of CysGly (-H <sub>2</sub> O)	0.372	1.644
870.44 <sup>a,b</sup>	25.30	870.4356	870.4345	-1.20	177.0324	+C <sub>5</sub> H <sub>9</sub> N <sub>2</sub> O <sub>3</sub> S	S-Addition of CysGly <sup>d</sup>	37.97	168.0
873.43 <sup>c</sup>	27.22	873.431			186.0187		Unknown	0.192	0.851
883.08 <sup>b</sup>	25.36	883.0836	883.0865	3.28	214.9764	+C <sub>5</sub> H <sub>8</sub> N <sub>2</sub> O <sub>3</sub> SK	K Adduct of 870.44	0.223	0.987
894.44 <sup>a,b</sup>	25.79	894.4426	894.4416	-1.09	249.0533	+C <sub>8</sub> H <sub>13</sub> N <sub>2</sub> O <sub>5</sub> S	S-Addition of GluCys <sup>d</sup>	1.599	7.073
913.45 <sup>a,b</sup>	25.79	913.4449	913.4487	-0.31	306.0727	+C <sub>10</sub> H <sub>16</sub> N <sub>3</sub> O <sub>6</sub> S	S-Addition of GSH <sup>d</sup>	1.585	7.012
931.82 <sup>b</sup>	24.71	931.8216			361.1905		Unknown	0.063	0.278
965.49 <sup>b</sup>	24.45	965.4915			462.2003		Unknown	1.218	5.387
974.51 <sup>b</sup>	24.96	974.5063			489.2445		Unknown	0.018	0.079

GSH, glutathione; hCys, homocysteine; MIM, monoisotopic mass; NQ, not quantified; TCEP, tris(2-carboxyethyl)phosphine.

<sup>a</sup>Also detected by Grigoryan et al. (11).

<sup>b</sup>Also detected by Lu et al. (30).

<sup>c</sup>Adduct was detected in samples that were reduced with TCEP prior to digestion.

<sup>d</sup>Annotation confirmed with a synthetic standard.

<sup>e</sup>Adduct was not present in samples that were tested for reduction with TCEP.

benzene-specific adducts derived from BO and BDE (837.10 and 854.44, respectively) that were correlated with aBen and targeted HSA adducts of BO and BQs from the previous investigation.

Cluster 2 is composed primarily of adducts of ROS and RCS and is dominated by nodes representing the three Cys34 oxidation products (816.42, 822.42 and 827.76) plus prominent truncations

