Chemical Research in Toxicology

Refinement of a Methodology for Untargeted Detection of Serum Albumin Adducts in Human Populations

George W. Preston,[†] Michelle Plusquin,[‡,§] Osman Sozeri,[†] Karin van Veldhoven,[‡] Lilian Bastian,[†] Tim S. Nawrot,[§,∥] Marc Chadeau-Hyam,[‡] and David H. Phillips*,[†]

[†]MRC-PHE Centre for Environment and Health, Department of Analytical, Environmental, and Forensic Sciences, Faculty of Life Sciences and Medicine, King's College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH, United Kingdom

[‡]MRC-PHE Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London W2 1PG, United Kingdom

[§]Centre for Environmental Sciences, Hasselt University, 3590 Hasselt, Belgium

[∥]Environment and Health Unit, Leuven University, 3000 Leuven, Belgium

S *Supporting Information*

ABSTRACT: Covalently modified blood proteins (e.g., serum albumin adducts) are increasingly being viewed as potential biomarkers via which the environmental causes of human diseases may be understood. The notion that some (perhaps many) modifications have yet to be discovered has led to the development of untargeted *adductomics* methods, which attempt to capture entire populations of adducts. One such method is fixed-step selected reaction monitoring (FS-SRM), which analyses distributions of serum albumin adducts via shifts in the mass of a tryptic peptide [Li et al. (2011) *Mol. Cell. Proteomics* 10, M110.004606]. Working on the basis that FS-SRM might be able to detect biological variation due to environmental factors, we aimed to scale the methodology for use in an epidemiological setting. Development of sample preparation methods led to a batch workflow with increased throughput and provision for quality control. Challenges posed by technical and biological variation were addressed in the processing and interpretation of the data. A pilot study of 20 smokers and 20 never-smokers provided evidence of an effect of smoking on levels of putative serum albumin adducts. Differences between smokers and never-smokers were most apparent in putative adducts with net gains in mass between 105 and 114 Da (relative to unmodified albumin). The findings suggest that our implementation of FS-SRM could be useful for studying other environmental factors with relevance to human health.
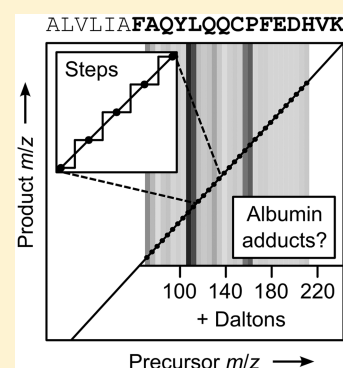
INTRODUCTION

Non-enzymatic covalent modifications to macromolecules represent a source of potential biomarkers with which to study human health and disease.[1−4] In some cases, modifications might be causally related to biological end points (e.g., a mutagenic lesion in DNA causing cancer).[5] Alternatively, modifications might be found in "off-pathway" products from which "on-pathway" events may be inferred. This latter possibility has motivated researchers to investigate adducts of blood proteins as potential biomarkers, particularly in studies seeking to understand the effects of environmental factors on human populations (biomonitoring, exposome studies). The rationale for this approach is that exogenous chemicals, or their derivatives, should tend to react with nucleophilic amino acid residues in the proteins.[6,7] Some exogenous chemicals possess intrinsic reactivity toward proteins (e.g., sulfur mustard),[8] while others require metabolic activation (e.g., aflatoxins).[2] Less direct mechanisms, in which the origin of the reactive chemical is endogenous rather than exogenous, may also occur.[9]

Hemoglobin and human serum albumin (HSA) are two blood proteins that have been used extensively for biomonitor-ing.[10−12] HSA, for example, is ideally suited for this purpose due to its reactivity (at Cys-34 and other nucleophilic loci),[13] abundance (approximately 40 g L^{-1} in serum),[14] and long physiological half-life (approximately 3 weeks).[15] Some adducts of blood proteins have been known for decades, and targeted methods for their detection have been established.[2] It is only recently, however, that "-ome" concepts have been applied to DNA and protein adducts and that corresponding "-omics" methods have been applied to their detection.[1,6,11,16] The idea that known adducts might exist within a wider *adductome* has led to the development of untargeted *adductomics* methods for the discovery of novel biomarkers. In 2011, Li et al. reported the use of fixed-step selected reaction monitoring (FS-SRM), a triple quadrupole mass spectrometry (TQ-MS) method for HSA adductomics.[17] The authors detected modifications within a short sequence of amino acid residues in HSA (residues 27−41) via shifts in a sequence tag[18] of the third-heaviest tryptic peptide ("T3"). Such a shift may be characterized by the mass of a putative "R" group[17] (presumably attached to the sulfur

atom of Cys-34) or, as below, by the net gain in molecular mass due to the modification (d).

Detection by TQ-MS involves the use of quadrupole mass analysers to exclude all but the ions of interest. Under appropriate conditions, a precursor ion gains a stable trajectory in the first quadrupole (Q1), and its product does likewise in the third quadrupole (Q3). If, for T3 peptides of HSA adducts, the product ions are always related to their precursors by a constant loss of mass (i.e., a loss that is independent of d), then the loss may be used as a basis on which to detect unknown adducts (Figure 1A). Thus, for untargeted adductomics, the
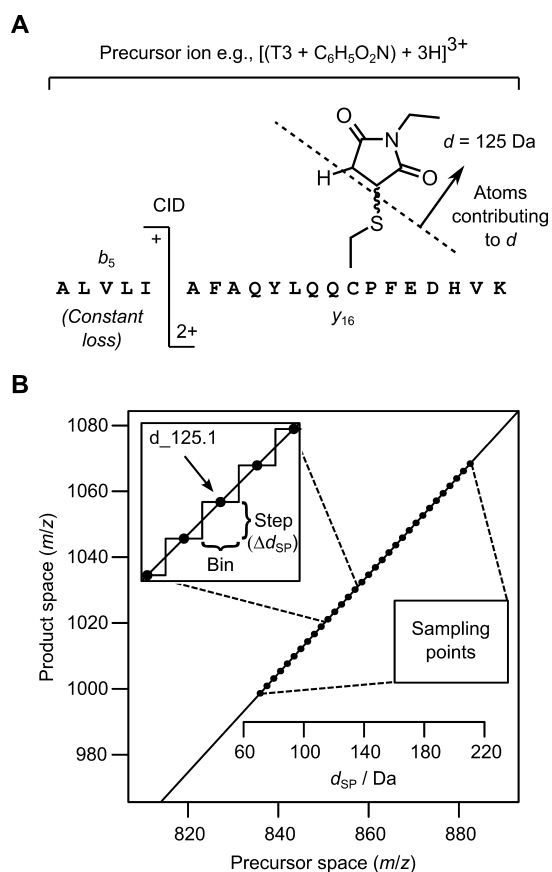
precision. These benefits are gained at the expense of resolution, meaning that the usefulness of FS-SRM comes not from an ability to identify particular adducts, but rather from the characterization of their distribution as a whole. If the closeness of the stepwise measurements is balanced with the range of masses transmitted by Q1, then the method should detect any and all relevant HSA adducts. Li et al.[17] refer to the range of masses captured by a single measurement as a *bin* (Figure 1B). We refer to the values of d at which measurements are made as *sampling points* (SPs), and we use the variable d_{SP} to distinguish them from the d-values of adducts (d_{adduct}). Each SP can be described in terms of the m/z values of a precursor ion and one or more product ions (Figure 1B).

