Analytical methods based on liquid chromatography for the analysis of albumin adducts involved in retrospective biomonitoring of exposure to mustard agents

Lorenzo Avigo, Florine Hallez, Audrey Combès, Charlotte Desoubries, Christine Albaret, Anne Bossée, Valérie Pichon

Analytical methods based on liquid chromatography for the analysis of albumin adducts involved in retrospective biomonitoring of exposure to mustard agents

Lorenzo Avigo[1,3], Florine Hallez[1], Audrey Combès[1], Charlotte Desoubries[2], Christine Albaret[2], Anne Bossée[2], Valérie Pichon[1,3,*]

[1]Department of Analytical, Bioanalytical Sciences and Miniaturization (LSABM) Chemistry, Biology and Innovation (CBI), ESPCI Paris, PSL University, CNRS, 10 rue Vauquelin, 75005 Paris, France

[2]DGA, CBRN Defence, 5 rue Lavoisier, 91710 Vert-le-Petit, France

[3]Sorbonne Université, 4 Place Jussieu, 75005 Paris, France

*Corresponding author: Professor Valérie Pichon

E-mail address : valerie.pichon@espci.fr

Abstract

The present review has for objective to list, describe, compare and critically analyze the main procedures developed in the last twenty years for the analysis of digested alkylated peptides, resulting from the adduction of albumin by different mustards, and that can be used as biomarkers of exposure to these chemical agents. While many biomarkers of sulfur mustard, its analogues and nitrogen mustards can easily be collected in urine such as their hydrolysis products, albumin adducts require blood or plasma collection to be analyzed. Nonetheless, it benefits of a wider period of detectability in human exposed patients than the previously mentioned metabolites with detection up to 25 days after exposure to the chemical agent. The detection of these digested alkylated peptides of adducted albumin constitutes an unambiguous proof of exposure. However, their determination, especially when they are present at very low concentration levels, can be very difficult due to the complexity of the biological matrices. Therefore, numerous sample preparation procedures to extract albumin and to recover alkylated peptides after a digestion step using enzymes have been proposed prior to the analysis of the targeted peptides by liquid chromatography coupled to mass spectrometry method with or without derivatization step. This review describes and compares the numerous procedures including a number of different steps for the extraction and purification of adducted albumin and its digested peptides described in the literature to achieve detection limits for biological samples exposed to sulfur mustard, its analogues and nitrogen mustards in the ng/mL range.

Keywords

Albumin adducts; Biomarkers; Mustard exposure; Vesicants; HPLC-MS/MS

Abbreviations

[N1-HETE]-His; N1-[2-[(2-hydroxyethyl)thio]ethyl]-Histidine
[N3-HETE]-His; N3-[2-[(2-hydroxyethyl)thio]ethyl]-Histidine
ACN; Acetonitrile
AE-[HETE]V-SKL; Alanine-Glutamic acid-[2-[(2-hydroxyethyl)thio]ethyl]-Valine-Serine-Lysine-Leucine
CWC; Chemical weapons convention
Cys34; 34-cystein residue
DNA; Desoxyribonucleic acid
GSH; Glutathion
HETE; hydroxyethylthioethyl
HETE-Asp; [2-[(2-hydroxyethyl)thio]ethyl]Asparagine
HETE-CP; [2-[(2-hydroxyethyl)thio]ethyl]-Cysteine-Proline
HETE-CPF; [2-[(2-hydroxyethyl)thio]ethyl]-Cysteine-Proline-Phenylalanine
HETE-CPY; [2-[(2-hydroxyethyl)thio]ethyl]-Cysteine-Proline-Tyrosine

HETE-Glu; [2-[(2-hydroxyethyl)thio]ethyl]-Glutamine

HETE-Lys; [2-[(2-hydroxyethyl)thio]ethyl]-Lysine

IMER; Immobilized Enzyme Reactor

LC-MS; Liquid chromatography coupled to mass spectrometry

LOD; Limit of detection

LQQC*PFED; Leucine-Glutamine-Glutamine-[2-[(2-hydroxyethyl)thio]ethyl]-Cysteine-Proline-Phenylalanine-Glutamic acid-Aspartic acid

LQQC*PFEDHVKL; Leucine-Glutamine-Glutamine-[2-[(2-hydroxyethyl)thio]ethyl]-Cysteine-Proline-Phenylalanine-Glutamic acid-Aspartic acid-Histidine-Valine-Lysine-Leucine

RBCs; Red Blood Cells

SM; Sulfur mustard

SMO; Sulfur mustard sulfoxide

TDG; Thiodiglycol

TDGO; Thiodiglycol sulfoxide

TRIS; Trisaminomethane

YLQQC*PFED; Tyrosine-Leucine-Glutamine-Glutamine-[2-[(2-hydroxyethyl)thio]ethyl]-Cysteine-Proline-Phenylalanine-Glutamic acid-Aspartic acid

Introduction

Sulfur mustard (SM) is a strong vesicant and alkylating compound used as a chemical warfare agent for the first time during World War I (1). Causing severe burns to the eyes, lungs and skin of exposed victims, sulfur mustard is also a carcinogenic and cytotoxic chemical whose synthesis, storage and use have been forbidden since 1997 with the entry into force of the Chemical Weapons Convention (CWC). However, it remains a threat for populations. Indeed, old sulfur mustard ammunitions are still present in the environment, thus causing accidental exposures to this chemical warfare agent (2–4). Moreover, because of its simplicity of production and its lack of antidote, sulfur mustard can also be used during military conflicts such as the Syrian Arab Republic conflict in 2015 (5–7) or can be a weapon used in chemical terrorism (8–10).

An episulfonium ion is quickly formed after administration of sulfur mustard in the organism by an intramolecular nucleophile substitution of one of the chloride atoms by the sulfur atom, which makes intact sulfur mustard really difficult to detect *in vivo* (11). The episulfonium ion is highly reactive and reacts with various nucleophiles present in the organism, such as water, glutathione, or nucleophile sites of DNA and proteins, leading to the formation of metabolites or adducts which life span is longer than the one of intact sulfur mustard (from several days up to several months) thus explaining the interest of their determination as biomarkers of exposure in biological samples (urine, blood, tissues).

The biomarkers of sulfur mustard are conventionally categorized depending on the targeted nucleophiles: (i) the metabolites resulting from the hydrolysis and/or oxidation reactions of sulfur mustard (12), (ii) the metabolites resulting from the conjugation with glutathione (3,13), (iii) the DNA adducts (14), (iv) the adducts resulting from the alkylation of proteins (15). A summary of formation pathways of the main sulfur mustard biomarkers is reported on the Fig. 1.

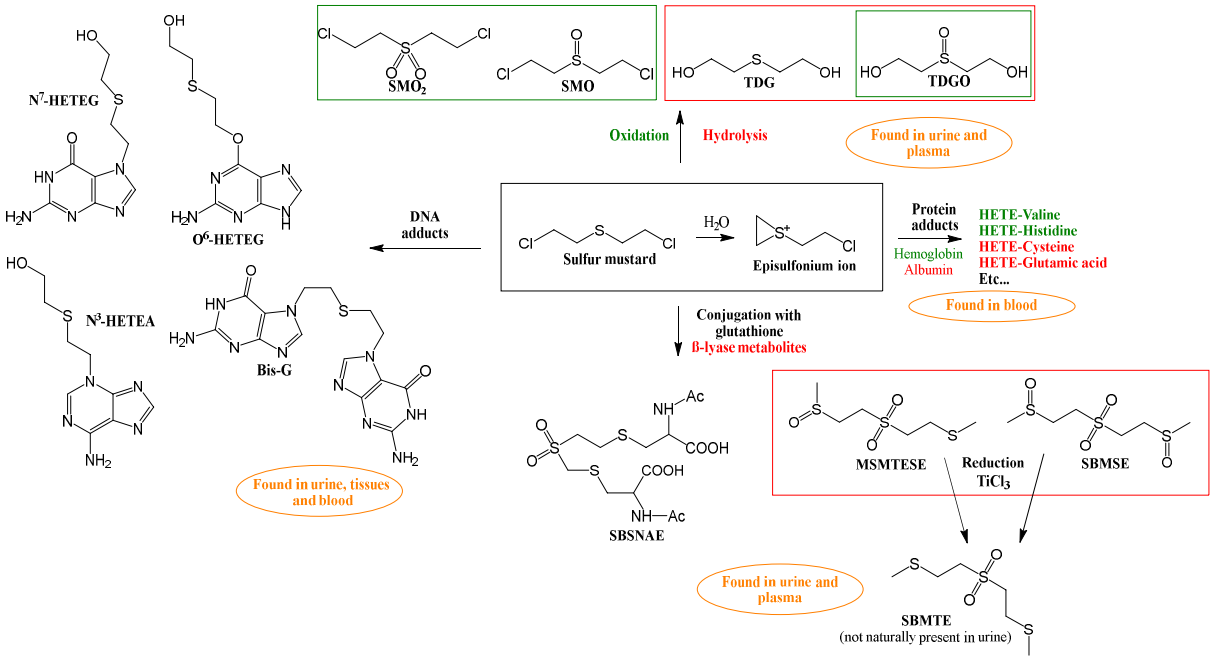


Fig. 1 *Representation of the main sulfur mustard biomarkers resulting from the hydrolysis and/or oxidation reactions, from conjugation with glutathione or from DNA or proteins alkylation. SMO_2: Sulfur mustard sulfone; SBSNAE: 1,1'-sulfonylbis[2-S-(N-acetylcysteinyl)ethane]; SBMTE: 1,1'-sulfonylbis[2-(methylthio) ethane]; MSMTESE: 1-methylsulfinyl-2-[2-(methylthio)ethylsulfonyl]ethane; SBMSE: 1,1'-sulfonylbis[2-(methylsulfinyl) ethane]; N^7-HETEG: N^7-[2-[(2-hydroxyethyl)thio]ethyl]-guanine; O^6-HETEG: O^6-[2-[(2-hydroxyethyl)thio]ethyl]-guanine; N^3-HETEA: N^3-[2-[(2-hydroxyethyl)thio]ethyl]-adenine; BIS-G: bis[2-(guanine-7-yl)ethyl] sulfide; HETE-Valine: Hydroxyethylthioethyl valine; HETE-Histidine: Hydroxyethylthioethyl histidine; HETE-Cysteine : Hydroxyethylthioethyl cysteine; HETE-Glutamic acid: Hydroxyethylthioethyl glutamic acid; TDG : Thiodiglycol; TDGO: Thiodiglycol sulfoxide; $TiCl_3$: Titanium trichloride; Ac: acetyl group ($COCH_3$). Adapted from (16).*

The free metabolites (TDG, TDGO, SMO, SMO_2, SBSNAE, MSMTESE, SBMSE) are mainly present in the urine, but can also be found in plasma (17). Their analysis in urine is one of the easiest ways to confirm an exposure to sulfur mustard (SM) (18,19) since most of the absorbed dose of the alkylating agent is eliminated in the form of urinary metabolites whose concentrations is higher than that of plasma or serum (20). In addition to that, the sample collection is noninvasive and does not require trained staff. Unfortunately, urinary excreted metabolites require quick sample collection after SM exposure due to their fast elimination by the organism (within the range of few days). This condition is of major importance since biomedical samples cannot always be collected and analyzed immediately after exposure to a chemical warfare agent.