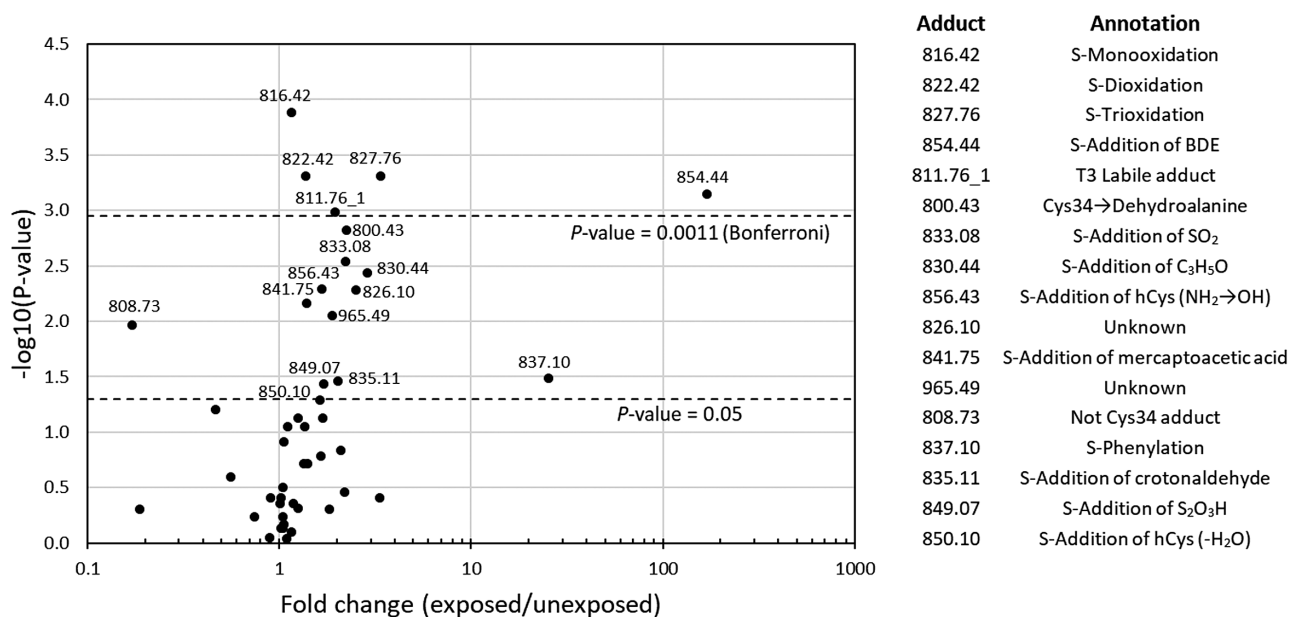


Figure 1. Volcano plot of the 47 adduct features listed in Table 1. The y-axis depicts the  $-\log_{10}(P\text{-value})$  for differences in adduct concentrations between benzene-exposed and control subjects. hCys, homocysteine.

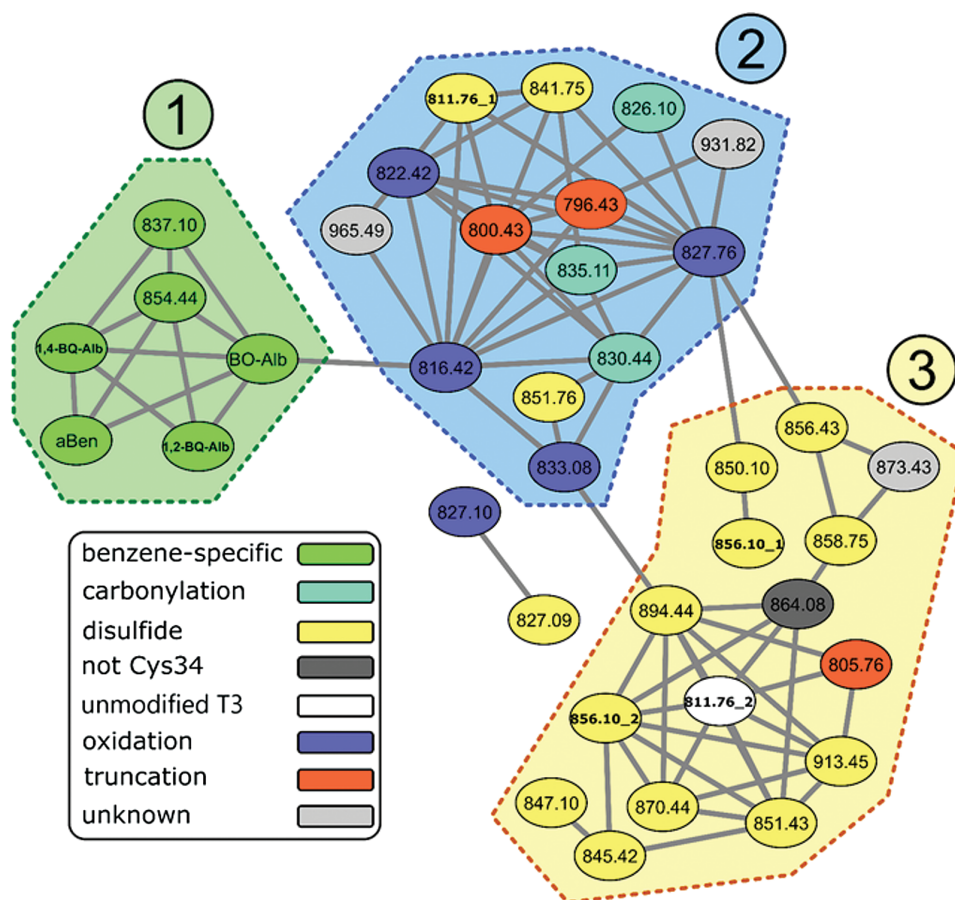


Figure 2. Correlation network for all pairs of adducts and covariates having Spearman coefficients  $\geq 0.72$ . The covariates represent independent measures of benzene exposure from the initial investigation, i.e. benzene in breathing-zone air (aBen) and targeted Cys34 adducts of BO (BO-Alb) and 1,2-benzoquinone (1,2-BQ-Alb) and 1,4-benzoquinone (1,4-BQ-Alb) (27) (see Table 1 for adduct annotations). Correlation networks were visualized with Cytoscape.