Li et al.[17] applied their methods (sample preparation and FS-SRM; Figure 2) to analyses of archived plasma protein that had

Figure 1. Theoretical basis of FS-SRM (e.g., detection of an *N*-ethylmaleimide adduct at Cys-34 of HSA): (A) Upon collision-induced dissociation (CID) of a triply protonated T3 peptide, a constant fragment (singly charged b_5) is lost, and a variable fragment (doubly charged y_{16}) is detected. A formal disconnection (dashed line) distinguishes atoms belonging to the T3 peptide from those contributing to the net gain in mass (d). (B) Sampling points (SPs) are values of d at which ion intensities are measured. Each SP specifies a pair of m/z values that permit stable ion trajectories in Q1 and Q3, respectively. The evenly spaced points on the plot relate to stable trajectories of $[M + 3H]^{3+}$ ions in Q1 and y_{16} ions in Q3. If "step" and "bin" are of equivalent size, each adduct should be detected once and only once. "d_125.1" is the SP at which we would expect to detect the *N*-ethylmaleimide adduct (d_{adduct} = 125 Da).
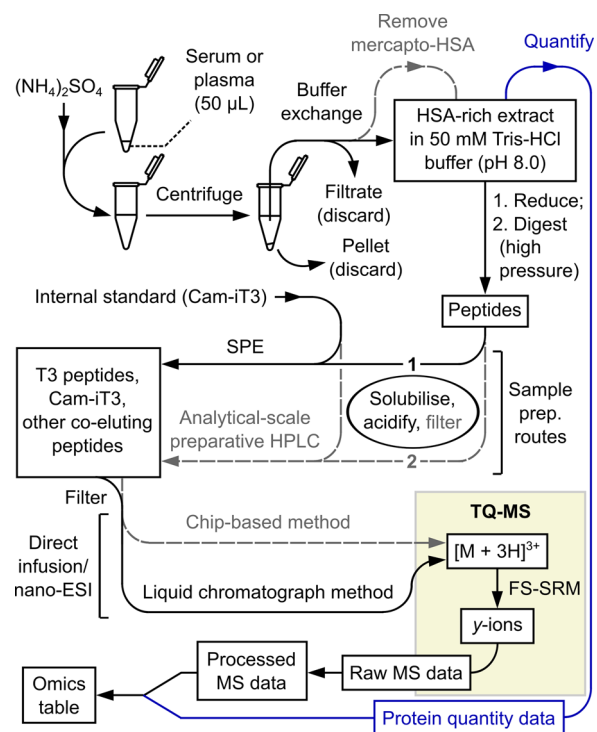

Figure 2. A workflow for HSA adductomics comprising sample preparation, TQ-MS, and data processing. Processes are represented by arrows (dashed gray line = initial method; solid black line = final method). HSA is extracted from serum or plasma and digested with trypsin. Relevant digestion products are purified collectively (one of two possible cleanup methods) and analyzed using TQ-MS (one of two possible sample delivery methods). The raw data are processed and tabulated in preparation for statistical analysis. Protein concentration data from a separate assay, carried out in parallel, can feed into the final stages of data processing.

conditions in Q3 are offset relative to those in Q1 according to the constant loss, and the respective conditions are covaried as a function of d (Figure 1B). In FS-SRM, this covarying of conditions is done in a stepwise fashion (pseudo constant neutral loss scanning).[1,16] By devoting more time to fewer measurements (cf. conventional scanning), stepped methods should benefit from enhanced sensitivity, accuracy and

been pooled according to subjects' ethnicities and tobacco smoking habits. Differences observed between pools suggested that FS-SRM might be able to detect statistically significant differences between groups of individual samples that had not been pooled. On this basis, we wished to use FS-SRM in an epidemiological setting,[19] which would necessitate the analysis of tens to hundreds of samples per group. In implementing FS-SRM for this purpose, the existing methods had to be modified in order to achieve an acceptable throughput of samples. Novel adaptations (e.g., use of solid-phase extraction for sample cleanup) enabled faster sample preparation. Focusing on a

narrower range of SPs enabled the throughput of TQ-MS to be increased. Further adaptations (e.g., quality control) were introduced to address challenges associated with technical variation, which are a particular concern when the number of samples is large. Methods were evaluated first using synthetic and semisynthetic standards (adducts of maleimides) and then by testing for effects of a model environmental factor (tobacco smoking) on the levels of putative adducts detected in human plasma (20 smokers and 20 never-smokers from the ENVIRONAGE cohort).[20] To our knowledge, this is the first example of FS-SRM being used to investigate associations between environmental factors and levels of putative HSA adducts in a quantitative fashion.

EXPERIMENTAL PROCEDURES

Chemicals. Synthetic peptides (T3, ALVLIAFAQYLQQCP-FEDHVK; and isotopically labeled T3, AL[valine-$^{13}C_5$,^{15}N]-LIAFAQYLQQCPFEDH[valine-$^{13}C_5$, ^{15}N]K) were purchased from Insight Biotechnology (Wembley, UK). N-(Naphthalen-1-yl)-maleimide was purchased from Santa Cruz Biotechnology (Heidelberg, Germany). Bovine serum albumin, 5,5′-dithiobis(2-nitrobenzoic acid) (Ellman's reagent), N-ethylmaleimide (NEM), GSH, human serum from male AB plasma, HSA, lyophilized citrated bovine plasma, lyophilized citrated human plasma, propranolol hydrochloride, and porcine trypsin (unmodified) were purchased from Sigma-Aldrich (Dorset, UK). Solvents (HPLC-grade and LC-MS-grade) and other general laboratory chemicals were purchased from Alfa Aesar (Lancashire, UK), Fisher Scientific (Loughborough, UK), Oxoid (Hampshire, UK), Sigma-Aldrich, or VWR International (Leicestershire, UK). All commercial substances were used without further preparation except for LC-MS-grade solvents (degassed by water-bath sonication) and lyophilized plasma (reconstituted with water). Water for general purposes (18.2 MΩ cm) was obtained using an ELGA purification apparatus (Veolia Water Technologies, High Wycombe, UK).

Preparation of Standards. An isotopically labeled internal standard (S-carbamidomethylated isotopically labeled T3, "Cam-iT3") was prepared as described previously.[21] Similar methods were used to prepare a N-(naphthalen-1-yl)succinimid-3-yl derivative of T3 (Nns-T3) as a mixture of putative isomers. Details can be found in the Supporting Information (SI). Both synthetic peptide adducts were obtained as eluates from reversed-phase HPLC. Cam-iT3 was quantified using analytical HPLC with fluorescence detection as described in the earlier report.[21] NEM-modified HSA (Nes-HSA; expected modification: S-[N-ethylsuccinimid-3-yl]), for use in the preparation of Nes+ plasma (i.e., "control plasma" containing Nes-HSA; see below), was prepared by reacting partially reduced commercial HSA with NEM in PBS (see SI). Nes+ plasma was prepared by adding a solution of Nes-HSA (1.4 mg) in PBS (50 μL) to commercial human plasma (5 mL). Nes− plasma (i.e., control plasma lacking Nes-HSA) was prepared in the same way, except that unmodified commercial HSA was added in place of Nes-HSA. Plasma preparations were stored as 50 μL aliquots at −80 °C.