To satisfy the sample collection duration criteria, it is possible to monitor the adduction of SM on macromolecules such as DNA and proteins as SM reacts with nucleophile sites to form long-lived hydroxyethylthioethyl (HETE) adducts compared to short-lived previously mentioned free metabolites (4). Most analytical procedures developed for urinary metabolites are simpler and less time-consuming compared to those used for the analysis of adducts formed with macromolecules present in blood and tissues (21) with the exception of DNA adducts that can be found in urines. Nevertheless, contrary to TDG and TDGO metabolites, no trace of DNA-SM or protein-SM adducts is detected in biological samples collected from people not exposed to this chemical agent making them more suitable for unambiguous identification of exposure (22).

Regarding proteins, several amino acids contain nucleophile groups, such as cysteine (SH), serine and tyrosine (OH), lysine (NH_2), glutamic and aspartic acids (CO_2H) (23) and methionine ($S-CH_3$) (24). Covalent adducts with proteins are potential long-lived biological indicators of exposure. Particularly useful biomarkers are the adducts formed with hemoglobin and albumin, two major blood proteins. Contrary to alkylated DNA, hemoglobin and albumin adducts are generally stable, DNA being repaired and DNA-adducts excreted through urine with

detectability of around 7 days to few weeks (14,25). Most of adducted proteins have life span similar to that of native proteins (approximatively 120 days for human hemoglobin and 20-25 days for human albumin) (20,23,26).

The most common and most adapted biomarkers of exposure to sulfur mustard for retrospective detection are protein adducts thanks to their long life time. In blood, sulfur mustard binds approximatively 1,000 times more with proteins (hemoglobin and albumin) than DNA due to the larger amount of proteins in blood (27). Albumin, the main plasma protein, is also very abundant in blood with an average concentration range from 30 to 45 $mg.mL^{-1}$. Human serum albumin is a protein composed of a single polypeptide chain containing 585 amino acids and 17 disulfide bonds, with a molecular weight of 66.5 kDa (28). A radiochemical study conducted by exposing blood in vitro with ^{14}C-labelled sulfur mustard proved that albumin adduction increased linearly with the mustard concentration (29). This conclusion was also reached using mass spectrometry to determine the alkylation sites of sulfur mustard on albumin (30). This study also highlighted the fact that, upon exposure to the chemical agent, almost 20% of the sulfur mustard dose was covalently linked to albumin. By analyzing the tryptic digest of albumin extracted from blood exposed to ^{14}C labelled sulfur mustard and isolated by sequential precipitation (31), it was found that the protein was alkylated on the 34-cystein residue (Cys^{34}) (32). The amino acid Cys^{34} is the only protein cysteine residue presenting a free thiol function and has a relatively low pK_a due to the intramolecular stabilization of the thiolate ion.

Numerous methods were developed in the past decades for the detection of protein-SM adducts in blood samples collected from victims exposed to this chemical agent. While in general the analytical method consists of reverse phase liquid chromatography-tandem mass spectrometry analysis of alkylated peptides which requires an enzymatic digestion step of albumin-SM adduct to obtain them, the complexity of blood and serum samples has led to the implementation of different sample preparation steps. The reported procedures include protein precipitation, affinity chromatography, filtration, and dilution, more or less combined with each other in different orders, carried out before and after the digestion step as illustrated in the schematic diagram in Fig. 2. This review discussed the reported methods applied firstly to the analysis of albumin-SM adducts that are summarized in the Table 1 and secondly to the analysis of SM-analogues and nitrogen mustards adducts which were less studied.

Table 1 *Sample preparation methods and LC-MS methods developed for the analysis of albumin-SM adducts in biological matrices.*

Albumin adduct	**Sample**	**Albumin isolation**	**Enzymatic digestion**	**Adduct isolation**	**Derivati-zation**	**Internal standard**	**LC conditions**	**MS analyzer**	**LOD of SM in real samples**	**Ref**
HETE-CPY	Blood (rat) 0.525 mL (after C)	C Dilution 1/4 (50 mM KH_2PO_4) Filtration (0.45 µm) AC	pronase, 37°C 2h	UF	-	d_8-SM adducted to albumin	C_{18} (150 x 1 mm, 3 µm), - V_{inj}= 50 µL	Q ToF	-	(33)
HETE-CPF	Blood (human) 5 mL	C PP (sodium acetate/EtOH (5/95, v/v))	pronase, 37°C 2.5h	UF (R=100%) SPE (C18) (R=87%)	-	-	C_{18} (350 x 0.32 mm, -), - V_{inj}= 40 µL	TQ	1.6 ng.mL^{-1}	(32)
HETE-CPF	Blood (human) 1 mL (after C)	C AC	pronase, 37°C 2h	UF	-	d_8-SM adducted to albumin	C_{18} (150 x 1 mm, 5 µm), - V_{inj}= 50 µL	TQ,	0.7 ng.mL^{-1} (LOQ)	(30)
HETE-CPF	Blood (human) 2 mL	C AC	pronase, 37°C, 2h	UF SPE (C_{18})	-	-	C_{18} (150 x 1 mm, 5 µm), 30°C V_{inj}= 20 µL	TQ	8 ng.mL^{-1}	(34)
HETE-CPF	Plasma (human) 10 µL	Dilution 1/10 (50 mM NH_4HCO_3)	proteinase K, 50°C, 1h	SPE (Oasis® HLB) (R=86.3%)	-	$^{13}C_9$ -SM peptide	C18 (50 x 2.1 mm, 1.3 µm), 70°C V_{inj}= 2,5 µL	TQ	1,06 ng.mL^{-1} (HETE-CPF spiked in matrix)	(35)
HETE-CPF	Plasma (human) 200 µL	C Dilution 1/4 Filtration (0.45 µm) AC PP (Acetone)	pronase, 37°C, 90 min	UF SPE (C_{18})	-	d_8-SM adducted to albumin	C_{18} (100 x 2.1 mm, 2.6 µm), 30°C V_{inj}= 50 µL	QTRAP	0.79 ng.mL^{-1}	(36)
HETE-CPF	Plasma (human) 100 µL	AC (R=96 ± 3.5 %) UF	proteinase K, 50°C, 4h	UF	-	2-CEES adducted to albumin	C_{18} (50 x 2.1 mm, 1.8 µm), 40°C V_{inj}= 20 µL	TQ	0.532 ng.mL^{-1}	(37)

HETE-CPF	Plasma (human) 50 µL	C PP (Acetone)	proteinase K, 50°C, 90 min, pH 7.8	SPE (Oasis® HLB)	-	d_8-SM adducted to albumin	C_{18} (50 x 2.1 mm, 1.3 µm), 70°C V_{inj}= 5 µL	TQ	1.74 ng.mL^{-1}	(38)
HETE-CPF	Plasma (human) 100 µL	PP (Acetone)	proteinase K, 50°C, 90 min, pH 7.8 (R=62% for pure albumin)	SPE (Oasis® HLB)	-	-	C_{18} (100 x 2.1 mm, 1.7 µm), 70°C V_{inj}= 50 µL	Q-Orbitrap (Exactive)	1 ng.mL^{-1}	(39)
HETE-CP	Plasma (human) 100 µL	AC UF	pronase, 37°C, 2h	UF	-	d_8-SM adducted to albumin	C_{18} (150 x 1.0 mm, 3 µm), 50°C V_{inj}= 20 µL	QTRAP	8-16 ng.mL^{-1}	(40)
HETE-CP	Plasma (human) 100 µL	UF	pronase, 37°C, 4h	UF	-	d_3-Atr	C_{18} (50 x 1.0 mm, 5 µm), 60°C V_{inj}= 20 µL	Q ToF	1.56 ng.mL^{-1}	(41)
HETE-CP	Plasma (human) 50 µL	PP (Acetone)	pronase, 50°C, 2h	PP (ACN)	Cbz-Cl (R=46%)	2-CEES adducted to albumin	C18 (150 x 2.1 mm, 1,8 µm), 30°C V_{inj}= 20 µL	TQ	0.500 ng.mL^{-1}	(42)
HETE-CP	Plasma (human) 50 µL	PP (Acetone)	pronase, 50°C, 2h	PP (ACN)	nic-NHS (R=81%)	d_3-Atr	C18 (50 x 1.0 mm, 1.8 µm), 60°C V_{inj}= 20 µL	Q ToF	12.7 ng.mL^{-1} (LOI)	(43)
HETE-CP	Plasma (human) 50 µL	PP (Acetone)	pronase, 50°C, 2h	PP (ACN)	PA (R=100%)	d_3-Atr	C18 (50 x 1.0 mm, 1.8 µm), 60°C V_{inj}= 20 µL	QTRAP	5.1 ng.mL^{-1} 25.4 ng.mL^{-1} (LOI)	(44)
HETE-Glu, HETE-His	Serum (chicken, duck, ostrich) 190 µL	Dilution 1/5 (50 mM NH_4HCO_3)	pronase, 50°C, 30 min (HETE-Glu), 180 min (HETE-His)	PP (ACN)	-	d_8-SM adducted to albumin	C18 (150 x 1 mm, 3 µm, 100 Å), 60°C V_{inj}= 20 µL	QTRAP	-	(45)