and rearrangements [796.43 (Cys34→Gly), 800.43 (Cys34→Dha) and 811.76\_1 (T3 labile adduct)]. In addition to the modifications involving crotonaldehyde (a known RCS from lipid peroxidation) and tiglic aldehyde, two additional adducts in Cluster 2—826.10 and 830.44, corresponding to elemental compositions of  $C_2H_4O$  (within 0.2 ppm) and  $C_2H_5O$  (within 1.4 ppm), respectively—appear to be modifications of RCS that were enhanced with benzene exposure (Figure 1). It is interesting that the Cys34 disulfide of mercaptoacetic acid (841.75) was also enhanced with benzene exposure (Figure 1) and was highly correlated with the Cys34 sulfoxidation products (816.42, 822.42 and 827.76) as well as the labile T3 adduct (811.76\_1). Clusters 1 and 2 are linked by BO-Alb and the mono-oxidation product of Cys34 (816.42), suggesting a connection between benzene metabolism via Cyp2E1 (34) and the subsequent cascade of oxidation reactions (35). The third cluster is comprised of nodes representing Cys34 disulfides of small thiols, notably Cys (851.43), hCys (856.10\_2), glutathione (GSH, 913.45), GluCys (894.44) and CysGly (870.44). The mono- and tri-oxidation products of Cys34 (816.42 and 827.76) appear to be the primary nodes connecting Clusters 1 and 3, consistent with production of Cys34 disulfides via oxidation by ROS (18).

## Discussion

We initially validated our adductomics pipeline with archived plasma from healthy smoking and nonsmoking subjects (11). This led to discovery of 43 adducts, several of which were significantly associated with smoking. Some associations had been anticipated; for example, adducts of two prominent constituents of cigarette smoke—ethylene oxide and acrylonitrile—were only detected in smokers. However, contrary to our expectations, smoking produced lower levels of the three Cys34 oxidation products (816.42, 822.42 and 827.76), possibly due to smoking-induced hypoxia. This combination of results points to separate windows that Cys34 modifications provide for viewing exposure-specific electrophiles and characteristics of the redox proteome (16–20). Since benzene is a carcinogen that produces several electrophilic metabolites (5,27) as well as free radicals, ROS and RCS (7), we used Cys34 adductomics to gain insight into the interplay across all sources of these reactive intermediates in humans. By using archived HSA from a well-curated investigation of workers with and without occupational exposure to benzene that had been documented with multiple air samples from each subject (21), we were able to maximize contrast with respect to benzene exposure.

Only two putative benzene-specific adducts were observed, representing theoretical reactions between Cys34 and BO (837.10) and BDE (854.44), and these were only detected in subsets of the 10 benzene-exposed workers ( $n = 5$  and 8 subjects, respectively). Although the identities of 837.10 and 854.44 were not confirmed—because analytical standards could not be synthesized—their annotations are supported by accurate masses within 1.2 ppm of theoretical values (Table 1), previous targeted experiments (5,36), and large pairwise correlations with personal benzene exposure (aBen) and the Cys34 adducts of BO and the BQs (i.e. BO-Alb and 1,2-BQ-Alb and 1,4-BQ-Alb, see Figure 2) that had been targeted in the previous investigation (27).

Detection of the putative BDE modification (854.44) in 8 of the 10 benzene-exposed subjects was a surprise because a BDE adduct has never been reported in humans or animals exposed to benzene. Despite their potent mutagenicity and carcinogenicity in rodents (22,23), BDEs have received the least attention of benzene's metabolites, and we encourage investigators to target these adducts as potentially important biomarkers in workers exposed to benzene.