Summary of Sample Preparation Methods. HSA was prepared from plasma using a method similar to that of Li et al.,[17] but with the difference that adducts were not enriched (see SI for full details). Briefly, most of the unwanted plasma protein was salted out with ammonium sulfate, and low-molecular-weight solutes were exchanged with Tris-HCl buffer (50 mM, pH 8.0). The resulting HSA-rich extract had a final volume of 550 μL (fresh plasma) or 450 μL (control plasma; smaller volume due to lower protein content) and contained protein recovered from the equivalent of 18 μL of plasma. An aliquot of each extract, containing protein from the equivalent of 2.68 μL of fresh plasma or 3.28 μL of control plasma, was subjected to further processing. The conditions used for protein denaturation (TCEP hydrochloride, methanol, heat) and digestion (trypsin, pressure cycling, heat) were similar to those described by Li et al., but samples were processed in batches rather than individually. The digestion

mixture was scaled down to fit into a smaller tube (PCT MicroTube; Pressure Biosciences, MA, USA). Up to 12 tubes could be accommodated in the pressure cycling instrument (Barocycler NEP2320; Pressure Biosciences/Constant Systems, Northants, UK). Digestion products were prepared for FS-SRM in one of two ways: by analytical-scale preparative HPLC (ASP-HPLC; see SI) or by solid-phase extraction (SPE; see below). This "decoupled" approach (cleanup followed by offline TQ-MS) is preferred over conventional hyphenated analysis because FS-SRM takes longer than the width of a typical chromatographic peak.[17] In the present study, ASP-HPLC was used initially during method development, with SPE being introduced later to increase the throughput of sample cleanup. For comparison, both methods were used to fractionate a "mock digest" containing Cam-iT3, Nns-T3, and the usual buffers and additives. Relevant eluates were analyzed using electrospray ionization ion trap mass spectrometry (see SI).

Solid-Phase Extraction. Digestion products were separated using reversed-phase SPE on a polymeric sorbent (Strata-X, 30 mg/1 mL tube; Phenomenex, Macclesfield, UK). SPE tubes were attached to a 10-port vacuum manifold (Biotage, Uppsala, Sweden), and conditioning/loading/washing/elution were achieved by drawing solutions through the sorbent under house vacuum. The sorbent was conditioned with a solution of 0.1% (v/v) formic acid in acetonitrile (1 mL), then equilibrated with a solution of 0.1% (v/v) formic acid in 17:3 (v:v) water:acetonitrile (2 × 1 mL). The reservoir of the SPE tube was charged with a 0.1% (v/v) solution of formic acid in 37:1 (v:v) water:acetonitrile (600 μL), and digestion products were added from the PCT MicroTube (see above). The MicroTube was washed with a 1% (v/v) solution of formic acid in 1:1 (v:v) water:acetonitrile (145 μL), and the washings were added to the SPE tube. Cam-iT3 (90 pmol) was added, and the tube was sealed with a poly-(tetrafluoroethylene) cap (Supelco; Sigma-Aldrich, Dorset, UK). The contents of the reservoir (pH ∼ 3) were mixed by inverting the tube once. The cap was removed, and the digestion products and Cam-iT3 were loaded onto the sorbent. The sorbent was then washed with a 0.1% (v/v) solution of formic acid in 4:1 (v:v) water:acetonitrile (5 × 1 mL). Peptides of interest were eluted with a 0.1% (v/v) solution of formic acid in 1:3 (v:v) water:acetonitrile (2 × 500 μL) into a 2 mL polypropylene microcentrifuge tube (Eppendorf, Stevenage, UK). The eluate was vortexed-mixed to homogeneity and centrifuged (16,000 × g, 2 min). Supernatant (300 μL) was transferred to a centrifugal filter unit (Costar Spin-X, nylon, pore size = 0.22 μm; Corning, New York, USA) that had been prerinsed with eluent, and the unit was centrifuged (10,000 × g, 1 min). Filtrate (250 μL) was transferred to a glass autosampler vial with a 300 μL fused insert (Fisher Scientific, Loughborough, UK).

FS-SRM. FS-SRM was performed using a Thermo TSQ Quantum Access triple quadrupole mass spectrometer (Thermo Fisher Scientific, Hemel Hempstead, UK) operated in SRM mode. Three major modifications were made to the method of Li et al. (see SI for the full method). First, the chip-based system used for sample delivery was replaced with a Dionex Ultimate 3000 liquid chromatograph and a Nanospray Flex ion source (both from Thermo Fisher Scientific). Second, only 32 of the 77 SPs were used (71.1 Da ≤ d_{SP} ≤ 210.6 Da). The other SPs were omitted as a compromise between mass-range coverage and analytical throughput. The chosen SPs represent a continuous range that is uninterrupted by potential artifacts (e.g., alkali metal ion adducts), and in which some signals of interest had been observed by Li et al.[17] Additionally, truncation of the range allowed measurements to be scheduled in a more consistent fashion (see Table S1). Third, different conditions were used to filter product ions in Q3 (see SI), resulting in fewer repeat measurements per 30-s "segment" of the analysis. As in the earlier method, putative adducts were detected via a sequence tag consisting of a triply charged precursor ion and three doubly charged y-ions (y_{15}, y_{16}, and y_{17}; standard peptide fragment nomenclature, see SI). Starting from 32 evenly spaced values of precursor m/z (836 to 882.5), corresponding values of product m/z were derived using eq 1 (x = 4, 5, 6), and corresponding values of d_{SP} were derived using eq 2. The use of molecular weight (M_r) accounts for the apparent averaging of m/z that occurs when multiply charged

ions are analyzed at low resolution. The step size, Δd_{SP}, was always 4.5 Da (equivalent to 1.5 m/z units of precursor space), except in an evaluation experiment where it was reduced to 0.9 Da (0.3 m/z units of precursor space). Sample delivery was monitored via targeted detection of the Cam-iT3 internal standard (~20 measurements per segment; see Figure S1 and Table S1).

$$\text{product } m/z = \frac{(\text{precursor } m/z) \times 3 - (M_r \text{ of } b_x)}{2} \quad (1)$$

$$d_{SP} = (\text{precursor } m/z) \times 3 - (M_r \text{ of } [T3 + 3H]^{3+}) \quad (2)$$

Pilot Study of Tobacco Smoke Exposure. The study included 40 mothers from the ENVIRonmental influence ON AGEing in early life (ENVIRONAGE) birth cohort in Belgium. Questionnaires provided information on age, education, smoking status, ethnicity, and prepregnancy body mass index (Table 1). All procedures were

Table 1. Characteristics of Participants Included in the Tobacco Smoke Study[a]

Characteristic		Value	
		Never-Smokers	Smokers
N		20	20
Ethnicity			
	European-Caucasian	20 (100)	19 (95)
	Non-European		1 (5)
Age (years)		28.15 (5.12)	25.90 (5.30)
Body mass index (kg m^{-2})		24.29 (4.78)	24.33 (5.88)
Education			
	Low	3 (15)	5 (25)
	Medium	6 (30)	9 (45)
	High	11 (55)	6 (30)
Number of cigarettes per day		0	10 (8.69)

[a]Counts and percentages for categorical variables; means and standard deviations for continuous variables.

approved by the Ethical Committee of Hasselt University and East-Limburg Hospital. The study design and procedures were described in detail previously.[20] Briefly, written informed consent was obtained from each participating mother who gave birth in the East-Limburg Hospital in Genk, Belgium. Twenty mothers who reported smoking during pregnancy and 20 mothers that had never smoked were selected based on their self-reported smoking behavior. For smokers, the median number of cigarettes per day during pregnancy was 7.25 (maximum: 40; minimum: 1; interquartile range: 4.88−10.00; see Table S4), and the median number of pack-years was six (interquartile range: 4.50−9.31). Peripheral blood was collected 1 day after parturition in Vacutainer plastic whole blood tubes with spray-coated K2EDTA (BD, Franklin Lakes, NJ, USA). Within 20 min of blood collection, samples were centrifuged (3200 rpm, 15 min) to isolate plasma. Plasma was collected and stored in microcentrifuge tubes (Eppendorf) at −80 °C. To assess the reliability of self-reported tobacco consumption, plasma cotinine concentrations (Table S4) were measured using an enzyme-linked immunosorbent assay kit (product ref CO096D; Calbiotech, Spring Valley, CA, USA). Smoking status was inferred using a cutoff value of 14 ng mL^{-1} (typical literature cutoff for serum cotinine concentrations).[22]

Each smoker's sample was paired with one from a never-smoker based on the storage time of the plasma. Twenty pairs were distributed randomly among five batches. For 10 randomly selected pairs, the default order of analysis (never-smoker then smoker) was reversed. To assess among-batch variation, an aliquot of Nes$^+$ plasma (positive control) was analyzed at the end of every batch. Analysts were blinded to the identities of the subjects' samples during sample preparation and TQ-MS. The five batches of subjects' samples were processed consecutively. To assess within-batch variation, a sixth batch consisting

of nine aliquots of Nes$^+$ plasma was also included. All 54 analyses were then replicated using different aliquots of the same plasma samples. For the second round of analyses, the only change was in the order of the batches (i.e., the order of samples within batches was the same). The complete schedule of analyses is depicted in Figure S3.