HETE-Asp, HETE-Glu, [N1-HETE]-His, [N3-HETE]-His, HETE-Lys	Plasma (human) 50 µL	PP (Acetone)	pronase, 50°C, 1h	PP (ACN)	PA (R=100%)	2-CEES adducted to albumin	C_{18} (150 x 2,1 mm, id 1,8 µm), 30°C V_{inj}= 20 µL	TQ	1.00 ng.mL^{-1}	(46)
LGM(-HETE)F	Serum (human) 100 µL	Dilution 1/2.5 (50 mM NH_4HCO_3)	pepsin, 37°C, 4h	UF	-	d_3-Atr	C_{18} (150 x 1,0 mm, id 3 µm, 100 Å), 40°C V_{inj} = 20 µL	QToF	5.14 ng.mL^{-1}	(24)
AE-(HETE)-VSKL	Blood (human) < 1 mL	C AC	pepsin, 5 min (on-line IMER)	None	-	-	C_{18} (150 x 1 mm, 3 µm), - V_{inj} = 10 µL	Q ToF	-	(47)

$^{13}C_9$-SM : SM containing ^{13}C isotopes; 2-CEES: 2-chloroethylethylsulfide; AC: Affinity Chromatography ; d_3-Atr : Deuterated atropine; C: Centrifugation; Cbz-Cl: Benzyl chloroformate; d_8-SM: deuterated SM; LC: Liquid chromatography; LOI: Limit of Identification; MS: Mass spectrometry; nic-NHS: 1-nicotinoyloxy-succinimide; PA: propionic anhydride; PP: Protein Precipitation; Q Exactive: Hybrid Quadrupole-Orbitrap; Q ToF: Hybrid Quadrupole-Time of Flight; QTRAP: Hybrid Quadrupole-Ion trap; R: Recovery yield; SM: Sulfur mustard; SPE: Solid phase extraction; TQ: Triple quadrupole; UF: UltraFiltration; V_{inj}: Injected volume.

The mobile phase used for reverse phase liquid chromatography was made of an ACN gradient with formic acid as an additive. All the LC-MS/MS methods used ElectroSpray Ionization in positive mode coupled with TQ analyzer used in Multiple Reaction Monitoring (MRM) mode or Q-ToF and Orbitrap analyzers used in pseudo-MRM mode. The LOD/LOQ/LOI values correspond to the concentration of exposure of whole blood or plasma to SM. The S/N ratio of LOI was not given in the articles.

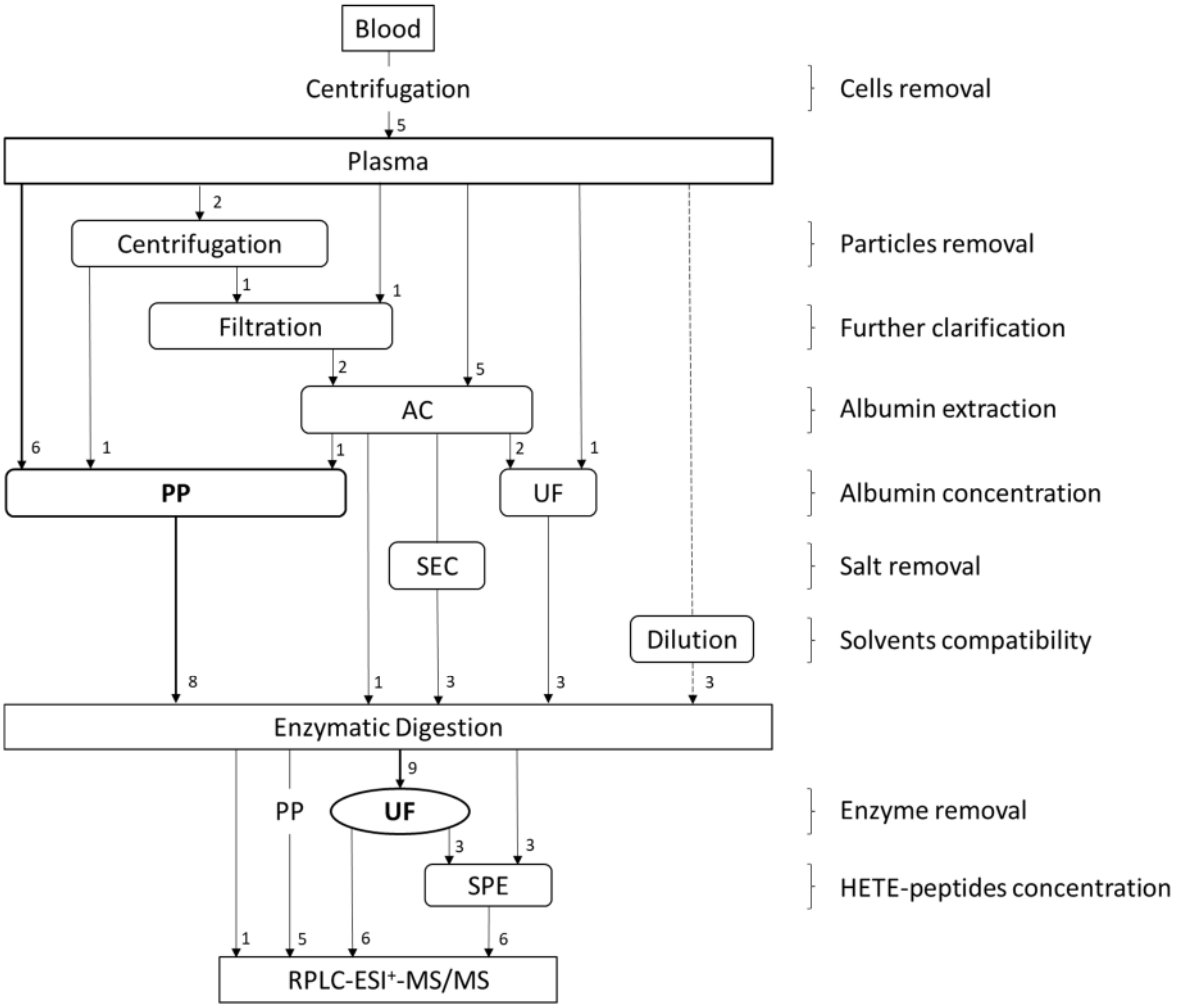


Fig. 2 *Synthesis of the sample preparation methods developed for the analysis of SM-albumin adducts in biological matrices. The figures next to the arrows indicate how many works used the pathway. PP: Protein Precipitation; AC: Affinity Chromatography; UF: Ultra-Filtration; SEC: Steric Exclusion Chromatography; SPE: Solid Phase Extraction; SM: Sulfur mustard*

Albumin extraction/purification

Albumin is located in the blood, more exactly in the plasma. A simple centrifugation step was then applied to isolate plasma from the red blood cells (RBCs) in the cases where studies were conducted using whole blood (32–34,47). The extraction of albumin from plasma was mostly performed by protein precipitation (32,39,42–44,46), mainly using acetone, followed by a centrifugation to form a pellet. The supernatant is then discarded and the protein pellet containing the adducted albumin is redissolved in a small volume of ammonium bicarbonate solution (32,36,38,39,42–44,46), i.e. a medium that is fully compatible with the enzymatic digestion using proteinase K or pronase of the recovered proteins.

Additionally, to the centrifugation performed to remove the cells for the teams working on blood, a secondary centrifugation can be achieved to remove any particles that may remain in the plasma (36,38). This was only witnessed when working on 96 well plates that can easily clog and require particle-less samples. A filtration on 0.45 µm filter was also performed for further clarification of the filtrate but this additional step was proposed before the treatment of the centrifuged plasma sample by affinity chromatography on Blue Sepharose sorbent that exists under different names depending on the supplier. This sorbent was selected for its high affinity for albumin thus allowing to isolate it from other proteins (36,48). This high affinity for albumin has led various groups to use this affinity sorbent directly for its extraction from plasma (30,33,34,36,37,40,47,48). The efficiency of this technique was measured and 96 ± 3.5 % of total non-adducted albumin was extracted from a non-exposed plasma (37). Albumin purity was also measured after extraction on different commercially available sorbents with values

ranging from 64% to 88% (48). It can be noticed that the structure of Blue Sepharose is not well communicated by the suppliers making it difficult to determine by which mechanism of retention albumin is retained on the sorbent and the degree of specificity of this sorbent for albumin compared to other proteins such as some interferons that can also be retained by this sorbent (49).

The solution used to elute albumin from the affinity sorbent consists of a phosphate buffer containing a huge amount of chloride salt (1.5-2 M). Therefore, the use of a desalting device, a filtration gel allowing to recover albumin by exclusion chromatography (30,33,34), was implemented to remove salts and to transfer albumin into a medium adapted to the digestion step, i.e. an ammonium bicarbonate buffer. An alternative is to perform protein precipitation instead of exclusion chromatography before the digestion step (36). Some groups also proposed to performed an ultrafiltration step on 10 kDa cut-off filters (37,40). The concentration here is made by getting rid of solvent and all the low molecular weight compounds in the plasma while retaining the purified albumin (Mw = 66.5 kDa) on the filter. Then, the filter is washed and the albumin can be concentrated by the same way as when using protein precipitation. This ultrafiltration step was also proposed as a unique sample preparation method directly applied to plasma to concentrate large molecules including albumin while removing the other plasma components (small size molecules, salts…) (41). Whereas Hallez *et al.* noted, when working on adducted SM-hemoglobin, differences in recoveries depending on the origin of filters for the same value of cut-off 100 kDa (50), no comments related to this problem have been raised in studies with albumin. Between different suppliers, yields up to 5 times greater were observed in their procedure so the efficiency of the filtration must be measured. Finally, simple dilution of plasma sample before the digestion step was also proposed using buffer adapted to the enzyme (24,35,45).