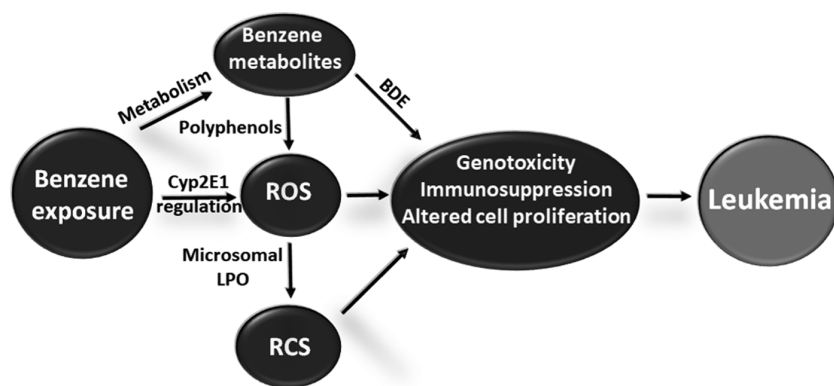
Because BO-Alb had been measured previously by GC-MS in the same HSA samples (27), we had anticipated that the S-phenylation product (837.10) would be found, and we detected this adduct in 5 of the 10 benzene-exposed subjects. The relatively low detection rate for 837.10 (i.e. 50%) could reflect the small amount of HSA analyzed here compared with our earlier study (0.1 versus 5 mg). For the five exposed subjects with both types of measurements, the mean levels of the S-phenylation adduct were 0.787 pmol/mg HSA in the current (untargeted) study versus 1.74 pmol/mg HSA in the original (targeted) study. This 2-fold difference is quite modest considering that the selected ion monitoring method used for GC-MS was specific to target ions of the S-phenylation product and its isotopically labeled internal standard, whereas our adductomics method generated data in scan mode and employed PARs for approximate quantitation relative to a housekeeping peptide.

We did not detect adducts of the BQs despite their previous measurement by GC-MS (27). This may reflect instability of these adducts in the serum-derived HSA we used for this study. Previous analyses of samples from the full study showed that levels of 1,2-BQ-Alb and 1,4-BQ-Alb were 10- and 3-fold lower, respectively, in serum than in plasma that had been stabilized with EDTA (27). (EDTA chelates iron, which is a source of Fenton chemistry that can modify proteins.) Such instability could have been exacerbated by prolonged storage (15 years) of the serum-derived HSA prior to analysis in the current study.

All three sulfoxidation products, corresponding to addition of one to three oxygen (816.42, 822.42 and 827.76), were present at significantly higher levels in benzene-exposed workers than controls, consistent with benzene-induced generation of ROS (Figure 1). Since benzene is a component of cigarette smoke, these results might appear to be contradict our earlier finding that Cys34 oxidation products were less abundant in smokers than nonsmokers (11). However, the amount of benzene absorbed from smoking was orders of magnitude less than that from occupational exposure to benzene. Nonsmoking and smoking workers from the full study had median levels of urinary benzene that were 1.36 and 2.47 nM, respectively, in controls compared with 267 and 197 nM, respectively, in exposed workers (29). Thus, the small dose of benzene derived from smoking was probably not sufficient to overwhelm the hypothesized effect of smoking-induced hypoxia, which could have diminished production of Cys34 oxidation products in smokers (11).

Because all three sulfoxidation products (816.42, 822.42 and 827.76) were highly associated with benzene exposure and the Cys34 sulfenic acid (–SOH) is an intermediate in the formation of mixed disulfides, it is curious that only disulfides of hCys (841.75 and 850.10) and mercaptoacetic acid (841.75) were significantly associated with benzene exposure. Indeed, the two most prominent Cys34 disulfides (Cys and CysGly) were not similarly associated. This apparent contradiction probably reflects the complexity of disulfide formation, which is a reversible process that depends on the nature and redox state of the environment, the disulfide stability and the kinetics of the forward and reverse reactions that can vary across circulating thiols (37–39). In contrast, the alternate oxidation pathways leading to formation of the three sulfoxidation products are essentially irreversible (18,33).