Quantity and Quality of Extracted Protein. The concentration of protein in each plasma extract was determined using the bicinchoninic acid (BCA) assay (Pierce/Thermo Fisher Scientific, Paisley, UK; manufacturer's protocol with minor modifications). Undiluted extracts (15 μL) were analyzed in triplicate on 96-well microplates (Greiner Bio One, Gloucester, UK), and each plate also included standards (commercial HSA in Tris-HCl buffer; five concentrations in duplicate) and quality controls (pooled extracts of Nes$^+$ plasma). The experimental design described above was transferred to the microplate format, so that extracts in the same batch, or from different aliquots of the same plasma, were always analyzed on the same microplate. BCA working reagent (185 μL) was added to each well, and the microplate was incubated at 37 °C for 30 min. After 10 min at ambient temperature, a microplate spectrophotometer (Biotek Instruments, Bedfordshire, UK) was used to measure the absorbances (λ = 595 nm). A calibration curve was established from the standards' concentrations and mean absorbances (linear regression analysis) and was used to estimate the concentration of protein in each extract. The absolute mass of protein in each extract was calculated by multiplying the estimated concentration by the volume. The quality of the protein was inferred from a set of three example extracts (one each of smoker's plasma, never-smoker's plasma and Nes$^+$ plasma), which were analyzed using SDS-PAGE (see SI).

Processing of FS-SRM Data. Initial processing of raw data was done using the methods of Li et al. with minor modifications. MATLAB code (see Acknowledgments) was used to calculate the apparent amounts of adducts in each eluate and to exclude any result coming from a missing or incomplete sequence tag (MATLAB, version 8.2.0.701, The MathWorks, Natick, MA, USA; see SI). The additional validation step described by Li et al. was used for a parallel analysis of aggregated data (see Results and Discussion), but not to exclude any individual data from statistical testing. Partially validated data were processed further in R (version 3.2.4; base package).[23] To account for variation due to different concentrations of HSA in the plasma extracts, adduct amounts were normalized using results from the BCA assay (see SI). Null responses (missing or incomplete sequence tags) were imputed using random estimates of baseline noise (see SI). Stability of the MS conditions was inferred from SRM of the internal standard (first measurement of every cycle; see Table S1). Unstable segments were identified by comparing stability metrics (segment-wise means and coefficients of variation) to threshold values (Table S5). Any data acquired in unstable segments were excluded at this stage (exclusion rate for ENVIRONAGE data: 4.3%). Best estimates of response variables for the ENVIRONAGE samples were derived from pairs of full technical replicates by averaging if possible. If a result was missing or appeared to be a false negative (null response with nonzero counterpart), its counterpart was taken as the best estimate. This was done using an R script based on Table S6, and the resulting table of best estimates was the primary data set used for statistical testing.

Statistical Methods. All statistical tests were carried out using R (version 3.2.4).[23] The significance level was $p = 0.05$ unless stated otherwise. The dependence of responses on concentration variables was assessed using linear regression analysis. A two-tailed independent samples t-test was used to test the hypothesis that the smokers and never-smokers differed in the masses of protein recovered from their plasma. The Pearson correlation coefficient (r) was used to compare masses of protein recovered in technical replicates. The similarity between pairs of adduct distributions from FS-SRM (spectral similarity;[24] one coefficient per pair, 32 SPs per comparison) was quantified as the Spearman rank correlation coefficient (ρ). Spearman's ρ was also used to test for agreement between technical replicates (one coefficient per SP, ≤40 subjects per comparison). The Wilcoxon rank-sum test (one test per SP, 40 subjects per comparison) was used to test for associations between smoking status and levels of

putative HSA adducts. Principal component analysis was used to assess the degree of correlation among the 32 sets of responses. Five principal components explained 99% of the variation in the data set. To account for making multiple comparisons while also taking the degree of correlation into consideration,[25] a corrected significance level of $p = 0.05/5 = 0.01$ was used.

■ RESULTS AND DISCUSSION

Method Development. In the present study, we adapted the FS-SRM workflow described previously by Li et al.[17] (Figure 2). For the steps leading to tryptic digestion of HSA, our only major modification was to omit the adduct enrichment procedure (removal of mercapto-HSA using covalent chromatography). In an earlier study, Funk et al.[26] showed that such a procedure can facilitate high-level adduct enrichment, but noted that better results were obtained when the protein starting material had been isolated from freshly collected plasma. When we attempted enrichment of a model adduct spiked into commercial serum (method similar to the one described by Li et al.;[17] see SI), only a modest proportion of the albumin was removed (mean ± s.d. = 39.5 ± 1.2%; four replicates). Being more similar to Funk and co-workers' results for archived plasma protein, this result pointed to a sample-dependent effect that would have been difficult to control for in the present study. A secondary consideration was the throughput of sample preparation, which was enhanced by omitting the enrichment procedure (see SI).

SDS-PAGE of three example extracts (see SI and Figure S7) indicated that HSA was always the major component and that the ratio of HSA to the major impurity (probably transferrin[14]) did not vary appreciably. From these results, we inferred uniformity of relative HSA concentration (fraction, w/w, of total protein) across all samples. This inference was supported by evidence from the literature.[27] In an extract consisting of only HSA and transferrin, for example, the relative HSA concentration should vary only between 92% and 96% (assumption: no respective enrichment of either protein). We considered whether it was necessary to quantify the extracted protein prior to digestion. Quantification is advantageous because it allows variation in the protein concentration (physiological and/or technical in origin) to be corrected for by dilution. Such corrections allow the amount of an adduct to be calculated as a fraction (w/w) of total protein. An alternative approach would be to dilute all extracts to the same fixed volume and allow the protein concentration to vary freely. Different amounts of protein would be digested, and the results of FS-SRM would relate to absolute concentrations instead of fractions of total protein. In the present work, we explored both possibilities: the volume was fixed, and the protein concentration was measured but not adjusted (see below). Results for standards and quality controls indicated that the BCA assay used to measure protein content was accurate, precise, and stable (Figure S8 and Table S10).

For tryptic digestion, the method of Li et al.[17] was adapted so that samples could be processed in parallel rather than individually. For cleanup of the digestion products, an ASP-HPLC method was implemented initially,[17,28] but was later replaced by SPE (parallel cleanup of up to 10 samples, no possibility of carry-over, and throughput at least twice that of ASP-HPLC). To ensure that the range of adducts purified by SPE was at least as wide as that of ASP-HPLC, a comparison was made using two standards (Cam-iT3 and Nns-T3). Mass spectrometric analyses indicated that both methods could elute

Cam-iT3 and Nns-T3 completely in a single, appropriately sized volume of eluate (Figure S9).