It is important to mention that Blue Sepharose shows different affinities towards albumin based on the animal species (49) as some teams worked on human and rat blood as well as avian serum. If targets are changed, it is thus required to measure the efficiency of the purification. Electronic Supplementary Material Table S1 illustrates the diversity of amino acid sequences containing the alkylable Cys site in albumin from different species (45,46). So, if pronase or proteinase K (specific cleavage after aliphatic and aromatic amino acids (51)) are used to monitor albumin adducts for other species, the CP and CPF peptides will not be observed and albumin digestion may give rise to different peptides.

Unfortunately, the efficiency for each step has not been calculated and it is therefore difficult to compare the efficiency of the different routes proposed by the different teams. Furthermore, in the case where four treatment steps prior to digestion were implemented, a decrease in the overall albumin recovery rate may be expected, so unnecessary steps should be avoided.

It should be noted that many groups have worked on SM-albumin adducts, and several papers have been published by the same teams. Unfortunately, the choices of adding or removing steps were never discussed in their papers. For example, one of these teams has proposed to replace the albumin concentration step by protein precipitation (38) by a simple dilution of the plasma before enzymatic digestion (35) without any explanation and the lack of data on albumin recovery makes it difficult to assess the value of this simplified procedure especially as limits of detection achieved at the end of the alkylated peptides analysis (using the same LC-MS method) are similar. The reason for this modification in the procedure is likely for reducing the overall process time. Ultrafiltration has been replaced by protein precipitation (41,43) but again there is a lack of quantitative data to assess whether the change in protocol has led to significant improvements, although by this method, SM-albumin adducts were detected up to 15 days after a real case of SM exposure.

Digestion of adducted albumin in peptides

The analysis of intact proteins by LC-MS is not trivial as it requires adapted LC column to provide either the retention or exclusion of proteins because of their large size and high-resolution MS (52). Originally, adducted proteins were modified to make the analysis of the biomarker easier using the Edman degradation procedure (29). This technique was originally developed for the study of SM-hemoglobin adducts. The alkylated N-terminal valine in hemoglobin splits off through Edman degradation. The cleavage of the alkylated amino acid is due to the use of pentafluorophenyl isothiocyanate (incubation at 45°C for 2h) followed by the derivatization of the resulting

thiohydantoin derivative with heptafluorobutyrylimidazole prior to chromatographic analysis of the alkylated-valine residue (53). An alternative of this approach was recently proposed by the bottom up approach that consists in digesting hemoglobin using a protease thus simplifying the whole analytical procedure and giving rise to more sensitive methods (50). For the study of SM-albumin adducts, the use of an enzymatic digestion has provided a less laborious procedure together with more sensitivity (32) compared to hemoglobin. The method based on the digestion with an enzyme allows to detect concentration levels 10-time lower than when using Edman procedure adapted for the albumin degradation to produce the alkylated peptides (32). Logically, Edman procedure has not been reported in the most recent papers. In return, as illustrated by data reported in Table 1, two enzymes were mainly used: proteinase K (Mw around 29 kDa) and pronase which is a mixture of several enzymes (Mw between 20 kDa and 60 kDa). When using proteinase K, albumin is digested leading to the alkylated tripeptide (S-[2-[(hydroxyethyl)thio]ethyl])-Cys-Pro-Phe (HETE-CPF) for human albumin and to the alkylated tripeptide (S-[2-[(hydroxyethyl)thio]ethyl])-Cys-Pro-Tyr (HETE-CPY) when working on rat albumin as the protein differs from the human form (33) (Fig. 3). When using pronase, the alkylated dipeptide (S-[2-[(hydroxyethyl)thio]ethyl])-Cys-Pro (HETE-CP) is obtained. With the exception of one case for which a TRIS buffer was mentioned (37), the enzymatic digestion was performed in an ammonium bicarbonate buffer (NH_4HCO_3, 50 mM).

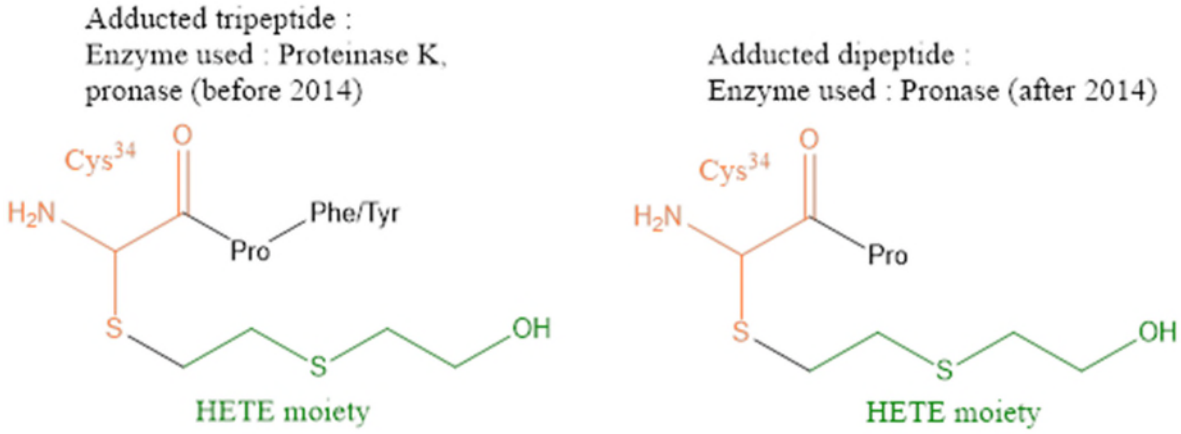


Fig. 3 *Structure of the tripeptide HETE-CPF/HETE-CPY (on the left) and dipeptide HETE-CP (on the right) obtained by tryptic digestion of albumin alkylated by sulfur mustard on the Cys^{34} residue. In rats, the phenylalanine residue (Phe) is replaced by a tyrosine residue (Tyr) leading to the tripeptide HETE-CPY.*

It is interesting to notice that for the papers published before 2014, pronase was used and the quantified alkylated peptide was the tripeptide HETE-CPF or HETE-CPY (30,32–34,36) and not the dipeptide HETE-CP. Past this date, all the studies involving pronase and studying the ^{34}Cys alkylated position focus on the detection of dipeptide HETE-CP. It has been proven that enzyme activity can change depending on the supplier (39). Indeed, Braun *et al*. have compared the recovery of the alkylated tripeptide between four different enzymes. Results are summarized in Electronic Supplementary Material Table S2. They showed up to 2-fold different yields depending on the origin of the enzyme. Indeed, with Qiagen® proteinase K, digestion rates of 62% were reached while this rate was 30% with SibEnzim® proteinase K or 42% with Applichen Panreac proteinase K. They also noticed that pronase did not allow to obtain the tripeptide. The digestions were performed in the same conditions. It was also reported that the activity may differ between batches from the same supplier (35). This may explain the difference in targeting the HETE-CPF and later the HETE-CP when using pronase as the composition of the mixture may have changed over the years. Indeed, as described earlier, pronase being a mixture of proteases, its activity can be variable depending of its composition.

Given that only one study reports digestion rates, it is impossible to conclude that one enzyme leads to better digestion rates than the other. The average digestion time is similar for both enzymes (around 2h) with up to 4h for two studies (37,41). One major difference is the temperature set for the digestion. Pronase tends to be used at 37°C while proteinase K is always used at 50°C. It must be pointed out that, recently, some teams also performed enzymatic digestion using pronase at 50°C (42–44,46). When performing the pronase digestion at higher temperatures (50-60°C), the concentration of alkylated peptides increased more rapidly with equivalent concentrations of albumin adducts obtained 5 times quicker up to 10 times quicker at higher temperatures

compared to the figures at 37°C (45). It should be noted that this observation was made when studying alkylated amino acids and not alkylated di- or tripeptides. In the case of John *et al.*, they performed digestion with pronase at 37°C for 4 hours (41) and later reported a duration of 2 hours but at 50°C (43,44) which is not the optimal temperature for the enzyme. Due to the lack of digestion rate data and the fact that the LC-MS conditions applied in these studies were very different, it was not possible to evaluate the performance of the digestion conditions using obtained limits of detection (LOD) values.

In all the previous articles the digestion is performed in solution, which means that all the reagents are placed in a vial and incubated for a given time and at a specific temperature. It is known that auto digestion of the enzyme can occur in solution which forces to work with a relatively low enzyme to substrate ratio, thus increasing the digestion times (54). One solution to this problem is to graft the enzyme onto a solid support and to apply the resulting reusable sorbent, called immobilized enzyme reactor (IMER), to the digestion of proteins. This approach allows to considerably reduce the amount of reagents used and the auto digestion phenomenon and greatly the digestion time while also increasing the repeatability of the method (47,54). Hallez *et al.* developed a trypsin reactor for the SM-hemoglobin adducts analysis that was reusable up to 60 times (54). For SM-albumin adducts, Carol Visser *et al*. proposed the use of a pronase-based IMER (47). However, they could not maintain good pronase activity after grafting, which led them to investigate another enzyme, pepsin, which generated other alkylated peptides that could be used as a biomarker than the dipeptide HETE-CP usually obtained through pronase digestion. Indeed, after a digestion carried out in formic acid on the commercialized IMER, i.e. a digestion medium adapted to this protease, they observed several alkylated peptides such as YLQQC-[HETE]-PFED, LQQC-[HETE]-PFED and LQQC-[HETE]-PFEDHVKL, these 3 peptides being alkylated on ^{34}Cys. The MS signal of these modified cysteine peptides were not the most intense ones leading the authors to choose another alkylated peptide AE-[HETE]-VSKL as the biomarker for their study. AE-[HETE]-VSKL is alkylated on a glutamic acid residue known to be a potential site for SM alkylation. Using the IMER, they managed to perform enzymatic digestion in only 5 min, the digestion temperature not being mentioned neither the sensitivity of the final method. The interest in amino acid adducts is recent. Different positions, other than Cys^{34}, can be alkylated such as aspartic acid, glutamic acid, histidine, methionine and lysine and lead to HETE adducts. Amino acid adducts on methionine residue were obtained after digestion using pepsin (24) where the other mentioned ones were obtained after pronase digestion (46). Glutamic acid adducts were also obtained after pepsin digestion but it was included in an hexapeptide (6). Amino acid adducts offer inter species analyses since the CPF sequence not observed in every species(Electronic Supplementary Material Table S12) (46). Digesting albumin up to the individual amino acid therefore allows inter species analysis as it is no longer specific to a given amino acid sequence. The overall LODs obtained can be equivalent to the ones reached when targeting Cys^{34} peptides as the signal can stack up if same amino acids get alkylated (46).