The association of the disulfide of mercaptoacetic acid with benzene exposure is also curious because mercaptoacetic acid (also called thioglycolic acid) is used as a corrosion inhibitor and anti-scaling agent in industry and as a depilatory and hair-treatment product in the general population (<https://pubchem.ncbi.nlm.nih.gov/compound/1133#section=Top>). However, we have no information regarding possible exposures to mercaptoacetic



**Figure 3.** Proposed scheme relating exposure and metabolism of benzene to leukemia via key characteristics of carcinogenicity—genotoxicity, immunosuppression and altered cell proliferation—that have been amply documented for benzene (44). LPO, lipid peroxidation.

acid across the workers in our study. There are probably endogenous sources of mercaptoacetic acid as well because this chemical has been measured in urine samples from the general public (40,41), and we had detected the same adduct (841.75) in two previous studies (11,30). Thus, we suspect that mercaptoacetic acid is ubiquitous in human blood and that benzene selectively enhanced production of the Cys34 disulfide of this circulating thiol in our subjects.

As noted above, Cys34 adductomics provides nuanced views of metabolism and redox biology. Because HSA is the most abundant protein in the extracellular space throughout the body (15), Cys34 is able to scavenge small reactive electrophiles from metabolism of exogenous exposures, such as benzene, and produce irreversible modifications [e.g. putative adducts of BO (837.10) and BDE (854.44)]. But Cys34 is also a prominent constituent of the thiol-redox proteome that undergoes reversible and irreversible reactions with ROS, leading to sulfoxidations and mixed disulfides, as previously discussed. Microsomal oxidative stress motivates production of RCS that can irreversibly modify Cys34 [i.e. adducts of crotonaldehyde (835.11) and tiglic aldehyde (839.78) and putative adducts of  $C_2H_4O$  (826.10) and  $C_3H_5O$  (830.44)]. All three of these classes of Cys34 modifications are represented in Table 1 and adducts from each category were significantly more abundant in benzene-exposed workers than controls (Figure 1).

Our results suggest that benzene metabolism can work via multiple mechanisms to modify critical macromolecules and alter redox biology in ways that may lead to leukemia. Given the potent mutagenicity and carcinogenicity of BDEs (22,23), evidence that at least one BDE isomer reacted with HSA in benzene-exposed workers is provocative. Likewise, evidence that human benzene exposure generates ROS—with concomitant production of RCS via oxidation of microsomal lipids—poses an interesting question. That is, does production of ROS and RCS primarily involve futile cycling of diol and quinone forms of polyphenolic molecules that are produced by Cyp2E1 metabolism of benzene [as generally assumed (7)] or the intrinsic oxidative stress of Cyp2E1 *per se* that can be regulated by benzene in metabolizing cells (6,35,42,43)? In either case, we postulate in Figure 3, that the combination of a potent genotoxicant (BDE) plus chronic oxidative and carbonyl stress contributes to key characteristics of benzene's carcinogenicity (i.e. genotoxicity, immunosuppression and altered cell proliferation) (44) that ultimately give rise to human leukemia.

This small discovery project has limitations. First, the number of subjects included only 10 workers with relatively high exposures to benzene and 10 workers who did not work with

benzene. Second, the processing of serum samples (27) was performed 15 years ago using methods that were substantially more involved than our current adductomics protocol (11). This could have led to formation of artifacts during sample processing or prolonged storage of purified HSA. Third, our adductomics pipeline is untargeted and employs relative quantitation of adduct levels. Finally, our annotations of adducts are incomplete because precursor electrophiles are unknown and can modify Cys34 *in vivo* with varied chemistries—including rearrangements and truncations—that are difficult to predict and reproduce for syntheses of reference standards. Follow-up analyses are planned for archived specimens of pristine plasma (containing EDTA) from the parent study with sufficient numbers of subjects to investigate exposure-adduct relationships.

## Supplementary material

Supplementary Table 1 and Figures 1 and 2 can be found at *Carcinogenesis* online.

## Acknowledgements

This work was supported by the National Institutes of Health through grant R33CA191159 from the National Cancer Institute and grant P42ES004705 from the National Institute for Environmental Health Sciences and by the Intramural Research Program of the National Cancer Institute.