Evaluation of FS-SRM. The capabilities of FS-SRM were explored using Nes[+] controls, which were mixtures of normal human material (plasma or serum) and NEM-treated HSA (Nes-HSA). Analysis of Nes[+] plasma produced a distinctive distribution of responses (Figure 3A). The response at one of
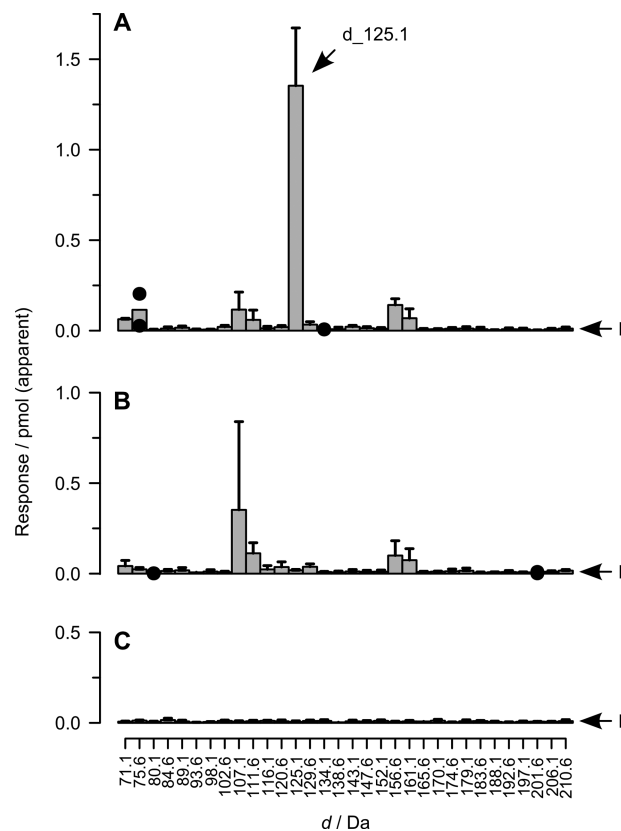

Figure 3. Distributions of putative HSA adducts (mean responses) observed in different plasma preparations: (A) human plasma spiked with Nes-HSA; (B) human plasma without Nes-HSA; (C) bovine plasma. "L" indicates the estimated detection limit. "d_125.1" is the SP at which detection of Nes-HSA was expected. Error bars represent s.d. for at least three technical replicates. Where only two replicates were available, the individual data are plotted.

the SPs (d_125.1) was consistent with the presence of NEM-modified Cys-34 (calculated $d = 125$ Da). The other responses were attributed tentatively to components of normal plasma. Analysis of Nes[−] plasma (no NEM treatment) confirmed that Nes-HSA was responsible for the majority of the response at d_125.1, but not for any of the other prominent responses (Figure 3B). Analysis of bovine plasma confirmed that none of the prominent responses were due to methods, reagents, or some general property of plasma (Figure 3C). Since Nes-HSA is a special case ($d_{adduct} = d_{SP}$), cases where d_{adduct} was offset relative to d_{SP} were also investigated. To do this, a Nes[+] serum extract was analyzed using SPs that were closer together ($\Delta d_{SP} = 0.9$ Da). From the results (Figure 4), it was possible to simulate the effect of varying d_{adduct} on the response at d_{SP}. This confirmed that $\Delta d_{SP} = 4.5$ Da was an appropriate spacing, but it also revealed that the magnitude of the response declined as the offset increased. The results in Figure 4 also highlight how different adducts with similar masses could potentially be detected at the same SP. The responsiveness of FS-SRM was
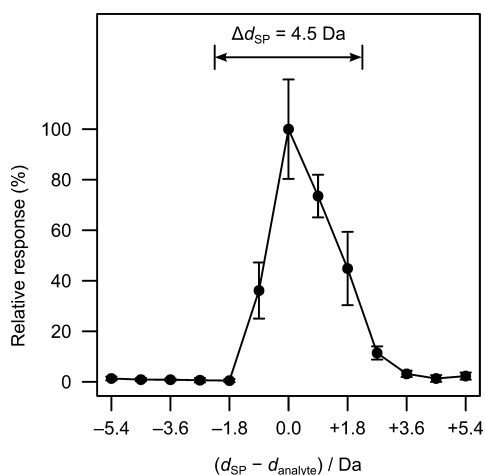
Figure 4. Effect of separation ($d_{SP} - d_{analyte}$) on the responsiveness of FS-SRM. The results indicate that if $d_{analyte}$ is more than half a step ($\Delta d_{SP}/2 = 2.25$ Da) away from d_{SP}, then the analyte will be detected at a different SP. This confirms that $\Delta d_{SP} = 4.5$ Da is an appropriate size of step.

investigated further by varying absolute and relative amounts of Nes-HSA. Dilution series were prepared from an extract of Nes⁺ serum (diluent: buffer or solution of Nes⁻ protein), and FS-SRM was performed as usual. In one experiment, the relative amount of Nes-HSA was fixed at 1.0% (w/w), and the amount of total protein was varied. The dependence of the d_125.1 response on the amount of total protein was positive and linear (Figure 5A), and a similar trend was observed at some of the other SPs (Figure S10). The observed proportionality became important when investigating human population samples in which the amount of protein was subject to biological variation (see below). In a different experiment, the amount of total protein was fixed, and the relative amount of Nes-HSA was varied. The dependence of the d_125.1 response on the relative amount of Nes-HSA showed clear evidence of linearity above ∼0.2% (w/w), but below this, the responsiveness appeared dampened (Figure 5B). Nevertheless, Nes-HSA was frequently still detectable in relative amounts as low as 0.04% (w/w), and, if suspected false negatives were

discounted, the response was still positively associated with the relative amount of Nes-HSA. The trend observed in Figure 5B for d_125.1 was not apparent at any other SPs because the concentration of total HSA (and therefore of other adducts) was constant (Figure S11).

Throughput Analysis. Method development reduced the time required for sample preparation from 85 to 22 min (average time per sample excluding protein quantification). Omitting the adduct enrichment step also eliminated an overnight incubation. Reducing the number of SPs from 77 to 32 increased the throughput of TQ-MS by an estimated 44%.

Pilot Study of Tobacco Smoke Exposure. Results for the ENVIRONAGE subjects revealed substantial variation in the recovery of protein (mean ± s.d. = 611 ± 138 μg). Amounts of protein recovered from different aliquots of the same plasma were, however, well-correlated ($r = 0.90$, $p = 7.53 \times 10^{-15}$; Figure S12), suggesting that much of the observed variation was biological rather than technical. The amount of protein recovered from self-reported smokers (mean ± s.d. = 667 ± 135 μg) was higher than that from never-smokers (mean ± s.d. = 554 ± 120 μg), and the difference was statistically significant ($p = 0.00794$). In other populations, investigators have observed either the opposite relationship[29] or no significant difference.[30] One possible explanation for our observations would be a modulation of the effect of smoking by pregnancy, although the findings of other recent work are generally inconsistent with this idea.[31] To simplify interpretation of the FS-SRM results, it was decided that statistical analyses should focus on variation that could not be formally explained by differences in protein concentration. Thus, an attempt was made to correct for these differences using simple normalization (eq S1), for which the compositional analysis (see above) and the result in Figure 5A provided some justification. Normalized responses were found to be distributed among the SPs in a reproducible fashion (e.g., response at d_107.1 tending to be among the highest, response at d_93.6 tending to be among the lowest). Reproducibility in this respect was quantified by comparing the distribution of median responses for a set of plasma aliquots with that for an identical second set. The results indicated that the shapes of the distributions were reproducible, irrespective of smoking status (spectral similarity
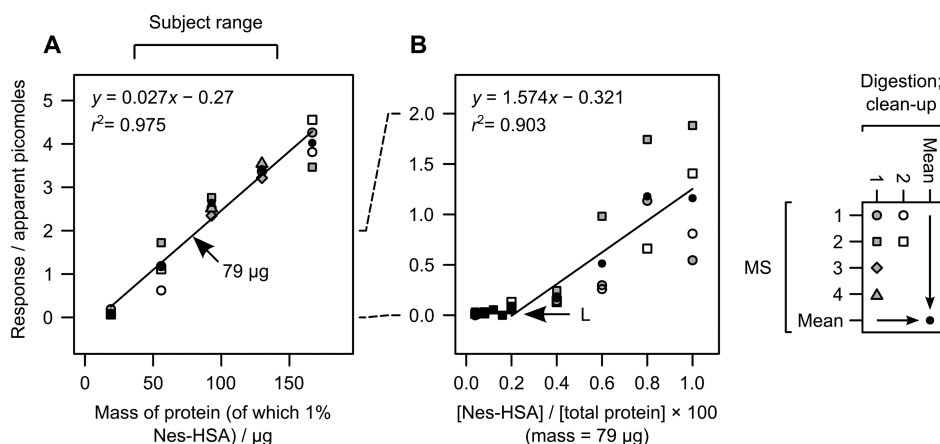