One notable phenomenon was witnessed with methionine adducts. Indeed, SM showed to relocate on other reactive sites of the protein (24) leading to methionine adducts with shorter lifetimes compared to Cys^{34} ones. They still can be used as short-term indicators of SM exposure and were monitored up to 5 days after the exposure.

Recovery of adducted peptides

Once digestion of the albumin has been performed, it is necessary to recover the alkylated peptides for their further LC-MS/MS determination while removing the enzyme from the analytical extract as well as any non-digested albumin. The different approaches are summarized in Fig 2. The removal of enzyme was mainly ensured by ultrafiltration (24,30,32–34,36,37,40,41) using a 10 kDa cut-off filter. During this procedure, the enzyme, either pronase, proteinase K or pepsin is retained on the filter which allows the alkylated peptides to pass through. Protein precipitation was also used to get rid of the enzyme (42–46). If acetone was applied to raw plasma samples, this precipitation carried out before LC-MS/MS analysis was achieved using acetonitrile (ACN), certainly selected for the peptide solubility and for its compatibility with LC-MS conditions as acetonitrile composed the LC mobile phase.

A supplementary concentration step of the adducted peptides using Solid Phase Extraction (SPE) was proposed to concentrate the filtrate after the ultrafiltration step (32,34,36) or directly on the digest of albumin (38,39) thus allowing both the concentration of the alkylated peptides and the removal of salt used for the digestion step. SPE was performed on hydrophobic sorbent, i.e. on C_{18} solid phases (32,34,36) or on Oasis® HLB (35,38,39), with the

use of ACN as eluting solvent. The addition of formic acid (FA) in ACN, that is also an additive of the mobile phase for LC-MS analysis, can also be noticed as it appeared in one protocol on C_{18} (36). The use of trifluoroacetic acid as an additive for the elution of the peptides was also witnessed (32,34). Interestingly, no additives were used when performing the SPE on Oasis® HLB phases. In one study, a recovery yield of 100% was reported for the ultrafiltration step and of 87% for the SPE step on C_{18} silica (32). However, no mention of assaying repeatability of the extraction procedure were mentioned in this work. A recovery of 86.3 ± 5.8% was also reported for a SPE step on Oasis® HLB sorbent (35) but without detail regarding number of assays.

LC-MS analysis of adducted peptides

All the works reported in Table 1 testify to the use of reverse phase liquid chromatography coupled to mass spectrometry for the determination of the alkylated peptides in the resulting extracts. Even though they possess hydrophilic properties, alkylated peptides were retained using C_{18} stationary phase using column of different dimensions ranging from 350 mm to 50 mm in length and ranging from 2.1 mm to 0.32 mm in internal diameter. Injected volumes vary greatly from 2.5 µL (35) up to 50 µL (30,33,36,39). Ratios of injected volumes on total column volumes up to 142 % (32) were witnessed meaning that the injection could disequilibrate the column and affect peak shape particularly when injecting complex samples. On top of that, no correlation was observed between the injected volume and the dimensions of the columns. All analyses were performed in ACN with FA as additive in variable quantities from 0.05% to 1% in total mobile phase. Only in one paper, the use of ammonium formate at 5 mmol/L added to FA was described (35). Analyses were mostly performed in gradient mode opposed to isocratic mode (36). Gradient duration varies from 3 min (38) up to around an hour (33,41,47). Gradients were often composed of an elution slope (values ranging around 0.8 (40) to 108 %ACN/min (42) with most values around 2 %ACN/min (30,32–35,40,41,44,45,47)) followed by an intense slope to elute any persisting compounds. Flow rate also varies from 30 µL/min to 600 µL/min, however, it doesn't appear to be related to the column dimensions. The column temperature for the analysis varies greatly, from 30°C (34,36,42,46) up to 70°C (35,38,39). The ionization of the alkylated peptides was achieved using electrospray source in positive mode.

The identification criteria mentioned in the OPCW BioPT04 procedure recommends at least two m/z transitions to confirm exposure to chemical agents. In the oldest papers this request is not fulfilled as the requirement for 2 transitions was not mandatory at this time. However, when looking at the mass spectrum of the adducts, there are enough transitions to satisfy the request (34). One transition could be used for quantification and at least one other must be used to confirm exposure. The choice for the quantification transition was made by using the most intense product ion to lower the limit of quantification of each analytical method. However, under certain conditions the OPCW BioPT04 procedure also accepts single but will result in fewer identification parameters for the chemical.

The m/z ratios for the precursor ions of HETE-CPF and HETE-CP are m/z 470 and m/z 323 respectively ($[M+H]^+$ ions). The most commonly used product ion is m/z 105 which corresponds to the HETE-moiety ($^+CH_2$-CH_2-S-CH_2-CH_2-OH) (33). This product ion is common to both the dipeptide and the tripeptide as it comes from the alkylated agent. Interestingly, the transition from the precursor ion to the product ion m/z 105 is exclusively used for the quantification because it is the transition leading to the strongest signal. It is interesting to mention that this product ion is also found when monitoring the fragments of adducts of SM-analogues (55,56).

Very recently, two groups proposed the derivatization of the alkylated peptides before their LC-MS/MS analysis. The derivatization can have different aims, such as improving the sensitivity of the detection or the volatility for further gas chromatography analysis. In the case of alkylated peptides of albumin, it was used to improve their retention in reverse phase LC. Indeed, the adducted peptides show low retention times during the chromatographic separation (42–44,46). Many derivatizing agents modify the N-terminal site of the ^{34}Cys thus reducing the hydrophilic properties of the analytes and increasing their retention on C_{18} silica stationary phase. Indeed, derivatized alkylated peptide elutes 5 minutes later than its underivatized analogues on a 15 min run (43). This increase in retention time could help reducing matrix effects. Different derivatization agents were evaluated all aiming to reduce hydrophilic properties of the alkylated peptides. Propionic anhydride seems the most suitable for the derivatization as yields of 100% were measured. When using benzylchloroformate or 1-nicotinoyloxy-succinimide only 46% and 81% yields were reached respectively. The use of these different derivatization agents also modifies the molecular mass of the analytes and therefore the m/z ratios of the precursor ions. When looking at the transition giving the most intense signal, fragment m/z 105 (product ion) was used (42,46) and also product

ion m/z 116 (proline, $C_5H_{10}NO_2$) when its intensity exceeded the m/z 105 one (43,44). Although, an important increase in the signal was witnessed by some authors when derivatizing (42), the advantage of using this overall analysis time extending step on the limits of quantification is still not clear as values are equivalent to non-derivatizing procedures.

Matrix effect performance of the methods

When working with biological samples such as whole blood and plasma, it is known that matrix constituents can interfere with the detection of the analytes of interest. To reduce matrix effect, as previously mentioned, numerous purification steps were proposed to isolate first adducted albumin and then alkylated peptides from other matrix components. The derivatization of peptides was also proposed to improve their retention which also allows to distance the elution of the peptides of interest from the polar interferents (buffer, salts) present in the injected extracts. These matrix effects, calculated by comparing the intensity of signal between an aqueous solution and the biological sample extract both spiked at the same concentration of the analyte of interest showed values of 93 ± 7 % (39). This means that low matrix effects were observed when combining a PP with acetone to recover albumin with an SPE purification of the alkylated peptides on an Oasis® HLB sorbent resulting from a digestion with proteinase K. This value was obtained without derivatization of the alkylated peptides.

In most of the studies, the use of internal standards (IS) was proposed to help quantification in complex matrices. Several molecules having similar physico-chemical properties than the alkylated peptides were used, obtained from deuterated SM (30,33,35,36,38,40,45) or by the use of 2-chloroethylethylsulfide which are similar alkylating agents both adducted to albumin (37,42,46). In both previous cases, the IS is added at the very beginning of the procedure, right after the eventual removal of the RBCs when working on whole blood. The IS undergoes the whole sample preparation steps as well as digestion to give deuterated alkylated peptidic IS. The use of deuterated atropine was also reported (24,41,43,44). It is added after the sample preparation procedure, right before the LC-MS/MS analysis. It must be noted that this type of internal standard was only used by one team and the use of deuterated atropine instead of regular atropine is not explained. The closer in physico-chemical properties the IS is, the more similar the matrix effect undergone by the molecules will be (57). This will help building the calibration curves as well as the quantification of the adducts. The use of deuterated analogues remains the best option as molecules are exactly the same only differing in their m/z ratios by a known value but is also the most expensive technique. In this case, d_8-SM (30,33,36,38,40) was more used than $^{13}C_9$-SM (35) as the latter one was only used once and is not well documented. However, the use of IS was not systematic (34,39,47).

Limits of detection reported for the different methods vary from 0.5 to 12.7 ng/mL of SM in plasma or blood. Those values result, as previously mentioned, from sample preparation procedures more or less sophisticated, from analytical column whose diameter may affect the sensitivity of the method and from the use of mass spectrometer made of different types of analyzer, mostly triple quadrupole (MRM mode) but also hydride analyzer (pseudo MRM mode) such as Q-time of flight or Q-ion trap up to Q-orbitrap from different generations and which we know that the intrinsic sensitivity can vary considerably. Nevertheless, it can be noticed that, although these values are very close, the sample volume of treated blood and plasma strongly varies, from 0.5 mL to 5 mL for blood, *i.e.* a factor 10, and from 10 µL to 200 µL for plasma, *i.e.* a factor 20. This aspect must be integrated when comparing the methods. In the most recent papers, LODs reach the ng/mL range while reducing considerably the sample volumes required for the analysis.