*Conflict of Interest Statement:* S.M.R. has received consulting and expert testimony fees from law firms representing plaintiffs' cases involving exposure to benzene, and he has received research support from the American Petroleum Institute and the American Chemistry Council. M.T.S. has received consulting and expert testimony fees from law firms representing both plaintiffs and defendants in cases involving exposure to benzene. G.L. has received funds from the American Petroleum Institute for consulting on benzene-related health research. The other authors declare they have no competing financial interests.

## References

- IARC (1987) Overall evaluations of carcinogenicity: an updating of IARC Monographs volumes 1 to 42. IARC Monogr Eval Carcinog Risks Hum Suppl., 7, 1–440.
- McHale, C.M. et al. (2012) Current understanding of the mechanism of benzene-induced leukemia in humans: implications for risk assessment. *Carcinogenesis*, 33, 240–252.
- Smith, M.T. et al. (2011) Benzene, the exposome and future investigations of leukemia etiology. *Chem. Biol. Interact.*, 192, 155–159.

4. Rappaport, S.M. et al. (2002) Albumin adducts of benzene oxide and 1,4-benzoquinone as measures of human benzene metabolism. *Cancer Res.*, 62, 1330–1337.
5. Waidyanatha, S. et al. (2005) Characterization and quantification of cysteinyl adducts of benzene diolepoxide. *Chem. Res. Toxicol.*, 18, 1178–1185.
6. Johansson, I. et al. (1988) Benzene metabolism by ethanol-, acetone-, and benzene-inducible cytochrome P-450 (IIE1) in rat and rabbit liver microsomes. *Cancer Res.*, 48, 5387–5390.
7. Subrahmanyam, V.V. et al. (1991) Potential role of free radicals in benzene-induced myelotoxicity and leukemia. *Free Radic. Biol. Med.*, 11, 495–515.
8. Aldini, G. et al. (2007) Intervention strategies to inhibit protein carbonylation by lipoxidation-derived reactive carbonyls. *Med. Res. Rev.*, 27, 817–868.
9. Rubino, F.M. et al. (2009) Toward an “omic” physiopathology of reactive chemicals: thirty years of mass spectrometric study of the protein adducts with endogenous and xenobiotic compounds. *Mass Spectrom. Rev.*, 28, 725–784.
10. Carlsson, H. et al. (2014) LC-MS/MS screening strategy for unknown adducts to N-terminal valine in hemoglobin applied to smokers and nonsmokers. *Chem. Res. Toxicol.*, 27, 2062–2070.
11. Grigoryan, H. et al. (2016) Adductomics pipeline for untargeted analysis of modifications to Cys34 of human serum Albumin. *Anal. Chem.*, 88, 10504–10512.
12. Li, H. et al. (2011) Profiling Cys34 adducts of human serum albumin by fixed-step selected reaction monitoring. *Mol. Cell. Proteomics*, 10, M110.004606.
13. Aldini, G. et al. (2008) Albumin is the main nucleophilic target of human plasma: a protective role against pro-atherogenic electrophilic reactive carbonyl species? *Chem. Res. Toxicol.*, 21, 824–835.
14. Sabbioni, G. et al. (2017) Biomonitoring human albumin adducts: the past, the present, and the future. *Chem. Res. Toxicol.*, 30, 332–366.
15. Merlot, A.M. et al. (2014) Unraveling the mysteries of serum albumin—more than just a serum protein. *Front. Physiol.*, 5, 299.
16. Go, Y.M. et al. (2014) Redox biology: interface of the exposome with the proteome, epigenome and genome. *Redox Biol.*, 2, 358–360.
17. Go, Y.M. et al. (2013) The redox proteome. *J. Biol. Chem.*, 288, 26512–26520.
18. Carballal, S. et al. (2007) Sulfenic acid in human serum albumin. *Amino Acids*, 32, 543–551.
19. Yang, J. et al. (2016) The expanding landscape of the thiol redox proteome. *Mol. Cell. Proteomics*, 15, 1–11.
20. Go, Y.M. et al. (2017) Redox theory of aging: implications for health and disease. *Clin. Sci. (Lond.)*, 131, 1669–1688.