Figure 5. Effect of varying the amount of Nes-HSA on the response at d_125.1: (A) The relative amount of Nes-HSA with respect to total protein was fixed at 1% (w/w), and the amount of total protein was varied; (B) the amount of total protein was fixed at 79 μg, and the amount of Nes-HSA was varied. Linear regression analysis was performed using mean responses as defined by the key (full replicate: digestion, cleanup, and MS; partial replicate: MS only). The range of amounts of total protein derived from ENVIRONAGE subjects' plasma is indicated in part A. "L" indicates the estimated detection limit.

for never-smokers: $\rho = 0.84$, $p = 4.29 \times 10^{-7}$; for smokers: $\rho = 0.85$, $p = 3.19 \times 10^{-7}$). Among the responses measured at a given SP, there was substantial technical variation, and it was likely that many responses lay outside of their respective linear dynamic ranges. False negatives were relatively common (e.g., no response detected at d_111.6 in 16% of controls), so steps were taken to minimize their influence on the results of statistical testing (see Experimental Procedures). The occurrence of false negatives partly explains why responses for different aliquots of the same plasma were often not correlated. Notably, however, a correlation was observed for responses at d_107.1 ($\rho = 0.56$, $p = 2.62 \times 10^{-4}$). Results from controls indicated that TQ-MS was a major source of technical variation (Figure 6 and Table 2). An effect of the order of analysis on the

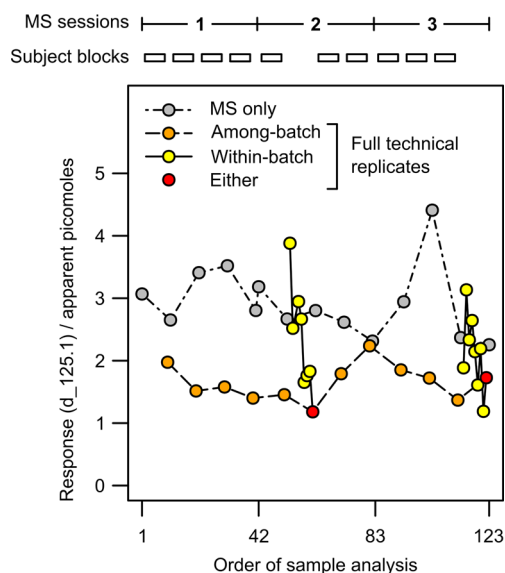**Figure 6.** Sources of technical variation inferred from apparent amounts of Nes-HSA in Nes$^+$ controls. Among-batch variation was inferred from Nes$^+$ controls processed at the ends of "subject blocks" (i.e., batches of ENVIR*O*NAGE samples). Within-batch variation was inferred from batches consisting of solely of Nes$^+$ controls. "Either" indicates overlap between sets of technical replicates. Periodic replacement of the nanoelectrospray emitter ("MS sessions") was considered as another possible source of technical variation ("MS only").

magnitude of the response was also apparent, although randomization should have mitigated any potential bias conferred by this effect.

The similarity between smokers' and never-smokers' distributions (spectral similarity, method as above: $\rho = 0.94$, $p < 2.2 \times 10^{-16}$) suggested that at least some of the same putative adducts were present in both groups. We tested for

associations between responses at SPs and smoking status and found that the response at one SP (d_111.6) was significantly positively associated with smoking ($p \leq 0.000836$; Figure 7 and
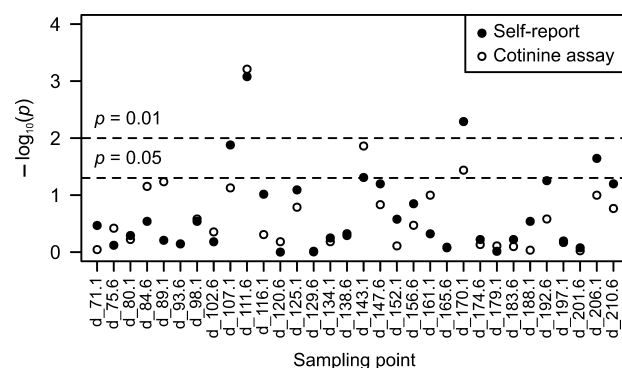

Figure 7. Distributions of p-values obtained from Wilcoxon rank-sum tests (testing for associations between smoking status and responses at SPs). The dashed lines indicate corrected and uncorrected significance levels ($p = 0.01$ and $p = 0.05$, respectively). The p-values showed a limited dependency on the method used to determine smoking status (self-report or cotinine assay).

Table S11). Significance was observed irrespective of the method by which smoking status was determined, but only when technical replicates were aggregated to exclude suspected false negatives (see Experimental Procedures). The response at another SP (d_170.1) was significantly negatively associated with smoking ($p = 0.00513$) when self-reported smoking statuses were used, but the association was only nominally significant when the classification was based on plasma cotinine concentration ($p = 0.0365$). A third SP (d_107.1) was notable because of a nominally significant positive association of the response with smoking ($p = 0.0132$, self-reported smoking statuses) and because of the agreement between nonaggregated technical replicates (see above).

The data set used for the above tests consisted of partially validated responses (detection of complete sequence tags) and imputed null responses (missing or incomplete sequence tags). An assessment of specificity was made by comparing partially and fully validated data sets. Full validation entailed quantitative analysis of the sequence tags and provided additional evidence that the responses were due to T3 peptides. For many of the SPs, partially validated responses were frequently retained in the fully validated data set. For d_107.1 and d_111.6, the frequencies were 68% and 84%, respectively. A notable exception was d_129.6, for which the frequency was only 3%. This low frequency suggested that the dominant substance(s) detected at d_129.6 were not T3 peptides.

Table 2. Technical Variation among Subsets of Nes$^+$ Controls at Selected SPs^a

Type of Variation	Coefficient of variation for responses at SP					
	d_71.1	d_107.1	d_111.6	d_125.1	d_156.6	d_161.1
Among-batch	0.45 (0.45)	0.49 (0.64)	0.83 (1.10)	0.18 (0.18)	0.53 (0.53)	0.48 (0.48)
Within-batch (A)	0.68 (0.68)	0.94 (1.21)	0.48 (0.48)	0.38 (0.38)	0.38 (0.38)	0.54 (0.54)
Within-batch (B)	0.95 (0.95)	0.34 (1.23)	0.26 (0.26)	0.28 (0.28)	0.23 (0.23)	0.61 (0.61)
TQ-MS only	0.71 (0.71)	0.34 (0.59)	0.31 (0.61)	0.20 (0.20)	0.34 (0.34)	0.61 (0.71)

^{a}Each value is a coefficient of variation, calculated with exclusion of false negatives (no parentheses), or from all relevant responses (in parentheses). The estimates of within-batch variation (A and B) come from different batches. The values for d_125.1 relate to the data in Figure 6.