Adducted peptide's stability

Adducted peptide stability throughout the procedure is of great concern in order to get a robust information. With that in mind, the stability of both HETE-CP (40) and derivatized HETE-CP (43,44) was tested for 24h in an autosampler at 15°C. The results showed that adducted peptides were stable over this period in the quality control samples *i.e.* referenced plasma spiked with SM. This experiment indicates that adducted peptides do not degrade during a chromatographic sequence. However, stability during storage is also of importance, so different storage temperatures were tested as well as freeze/thaw cycle stability. Storage at -70°C, -20°C, 4°C, room temperature and 37°C (37,38) was investigated for 4 weeks with weekly analyses. Tests were conducted on spiked unexposed plasma and stability appeared satisfying with variations under 10% observed for HETE-CPF adducts (37). Freeze/thaw cycle stability was performed on HETE-CPF which was stable for 2 cycles from -70°C to room

temperature (38), and derivatized adducted dipeptide in spiked plasma was stable for up to 10 cycles between -20°C and 4°C (42). Data seem to show that samples have good stability at room temperature which is required for an analysis in a chromatographic sequence as well as good storage stability. Adducted amino acids in plasma were also tested and showed high storage stability at -20°C and appeared stable after 10 freeze/thaw cycles (-20°C to room temperature) (46). However, these experiments were carried out under specific conditions and stability tests should be carried out when working with adducted peptides, as stability may vary depending on the type of solvents used.

Towards the development of similar analytical methods for nitrogen mustards and SM analogues

The developed sample preparation procedures described earlier could also be used for the study of other mustards such as nitrogen mustards (HN-1, HN-2, HN-3) (34,35) and SM-analogues (methyl/ethyl/propyl/butyl/pentyl) (55,56) which have however been less studied. These compounds are registered in the Schedule 1 of the Annex of Chemicals of the Chemical Weapons Convention. The structures of the nitrogen mustards and SM-analogues are shown in Fig. 4.

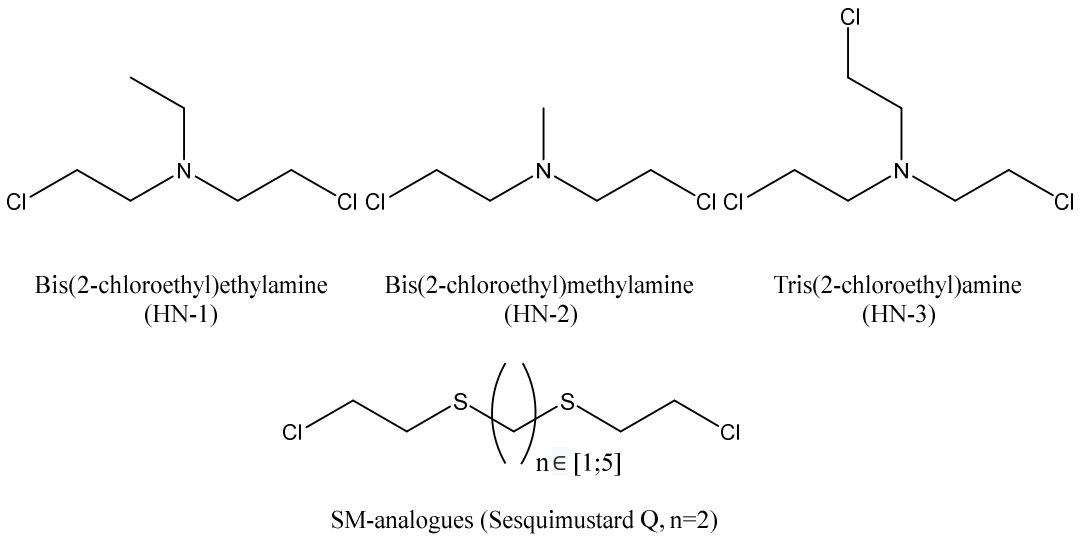


Fig. 4 *Structure of the nitrogen mustards and SM-analogues*

Concerning nitrogen mustards, they can be used as chemotherapeutic drugs such as HN-2 also known as mechlorethamine (58). Their carcinogenic and alkylating properties are sought for therapeutic uses, which leads to little research concerning their use as chemical warfare agents, although their reactivity is well established (59). Sesquimustard (Q) is part of the SM-analogues family meaning it differs from SM only by the presence of an extra sulfur atom and the length of the aliphatic chain between the two sulfur atoms. They are formed during the synthesis of the SM thus providing information about it. They could even be used as fingerprints of different batches and help determine the origin of SM (55).

As they are alkylating agents they can interact with the same sites as SM. In the case of Q, it showed higher reactivity than SM towards Cys^{34} (factor 1.33) supposedly due to a higher stability with the surrounding amino acids around reactive site. (56). This shows the great importance of the study of these compounds.

HN-1, HN-2 and HN-3 were detected simultaneously with SM-tripeptide adducts (34,35) in a same treated plasma initially spiked with SM as well as the three HN. Biological samples were even exposed to both SM and SM-analogues (methyl, Q, propyl, butyl, pentyl analogues) at the same time enabling detection of both CWs (55,56). The sample preparation procedures developed are summarized in Table 2.

Albumin purification steps are equivalent to those of SM-adducts as the target is the same, differing only by the alkylated molecules. Thus, affinity chromatography (34) as well as protein precipitation using acetone (55,56) or ethanol (60) were used. Simple dilution of the plasma was also performed prior to the enzymatic digestion (35). Albumin digestion was also performed with pronase or proteinase K under equivalent conditions as the ones

described for SM exposure. Pronase digestion was mostly done at 37°C (34,55,60) and proteinase K digestion was done at 50°C (35,56). Pronase was used to either produce alkylated tripeptides of nitrogen mustards (34,60), alkylated dipeptides of sesquimustard (56) or alkylated histidine of SM-analogues (55). Proteinase K on the other hand was exclusively used to produce alkylated tripeptides of nitrogen mustards (35) or SM-analogues (55,56). Here, the same behavior as previously described is visible, pronase leads to tripeptides in the oldest works where in the most recent ones, it leads to dipeptides. In the case of histidine adducts, digestion times were greatly increased (19h) to reach a more thorough digestion of albumin up to the amino acid using pronase (55).

Derivatization, already little used for the analysis of SM-adducts, has not been reported for the analysis of nitrogen mustards and analogous mustards. LC-MS/MS analysis conditions are also equivalent to the ones developed for the analysis of SM adducts. Indeed, RPLC using C_{18} as the stationary phase associated to water/ACN gradients containing formic acid as additive was exclusively used. Great variations regarding the injected volumes the same way as for the SM-adducts were noticed. Finally, the detection in multiple reaction monitoring or pseudo multiple reaction monitoring modes were used depending of the type of MS analyzers. CPF adducts of sesquimustard (HETETE-CPF) showed better detectability than histidine adducts (55) with LODs in the range of ng/mL, as did CPF adducts from the nitrogen mustards. HETETE-CP showed an LOI slightly higher than HETETE-CPF but still in the same range (56).

As described here, the study of nitrogen mustard and SM analogues is strikingly similar to the analysis of SM adducts meaning that methods can easily be applied to them extending the potential of chemical warfare exposure detection and monitoring with the same sample preparation methodology.

Table 2 *Sample preparation methods and LC-MS/MS methods developed for the analysis of adducts with different mustards in biological matrices.*

Chemical weapon	**Albumin adduct**	**Sample**	**Albumin isolation**	**Enzymatic digestion**	**Adduct isolation**	**Internal standard**	**LC conditions**	**MS analyzer**	**LOD of CW in real samples**	**Ref**
HN-1, HN-2, HN-3	HEEAE-CPF, HEMAE-CPF, DHEAE-CPF	Blood (human) 2 mL	C AC	pronase, 37°C, 2h	UF SPE (C_{18})	-	C_{18} (150 x 1 mm, 5 µm), 30°C V_{inj}= 20 µL	TQ	31.8 ng.mL-1 (HN-1), 16.0 ng.mL-1 (HN-2), 31.8 ng.mL-1 (HN-3)	(34)
HN-1, HN-2, HN-3	HEEAE-CPF, HEMAE-CPF, DHEAE-CPF	Plasma (human) 10 µL	Dilution 1/10 (50 mM NH_4HCO_3)	proteinase K, 50°C, 1h	SPE (Oasis® HLB) (R=94.5% (HN-1), 85.8% (HN-2), 84.0 % (HN-3))	^{15}N -HN peptides	C_{18} (50 x 2.1 mm, 1.3 µm), 70°C V_{inj}= 2,5µL	TQ	2.20 ng.mL-1 (HN-1), 1.76 ng.mL-1 (HN-2), 1.32 ng.mL-1 (HN-3) (HN-CPF spiked in matrix)	(35)
HN-2	HEMAE-CPF	Blood (human) 5 mL	C PP (sodium acetate/EtOH (5/95, v/v))	pronase, 37°C, 2h	UF SPE (C18)	-	C_{18} (150 x 0.3 mm), - V_{inj}= 50 µL	QToF	-	(60)
Sesquimustard (Q)	HETETE-CP, HETETE-CPF	Plasma (human) 50 µL	PP (Acetone)	pronase / proteinase K, 50°C, 2h	PP (ACN)	d_3-Atr	C_{18} (50 x 1 mm, id 1.8 µm, 100 Å), 60°C V_{inj}= 20 µL	QTRAP	LOI : HETETE-CP : 5.09 ng.mL-1 HETETE-CPF : 1.02 ng.mL-1	(56)
Methyl/ethyl (Q)/propyl/ butyl/pentyl -SM analogues	CPF-adducts N^1/N^3-His-adducts	Plasma (human) 2 mL	PP (Acetone)	proteinase K, 50°C, 90 min pronase, 37°C, 19h	SPE (Oasis ® HLB) (proteinase K) UF (pronase)	-	C_{18} (100 x 2.1 mm, 1.8 µm), ambient temperature V_{inj}= 5 µL (CPF-adducts), 1 µL (His-adducts)	TQ	-	(55)

^{15}N-NH: nitrogen mustard containing a nitrogen isotope; AC: Affinity Chromatography; C: Centrifugation; CW: Chemical weapon; d_3-Atr: Deuterated atropine; HETETE-CP: Hydroxyethylthioethylthioethyl-Cysteine-Proline; HETETE-CPF: Hydroxyethylthioethylthioethyl-Cysteine-Proline-Phenylalanine; His: Histidine; HN-1: Bis(2-chloroethyl)ethylamine; HN-2: Bis(2-chloroethyl)methylamine; HN-3: Tris(2-chloroethyl)amine; HN-1-CPF (HEEAE-CPF): Hydroxyethylethylaminoethyl-Cysteine-Proline-Phenylalanine; HN-2-CPF (HEMAE-CPF): Hydroxyethylmethylaminoethyl-Cysteine-Proline-Phenylalanine; HN-3-CPF (DHEAE-CPF): Dihydroxyethylaminoethyl-Cysteine-Proline-Phenylalanine; LC: Liquid chromatography; LOD/I: Limit of detection/identification; MS: Mass spectrometry; PP: Protein Precipitation; Q: Sesquimustard; Q ToF:

Hybrid Quadrupole-Time of Flight; QTRAP: Hybrid Quadrupole-Ion trap; R: Recovery yield; SPE: Solid phase extraction; TQ: Triple quadrupole; UF: UltraFiltration; Vinj: Injected volume.