21. Lan, Q. et al. (2004) Hematototoxicity in workers exposed to low levels of benzene. *Science*, 306, 1774–1776.
22. Glatt, H. et al. (1989) Multiple activation pathways of benzene leading to products with varying genotoxic characteristics. *Environ. Health Perspect.*, 82, 81–89.
23. Busby, W.F. Jr et al. (1990) Lung tumorigenicity of benzene oxide, benzene dihydrodiols and benzene diolepoxides in the BLU:ha newborn mouse assay. *Carcinogenesis*, 11, 1473–1478.
24. Reuter, S. et al. (2010) Oxidative stress, inflammation, and cancer: how are they linked? *Free Radic. Biol. Med.*, 49, 1603–1616.
25. Liou, G.Y. et al. (2010) Reactive oxygen species in cancer. *Free Radic. Res.*, 44, 479–496.
26. Cejas, P. et al. (2004) Implications of oxidative stress and cell membrane lipid peroxidation in human cancer (Spain). *Cancer Causes Control*, 15, 707–719.
27. Lin, Y.S. et al. (2007) Albumin adducts of electrophilic benzene metabolites in benzene-exposed and control workers. *Environ. Health Perspect.*, 115, 28–34.
28. Vermeulen, R. et al. (2004) Detailed exposure assessment for a molecular epidemiology study of benzene in two shoe factories in China. *Ann. Occup. Hyg.*, 48, 105–116.
29. Kim, S. et al. (2006) Using urinary biomarkers to elucidate dose-related patterns of human benzene metabolism. *Carcinogenesis*, 27, 772–781.
30. Lu, S.S. et al. (2017) Profiling the serum Albumin Cys34 adductome of solid fuel users in xuanwei and fuyuan, China. *Environ. Sci. Technol.*, 51, 46–57.
31. Shannon, P. et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, 13, 2498–2504.
32. Yeowell-O’Connell, K. et al. (1998) Hemoglobin and albumin adducts of benzene oxide among workers exposed to high levels of benzene. *Carcinogenesis*, 19, 1565–1571.
33. Grigoryan, H. et al. (2012) Cys34 adducts of reactive oxygen species in human serum albumin. *Chem. Res. Toxicol.*, 25, 1633–1642.
34. Guengerich, F.P. et al. (1991) Role of human cytochrome P-450 IIE1 in the oxidation of many low molecular weight cancer suspects. *Chem. Res. Toxicol.*, 4, 168–179.
35. Caro, A.A. et al. (2004) Oxidative stress, toxicology, and pharmacology of CYP2E1. *Annu. Rev. Pharmacol. Toxicol.*, 44, 27–42.
36. Rappaport, S.M. et al. (2002) Non-linear production of benzene oxide-albumin adducts with human exposure to benzene. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.*, 778, 367–374.
37. Narayan, M. (2012) Disulfide bonds: protein folding and subcellular protein trafficking. *FEBS J.*, 279, 2272–2282.
38. Chandrasekhar, S. et al. (2014) Thiol-disulfide exchange in peptides derived from human growth hormone. *J. Pharm. Sci.*, 103, 1032–1042.
39. Glowacki, R. et al. (2004) Cross-talk between Cys34 and lysine residues in human serum albumin revealed by N-homocysteinylation. *J. Biol. Chem.*, 279, 10864–10871.
40. Kuśmierk, K. et al. (2008) Measurement of reduced and total mercaptamine in urine using liquid chromatography with ultraviolet detection. *Biomed. Chromatogr.*, 22, 441–445.
41. Wroński, M. (1996) Separation of urinary thiols as tributyltinmercaptides and determination using capillary isotachopheresis. *J. Chromatogr. B. Biomed. Appl.*, 676, 29–34.
42. Novak, R.F. et al. (2000) The alcohol-inducible form of cytochrome P450 (CYP 2E1): role in toxicology and regulation of expression. *Arch. Pharm. Res.*, 23, 267–282.
43. Song, B.J. (1996) Ethanol-inducible cytochrome P450 (CYP2E1): biochemistry, molecular biology and clinical relevance: 1996 update. *Alcohol. Clin. Exp. Res.*, 20(suppl. 8), 138A–146A.
44. Smith, M.T. et al. (2016) Key characteristics of Carcinogens as a basis for organizing data on mechanisms of carcinogenesis. *Environ. Health Perspect.*, 124, 713–721.