In terms of magnitude, our responses were clustered at the lower end of the range reported by Li et al.[17] The apparent amounts of putative adducts (fractions, mol/mol, of total protein) were between ∼5 ppm and ∼0.1% in the present study (note, however, that the apparent amount was about one-tenth of the actual amount; see Figure 5B) and between ∼5 ppm and ∼5% in the earlier study. In both studies there were notable responses at d_72.1 and d_156.6 (our nomenclature). Li et al.[17] observed their highest responses at d_80.1, whereas we observed low responses at this SP; and the opposite was true of the responses at d_107.1. The origins of these inconsistencies (presumably subtle differences in samples and/or methods) remain unclear. We also found it instructive to consider the results of others' recent analyses using alternative methods.[32−34] Assuming no effects of differences in sample preparation or analytical conditions (exception: use of reducing agents precluding detection of S-thiolation), eight of the features detected by alternative methods would, if present, be detectable using FS-SRM. Notably, responses at d_107.1 are consistent with feature A25 observed by Grigoryan et al.[33] (d equivalent to C_7H_6O) and/or an unidentified feature detected by Chung et al.[32] Responses at d_156.1 are consistent with a putative tryptic transpeptidation product suggested by Chung et al.[32]

Whatever their specific mechanisms of formation, the putative adducts we observed must presumably have derived from reactive small molecules in the blood or been formed somehow during sample preparation. Reactive small molecules, or their precursors, may enter the blood from a variety of sources (pollution, drugs, endogenous sources, diet). The detection limit of our method corresponds to a low-μM concentration of adduct in plasma, which is similar to reported plasma concentrations of drugs, endogenous chemicals, and dietary chemicals, and is much greater than the plasma concentrations of most pollutants.[35] On this basis, a pollutant-derived adduct would not be detected by our method unless the adduct could accumulate to a level much higher than that of the pollutant. The naphthoquinone adducts reported by Lin et al.,[36] for example, would be need to be enriched approximately 10^3- or 10^4-fold to be detectable. On the other hand, Waidyanatha et al.[37] detected 1,4-benzoquinone adducts at a level that, providing d_{adduct} was close to d_{SP}, could potentially be detected by our method (although enrichment would be required in order to quantify the adduct reliably). The above inferences suggest that the putative adducts we observed are less likely to be of pollutants than of chemicals from the other sources, but further characterization will be needed to confirm their origins.

CONCLUSIONS

The FS-SRM methodology for analyzing distributions of HSA adducts was adapted for increased throughput. Technical aspects of the workflow were evaluated, and its ability to detect biological variation in samples of human plasma was explored. The mass coverage (69 Da ≤ d ≤ 213 Da) appeared to be essentially uninterrupted and therefore suitable for untargeted analysis. Responsiveness was non-uniform with respect to d, but this did not preclude comparisons between data sets. Accuracy and precision were best when the amount of adduct was >0.2% (w/w) of the total HSA and when d_{adduct} was close to d_{SP}. Levels of unknowns were often lower than 0.2% (w/w), and the high technical variation encountered at these levels (scatter, false negatives) had a negative effect on

reproducibility. This was addressable to an extent using novel data processing methods to aggregate sets of replicates. As a model environmental factor, tobacco smoking appeared to explain some of the variation observed in the human samples, even when smoking-associated differences in protein concentration were accounted for. The pilot study of tobacco smoke exposure provides a proof of principle; despite being underpowered, it demonstrates the possibility of detecting significant exposure-related differences in levels of putative adducts. Hence, FS-SRM could also be useful for detecting HSA adducts associated with exposures to other environmental factors, providing that the adducts are present in relatively high amounts. Knowledge of such associations will be valuable for studying human diseases because of the mechanistic relevance of the chemicals from which adducts derive. Since FS-SRM does not support detailed structural characterization, complementary analytical platforms will be required to fill in the mechanistic details. This endeavor will benefit from renewed efforts to enrich HSA adducts.

■ ASSOCIATED CONTENT

ⓈSupporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.chemrestox.7b00186.

> Full methods, characterization of Nns-T3, method validation results, tobacco smoke exposure data, BCA assay results, SDS-PAGE results, results of statistical analyses (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +44 (0) 20 7848 4569. E-mail: david.phillips@kcl.ac.uk.

ORCID

George W. Preston: 0000-0003-4812-139X

Marc Chadeau-Hyam: 0000-0001-8341-5436

■ ABBREVIATIONS

ASP-HPLC, analytical-scale preparative HPLC; BCA, bicinchoninic acid; Cam-iT3, isotopically labeled *S*-carbamidomethylated T3; *d*, net gain in mass; d_{adduct}, net gain in mass (adduct relative to unmodified protein); d_{SP}, net gain in mass (sampling point relative to unmodified protein); FS-SRM, fixed-step selected reaction monitoring; HSA, human serum albumin; NEM, *N*-ethylmaleimide; Nes, *N*-ethylsuccinimid-3-yl; Nns, *N*-(naphthalen-1-yl)succinimid-3-yl; Q1, first quadrupole; Q3, third quadrupole; ρ, Spearman's rank correlation coefficient; SI, Supporting Information; SP, sampling point; SPE, solid-phase extraction; T3, third-heaviest tryptic peptide; TCEP, tris(2-carboxyethyl)phosphine; TQ-MS, triple-quadrupole mass spectrometry.

■ REFERENCES

(1) Balbo, S., Turesky, R. J., and Villalta, P. W. (2014) DNA Adductomics. *Chem. Res. Toxicol. 27*, 356−366.

(2) Gan, L. S., Skipper, P. L., Peng, X. C., Groopman, J. D., Chen, J. S., Wogan, G. N., and Tannenbaum, S. R. (1988) Serum Albumin Adducts in the Molecular Epidemiology of Aflatoxin Carcinogenesis - Correlation with Aflatoxin-B1 Intake and Urinary-Excretion of Aflatoxin M1. *Carcinogenesis 9*, 1323−1325.

(3) Phillips, D. H., and Venitt, S. (2012) DNA and protein adducts in human tissues resulting from exposure to tobacco smoke. *Int. J. Cancer 131*, 2733−2753.

(4) Selvin, E., Steffes, M. W., Zhu, H., Matsushita, K., Wagenknecht, L., Pankow, J., Coresh, J., and Brancati, F. L. (2010) Glycated Hemoglobin, Diabetes, and Cardiovascular Risk in Nondiabetic Adults. *N. Engl. J. Med. 362*, 800−811.

(5) Ames, B. N., Shigenaga, M. K., and Gold, L. S. (1993) DNA Lesions, Inducible DNA-Repair, and Cell-Division: Three Key Factors in Mutagenesis and Carcinogenesis. *Environ. Health Perspect. 101*, 35−44.

(6) Rubino, F. M., Pitton, M., Di Fabio, D., and Colombi, A. (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.

(7) LoPachin, R. M., Gavin, T., DeCaprio, A., and Barber, D. S. (2012) Application of the Hard and Soft, Acids and Bases (HSAB) Theory to Toxicant-Target Interactions. *Chem. Res. Toxicol. 25*, 239−251.

(8) Noort, D., Hulst, A. G., de Jong, L. P. A., and Benschop, H. P. (1999) Alkylation of human serum albumin by sulfur mustard in vitro and in vivo: Mass spectrometric analysis of a cysteine adduct as a sensitive biomarker of exposure. *Chem. Res. Toxicol. 12*, 715−721.

(9) Pathak, K. V., Bellamri, M., Wang, Y., Langouet, S., and Turesky, R. J. (2015) 2-Amino-9*H*-pyrido[2,3-*b*]indole (AC) Adducts and Thiol Oxidation of Serum Albumin as Potential Biomarkers of Tobacco Smoke. *J. Biol. Chem. 290*, 16304−16318.