The mobile phase used for reverse phase liquid chromatography was made of an ACN gradient with formic acid as an additive. All the LC-MS/MS methods used ElectroSpray Ionization in positive mode coupled with TQ analyzer used in Multiple Reaction Monitoring (MRM) mode or Q-ToF and Orbitrap analyzers used in pseudo-MRM mode. The LOD/LOI values correspond to the concentration of exposure of whole blood or plasma to the CW. The S/N ratio of LOI was not given in the articles.

Conclusion

The review is an exhaustive list of the main analytical procedures developed in the last twenty years for the analysis of albumin adducts in biological samples in order to confirm an exposure to sulfur mustard. Albumin adducts are interesting biomarkers of exposure as they offer a wider period of time for sampling than the free metabolites of sulfur mustard (TDG, TDGO, SMO, GSH conjugation products) and unequivocal proof of exposure.

In the last few years, sample preparation was greatly improved and several analytical methods can detect biomarkers at concentration around to 1 $ng.mL^{-1}$. The improvement of the methods sensitivity makes it possible to confirm an exposure to sulfur mustard at very low concentrations and up to several weeks after the exposure. Indeed, albumin biomarkers were quantified in real case human samples 3 to 4 weeks post exposure(61). An important improvement lies on the reduction of the sample volume required for the analysis as factors of 10 for blood and 20 for plasma were noticed.

As illustrated by this review, many steps are required prior to the analysis of the peptide biomarker of exposure of albumin to SM. While the quantification of alkylated peptides relies mainly on the use of reverse phase LC-MS/MS, with or without a derivatization step, up to four methods can be combined, to isolate albumin from plasma or processed blood sample, to which two additional methods can be added to recover the peptides after protein digestion that can be performed with different proteases. It is undeniable that the multiplicity of steps is a source of loss of analytes and can lead to an increase in the uncertainty of the results provided. Rather than simply combining multiple steps to the procedure, the main focus should be to favor the more selective approaches. Despite this, little data is available to evaluate the performance of the different stages. Only data relating to the limits of detection reached are provided. These data are not sufficient to compare the potential of the reported methods as they strongly depend on the sensitivity of the multiple mass analyzers used. It would have been necessary to have quantitative data such as recovery rates, enrichment factors and digestion rates to fully compare the different procedures. Despite the fact that quantitative methods are not required by the OPCW BioTP04, quantitative data are necessary to compare procedures between one another, in order to determine the most reliable.

Nevertheless, this review suggests new possibilities for improvement. A more systematic use of IMER to be coupled on-line with LC-MS/MS as it was done for the quantitative analysis of peptide adducts of butyrylcholinesterase submitted to organophosphorous compounds or hemoglobin submitted to SM could be considered (50). The use of HILIC mode in LC to favor the retention of polar peptide without requiring a derivatization step could also be envisaged. Nevertheless, all new sample preparation developments should be carried out by focusing on the evaluation of the performance of each stage in order to establish their contribution to the global method allowing easier intercomparisons of the different approaches. It would also be interesting to study the impact of procedures on the stability of adducted peptides particularly as a function of different solvents, the use of rapid methods with reduced number of steps would limit the potential risk of degradation.

This review includes also a list of analytical procedures for sulfur mustard analogues and nitrogen mustards. Indeed, the monitoring of other mustard adducts is also of importance as similar sample preparation procedures can be developed for an unambiguous proof of exposure but also to obtain information about the synthesis pathway of mustard agent as some of them could be impurities created during synthesis of blister agents.

Declaration

Acknowledgment

The authors would like to thank the French Defense Procurement Agency for funding this research study.

Conflict of Interest

The authors declare no competing interests.

Credit author statement

Lorenzo Avigo, Florine Hallez: Writing the original draft. Charlotte Desoubries, Anne Bossée, Christine Albaret, Audrey Combès, Valérie Pichon: Visualization, Methodology, Writing – review & editing. Charlotte Desoubries, Valerie Pichon,: Project administration, Funding acquisition, Supervision.

Funding

This work was supported by DGA, CBRN Defense [Grant number 2022470173]

References

1. Fitzgerald GJ. Chemical Warfare and Medical Response During World War I. Am J Public Health. 2008;98(4):611-25.

2. Ruhl CM, Park SJ, Danisa O, Morgan RF, Papirmsister B, Sidell FR, et al. A serious skin sulfur mustard burn from an artillery shell. J Emerg Med. 1994;12(2):159-66.

3. Barr JR, Pierce CL, Smith JR, Capacio BR, Woolfitt AR, Solano MI, et al. Analysis of Urinary Metabolites of Sulfur Mustard in Two Individuals after Accidental Exposure. J Anal Toxicol. 2008;32(1):10-6.

4. Smith JR, Capacio BR, Korte WD, Woolfitt AR, Barr JR. Analysis for Plasma Protein Biomarkers Following an Accidental Human Exposure to Sulfur Mustard. J Anal Toxicol. 2008;32(1):17-24.

5. Quillen C. The Islamic State's Evolving Chemical Arsenal. Stud Confl Terror. 2016;39(11):1019-30.

6. John H, Koller M, Worek F, Thiermann H, Siegert M. Forensic evidence of sulfur mustard exposure in real cases of human poisoning by detection of diverse albumin-derived protein adducts. Arch Toxicol. 2019;93(7):1881-91.

7. Sezigen S, Ivelik K, Ortatatli M, Almacioglu M, Demirkasimoglu M, Eyison RK, et al. Victims of chemical terrorism, a family of four who were exposed to sulfur mustard. Toxicol Lett. 2019;303:9-15.

8. Wattana M, Bey T. Mustard Gas or Sulfur Mustard: An Old Chemical Agent as a New Terrorist Threat. Prehospital Disaster Med. 2009;24(01):19-29.

9. Sezigen S, Eyison RK, Kenar L. Bioanalytical verification of sulfur mustard exposure in a syrian family. GULHANE Med J. 2019;61(2):64.

10. Sezigen S, Eyison RK, Kilic E, Kenar L. Evidence of sulfur mustard exposure in victims of chemical terrorism by detection of urinary β-lyase metabolites. Clin Toxicol. 2020;58(1):36-44.

11. Noort D, Verheij ER, Hulst AG, De Jong LPA, Benschop HP. Characterization of Sulfur Mustard Induced Structural Modifications in Human Hemoglobin by Liquid Chromatography−Tandem Mass Spectrometry. Chem Res Toxicol. 1996;9(4):781-7.

12. Riches J, Read RW, Black RM. Analysis of the sulphur mustard metabolites thiodiglycol and thiodiglycol sulphoxide in urine using isotope-dilution gas chromatography–ion trap tandem mass spectrometry. J Chromatogr B. 2007;845(1):114-20.

13. Li C, Chen J, Liu Q, Xie J, Li H. Simultaneous quantification of seven plasma metabolites of sulfur mustard by ultra high performance liquid chromatography–tandem mass spectrometry. J Chromatogr B. 2013;917-918:100-7.

14. Zhang Y, Yue L, Nie Z, Chen J, Guo L, Wu B, et al. Simultaneous determination of four sulfur mustard–DNA adducts in rabbit urine after dermal exposure by isotope-dilution liquid chromatography–tandem mass spectrometry. J Chromatogr B. 2014;961:29-35.

15. Noort D, Fidder A, Benschop HP, De Jong LPA, Smith JR. Procedure for Monitoring Exposure to Sulfur Mustard Based on Modified Edman Degradation of Globin. J Anal Toxicol. 2004;28(5):311-5.

16. Xu H, Nie Z, Zhang Y, Li C, Yue L, Yang W, et al. Four sulfur mustard exposure cases: Overall analysis of four types of biomarkers in clinical samples provides positive implication for early diagnosis and treatment monitoring. Toxicol Rep. 2014;1:533-43.

17. Capacio BR, Smith JR, DeLion MT, Anderson DR, Graham JS, Platoff GE, et al. Monitoring Sulfur Mustard Exposure by Gas Chromatography-Mass Spectrometry Analysis of Thiodiglycol Cleaved from Blood Proteins. J Anal Toxicol. 2004;28(5):306-10.

18. Black RM, Brewster K, Clarke RJ, Hambrook JL, Harrison JM, Howells DJ. Biological fate of sulphur mustard, 1,1'-thiobis(2-chloroethane): isolation and identification of urinary metabolites following intraperitoneal administration to rat. Xenobiotica. 1992;22(4):405-18.

19. Black RM, Read RW. Biological fate of sulphur mustard, 1,1'-thiobis(2-chloroethane): identification of β-lyase metabolites and hydrolysis products in human urine. Xenobiotica. 1995;25(2):167-73.

20. von Stedingk H, Osterman-Golkar S, Törnqvist M. Chapter 4. Biomarkers of Exposure: Hemoglobin Adducts. In: Knudsen L, Merlo DF, éditeurs. Issues in Toxicology [Internet]. Cambridge: Royal Society of Chemistry; 2011. p. 1-22.