(10) Carlsson, H., and Tornqvist, M. (2016) Strategy for identifying unknown hemoglobin adducts using adductome LC-MS/MS data: Identification of adducts corresponding to acrylic acid, glyoxal, methylglyoxal, and 1-octen-3-one. *Food Chem. Toxicol. 92*, 94−103.

(11) Rappaport, S. M., Li, H., Grigoryan, H., Funk, W. E., and Williams, E. R. (2012) Adductomics: Characterizing exposures to reactive electrophiles. *Toxicol. Lett. 213*, 83−90.

(12) Sabbioni, G., and Turesky, R. J. (2017) Biomonitoring Human Albumin Adducts: The Past, the Present, and the Future. *Chem. Res. Toxicol. 30*, 332−366.

(13) Aldini, G., Regazzoni, L., Orioli, M., Rimoldi, I., Facino, R. M., and Carini, M. (2008) A tandem MS precursor-ion scan approach to identify variable covalent modification of albumin Cys34: a new tool for studying vascular carbonylation. *J. Mass Spectrom. 43*, 1470−1481.

(14) Peters, T., Jr. (1975) Serum Albumin. In *The Plasma Proteins* (Putnam, F. W., Ed.) pp 133−181, Academic Press, Inc., London.

(15) Peters, T., Jr (1970) Serum Albumin. *Adv. Clin. Chem. 13*, 37−111.

(16) Kanaly, R. A., Hanaoka, T., Sugimura, H., Toda, H., Matsui, S., and Matsuda, T. (2006) Development of the adductome approach to detect DNA damage in humans. *Antioxid. Redox Signaling 8*, 993−1001.

(17) Li, H., Grigoryan, H., Funk, W. E., Lu, S. S., Rose, S., Williams, E. R., and Rappaport, S. M. (2011) Profiling Cys(34) Adducts of Human Serum Albumin by Fixed-Step Selected Reaction Monitoring. *Mol. Cell. Proteomics 10*, M110.004606.

(18) Mann, M., and Wilm, M. (1994) Error Tolerant Identification of Peptides in Sequence Databases by Peptide Sequence Tags. *Anal. Chem. 66*, 4390−4399.

(19) Vineis, P., Chadeau-Hyam, M., Gmuender, H., Gulliver, J., Kleinjans, J., Kogevinas, M., Kyrtopoulos, S., Nieuwenhuijsen, M., Phillips, D. H., Probst-Hensch, N., Scalbert, A., Vermeulen, R., Wild, C. P., and Consortium, The EXPOsOMICS (2017) The exposome in practice: design of the EXPOsOMICS project. *Int. J. Hyg. Environ. Health 220*, 142−151.

(20) Janssen, B. G., Madhloum, N., Gyselaers, W., Bijnens, E., Clemente, D. B., Cox, B., Hogervorst, J., Luyten, L., Martens, D. S., Peusens, M., Plusquin, M., Provost, E. B., Roels, H. A., Saenen, N. D., Tsamou, M., Vriens, A., Winckelmans, E., Vrijens, K., and Nawrot, T. S. (2017) Cohort Profile: The ENVIRonmental influence ON early AGEing (ENVIRONAGE): a birth cohort study. *Int. J. Epidemiol. dyw269*, dyw269.

(21) Preston, G. W., and Phillips, D. H. (2016) Quantification of a peptide standard using the intrinsic fluorescence of tyrosine. *Anal. Bioanal. Chem. 408*, 2187−2193.

(22) Kim, S. (2016) Overview of Cotinine Cutoff Values for Smoking Status Classification. *Int. J. Environ. Res. Public Health 13*, 1236.

(23) R Core Team. (2016) *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.

(24) Kim, S., and Zhang, X. (2013) Comparative Analysis of Mass Spectral Similarity Measures on Peak Alignment for Comprehensive Two-Dimensional Gas Chromatography Mass Spectrometry. *Comput. Math. Methods Med. 2013*, 509761.

(25) Castagne, R., Delpierre, C., Kelly-Irving, M., Campanella, G., Guida, F., Krogh, V., Palli, D., Panico, S., Sacerdote, C., Tumino, R., Kyrtopoulos, S., Hosnijeh, F. S., Lang, T., Vermeulen, R., Vineis, P., Stringhini, S., and Chadeau-Hyam, M. (2016) A life course approach to explore the biological embedding of socioeconomic position and social mobility through circulating inflammatory markers. *Sci. Rep. 6*, 25170.

(26) Funk, W. E., Li, H., Iavarone, A. T., Williams, E. R., Riby, J., and Rappaport, S. M. (2010) Enrichment of cysteinyl adducts of human serum albumin. *Anal. Biochem. 400*, 61−68.

(27) Shultze, H. E., and Heremans, J. F. (1966) *Molecular Biology of Human Proteins*, Vol. 1, Elsevier, Amsterdam.

(28) Grigoryan, H., Li, H., Iavarone, A. T., Williams, E. R., and Rappaport, S. M. (2012) Cys34 Adducts of Reactive Oxygen Species in Human Serum Albumin. *Chem. Res. Toxicol. 25*, 1633−1642.

(29) Shaper, A. G., Wannamethee, S. G., and Whincup, P. H. (2004) Serum albumin and risk of stroke, coronary heart disease, and mortality: the role of cigarette smoking. *J. Clin. Epidemiol. 57*, 195−202.

(30) Ganotakis, E. S., Vrentzos, G. E., Gazi, I. F., Papadakis, J. A., Jagroop, A., Paraskevas, K. I., Nair, D. R., and Mikhailidis, D. P. (2007) Fibrinogen, lipoprotein (a), albumin and bilirubin (F-L-A-B) levels and cardiovascular risk calculated using the Framingham equation. *In Vivo 21*, 685−694.

(31) Wrzesniak, M., Kepinska, M., Krolik, M., and Milnerowicz, H. (2016) The Influence of Tobacco Smoke on Protein and Metal Levels in the Serum of Women during Pregnancy. *PLoS One 11*, e0161342.

(32) Chung, M. K., Grigoryan, H., Iavarone, A. T., and Rappaport, S. M. (2014) Antibody Enrichment and Mass Spectrometry of Albumin-Cys34 Adducts. *Chem. Res. Toxicol. 27*, 400−407.

(33) Grigoryan, H., Edmands, W., Lu, S. X. S., Yano, Y., Regazzoni, L., Iavarone, A. T., Williams, E. R., and Rappaport, S. M. (2016) Adductomics Pipeline for Untargeted Analysis of Modifications to Cys34 of Human Serum Albumin. *Anal. Chem. 88*, 10504−10512.

(34) Porter, C. J., and Bereman, M. S. (2015) Data-independent-acquisition mass spectrometry for identification of targeted-peptide site-specific modifications. *Anal. Bioanal. Chem. 407*, 6627−6635.

(35) Rappaport, S. M., Barupal, D. K., Wishart, D., Vineis, P., and Scalbert, A. (2014) The Blood Exposome and Its Role in Discovering Causes of Disease. *Environ. Health Perspect. 122*, 769−774.

(36) Lin, P.-H., Chen, D.-R., Wang, T.-W., Lin, C.-H., and Chuang, M.-C. (2009) Investigation of the cumulative tissue doses of naphthoquinones in human serum using protein adducts as a biomarker of exposure. *Chem.-Biol. Interact. 181*, 107−114.

(37) Waidyanatha, S., Yeowell-O'Connell, K., and Rappaport, S. M. (1998) A new assay for albumin and hemoglobin adducts of 1,2- and 1,4-benzoquinones. *Chem.-Biol. Interact. 115*, 117−139.