21. van der Schans GP, Noort D, Mars-Groenendijk RH, Fidder A, Chau LF, de Jong LPA, et al. Immunochemical Detection of Sulfur Mustard Adducts with Keratins in the Stratum Corneum of Human Skin. Chem Res Toxicol. 2002;15(1):21-5.

22. Boyer AE, Ash D, Barr DB, Young CL, Driskell WJ, Whitehead RD, et al. Quantitation of the Sulfur Mustard Metabolites 1,1'-Sulfonylbis[2-(methylthio)ethane] and Thiodiglycol in Urine Using Isotope-Dilution Gas Chromatography-Tandem Mass Spectrometry. J Anal Toxicol. 2004;28(5):327-32.

23. Noort D, Black RM. Methods for Retrospective Detection of Exposure to Toxic Scheduled Chemicals. Part B: Mass Spectrometric and Immunochemical Analysis of Covalent Adducts to Proteins and DNA. In: Mesilaakso M, éditeur. Chemical Weapons Convention Chemicals Analysis [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2006. p. 433-51.

24. Siegert M, Gandor F, Kranawetvogl A, Börner H, Thiermann H, John H. Methionine ^{329} in human serum albumin: A novel target for alkylation by sulfur mustard. Drug Test Anal. 2019;11(5):659-68.

25. Orlova OI, Karakashev GV, Savel'eva EI. Simultaneous Determination of Sulfur Mustard Adducts with Guanine and Acetylcysteine in Urine by High-Resolution High-Performance Liquid Chromatography–Tandem Mass Spectrometry. J Anal Chem. 2020;75(8):1011-7.

26. Törnqvist M, Fred C, Haglund J, Helleberg H, Paulsson B, Rydberg P. Protein adducts: quantitative and qualitative aspects of their formation, analysis and applications. J Chromatogr B. 2002;778(1-2):279-308.

27. Benschop HP, Van Der Schans GP. Verification, Dosimetry, and Biomonitoring of Mustard Gas Exposure via Immunochemical Detection of Mustard Gas Adducts to DNA and Proteins. Prins Maurits Laboratorium Tno Rijswijk (Netherlands); 1993.

28. Meloun B, Morávek L, Kostka V. Complete amino acid sequence of human serum albumin. FEBS Lett. 1975;58(1-2):134-7.

29. Noort D, Fidder A, Hulst AG, de Jong LPA, Benschop HP. Diagnosis and dosimetry of exposure to sulfur mustard: development of a standard operating procedure for mass spectrometric analysis of haemoglobin adducts: exploratory research on albumin and keratin adducts. J Appl Toxicol. 2001;20(S1):S187-92.

30. Noort D, Fidder A, Hulst AG, Woolfitt AR, Ash D, Barr JR. Retrospective Detection of Exposure to Sulfur Mustard: Improvements on an Assay for Liquid Chromatography-Tandem Mass Spectrometry Analysis of Albumin/Sulfur Mustard Adducts. J Anal Toxicol. 2004;28(5):333-8.

31. Bechtold WE, Willis JK, Sun JD, Griffith WC, Reddy TV. Biological markers of exposure to benzene: 5-phenylcysteine in albumin. Carcinogenesis. 1992;13(7):1217-20.

32. Noort D, Hulst AG, de Jong LP, Benschop HP. 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. 1999;12(8):715-21.

33. Noort D, Fidder A, Degenhardt-Langelaan CEAM, Hulst AG. Retrospective Detection of Sulfur Mustard Exposure by Mass Spectrometric Analysis of Adducts to Albumin and Hemoglobin: An In Vivo Study. J Anal Toxicol. 2008;32(1):25-30.

34. Yeo TH, Ho ML, Loke WK. Development of a Liquid Chromatography-Multiple Reaction Monitoring Procedure for Concurrent Verification of Exposure to Different Forms of Mustard Agents. J Anal Toxicol. 2008;32(1):51-6.

35. Pantazides BG, Quiñones-González J, Rivera Nazario DM, Crow BS, Perez JW, Blake TA, et al. A quantitative method to detect human exposure to sulfur and nitrogen mustards via protein adducts. J Chromatogr B. 2019;1121:9-17.

36. Andacht TM, Pantazides BG, Crow BS, Fidder A, Noort D, Thomas JD, et al. An Enhanced Throughput Method for Quantification of Sulfur Mustard Adducts to Human Serum Albumin Via Isotope Dilution Tandem Mass Spectrometry. J Anal Toxicol. 2014;38(1):8-15.

37. Liu C, Liang L, Xiang Y, Yu H, Zhou S, Xi H, et al. An improved method for retrospective quantification of sulfur mustard exposure by detection of its albumin adduct using ultra-high pressure liquid chromatography-tandem mass spectrometry. Anal Bioanal Chem. 2015;407(23):7037-46.

38. Pantazides BG, Crow BS, Garton JW, Quiñones-González JA, Blake TA, Thomas JD, et al. Simplified Method for Quantifying Sulfur Mustard Adducts to Blood Proteins by Ultrahigh Pressure Liquid Chromatography–Isotope Dilution Tandem Mass Spectrometry. Chem Res Toxicol. 2015;28(2):256-61.

39. Braun AV, Rybal'chenko IV, Ponsov MA, Stavitskaya YaV, Tikhomirov LA, Grechukhin AP. Optimization of a method for the determination of a mustard gas biomarker in human blood plasma by liquid chromatography–high-resolution mass spectrometry. J Anal Chem. 2017;72(3):303-8.

40. Gandor F, Gawlik M, Thiermann H, John H. Evidence of Sulfur Mustard Exposure in Human Plasma by LC-ESI-MS-MS Detection of the Albumin-Derived Alkylated HETE-CP Dipeptide and Chromatographic Investigation of Its Cis/Trans Isomerism. J Anal Toxicol. 2015;39(4):270-9.

41. John H, Siegert M, Gandor F, Gawlik M, Kranawetvogl A, Karaghiosoff K, et al. Optimized verification method for detection of an albumin-sulfur mustard adduct at Cys34 using a hybrid quadrupole time-of-flight tandem mass spectrometer after direct plasma proteolysis. Toxicol Lett. 2016;244:103-11.

42. Chen B, Yu HL, Liu SL, Liu CC, Liang LH, Li XH, et al. A sensitive quantification approach for detection of HETE-CP adduct after benzyl chloroformate derivatization using ultra-high-pressure liquid chromatography tandem mass spectrometry. Anal Bioanal Chem. 2019;411(15):3405-15.

43. John H, Richter A, Thiermann H. Evidence of sulfur mustard poisoning by detection of the albumin-derived dipeptide biomarker C(-HETE)P after nicotinylation. Drug Test Anal. 2021;13(9):1593-602.

44. Richter A, Siegert M, Thiermann H, John H. Alkylated albumin-derived dipeptide C(-HETE)P derivatized by propionic anhydride as a biomarker for the verification of poisoning with sulfur mustard. Anal Bioanal Chem. 2021;413(19):4907-16.

45. John H, Hörmann P, Schrader M, Thiermann H. Alkylated glutamic acid and histidine derived from protein-adducts indicate exposure to sulfur mustard in avian serum. Drug Test Anal. 2022;14(6):1140-8.

46. Chen B, Ren Z, Zhang T, Yu H, Shu Z, Liu C, et al. Simultaneous quantification of multiple amino acid adducts from sulfur mustard-modified human serum albumin in plasma at trace exposure levels by ultra-high performance liquid chromatography–triple quadrupole mass spectrometry after propionyl derivatization. J Chromatogr A. 2022;1678:463354.

47. Carol-Visser J, van der Schans M, Fidder A, Hulst AG, van Baar BLM, Irth H, et al. Development of an automated on-line pepsin digestion–liquid chromatography–tandem mass spectrometry configuration for the rapid analysis of protein adducts of chemical warfare agents. J Chromatogr B. 2008;870(1):91-7.

48. Young CL, Woolfitt AR, McWilliams LG, Moura H, Boyer AE, Barr JR. Survey of albumin purification methods for the analysis of albumin-organic toxicant adducts by liquid chromatography-electrospray ionization-tandem mass spectrometry. PROTEOMICS. 2005;5(18):4973-9.

49. Mahany T, Khirabadi BS, Gersten DM, Kurian P, Ledley RS, Ramwell PW. Studies on the affinity chromatography of serum albumins from human and animal plasmas. Comp Biochem Physiol Part B Comp Biochem. 1981;68(2):319-23.

50. Hallez F, Combès A, Desoubries C, Bossée A, Pichon V. Analysis of long-lived sulfur mustard-human hemoglobin adducts in blood samples by red blood cells lysis and on-line coupling of digestion on an immobilized-trypsin reactor with liquid chromatography-tandem mass spectrometry. J Chromatogr A. 2022;1665:462830.

51. Ebeling W, Hennrich N, Klockow M, Metz H, Orth HD, Lang H. Proteinase K from Tritirachium album Limber. Eur J Biochem. 1974;47(1):91-7.

52. Camperi J, Pichon V, Delaunay N. Separation methods hyphenated to mass spectrometry for the characterization of the protein glycosylation at the intact level. J Pharm Biomed Anal. 2020;178:112921.

53. Mowrer J, Törnqvist M, Jensen S, Ehrenberg L. Modified Edman degradation applied to hemoglobin for monitoring occupational exposure to alkylating agents. Toxicol Environ Chem. 1986;11(3):215-31.

54. Hallez F, Combès A, Desoubries C, Bossée A, Pichon V. Development of an immobilized-trypsin reactor coupled to liquid chromatography and tandem mass spectrometry for the analysis of human hemoglobin adducts with sulfur mustard. J Chromatogr B. 2021;1186:123031.

55. Hemme M, Fidder A, van der Riet-van Oeveren D, van der Schans MJ, Noort D. Mass spectrometric analysis of adducts of sulfur mustard analogues to human plasma proteins: approach towards chemical provenancing in biomedical samples. Anal Bioanal Chem. 2021;413(15):4023-36.

56. Blum MM, Richter A, Siegert M, Thiermann H, John H. Adduct of the blistering warfare agent sesquimustard with human serum albumin and its mass spectrometric identification for biomedical verification of exposure. Anal Bioanal Chem. 2020;412(28):7723-37.

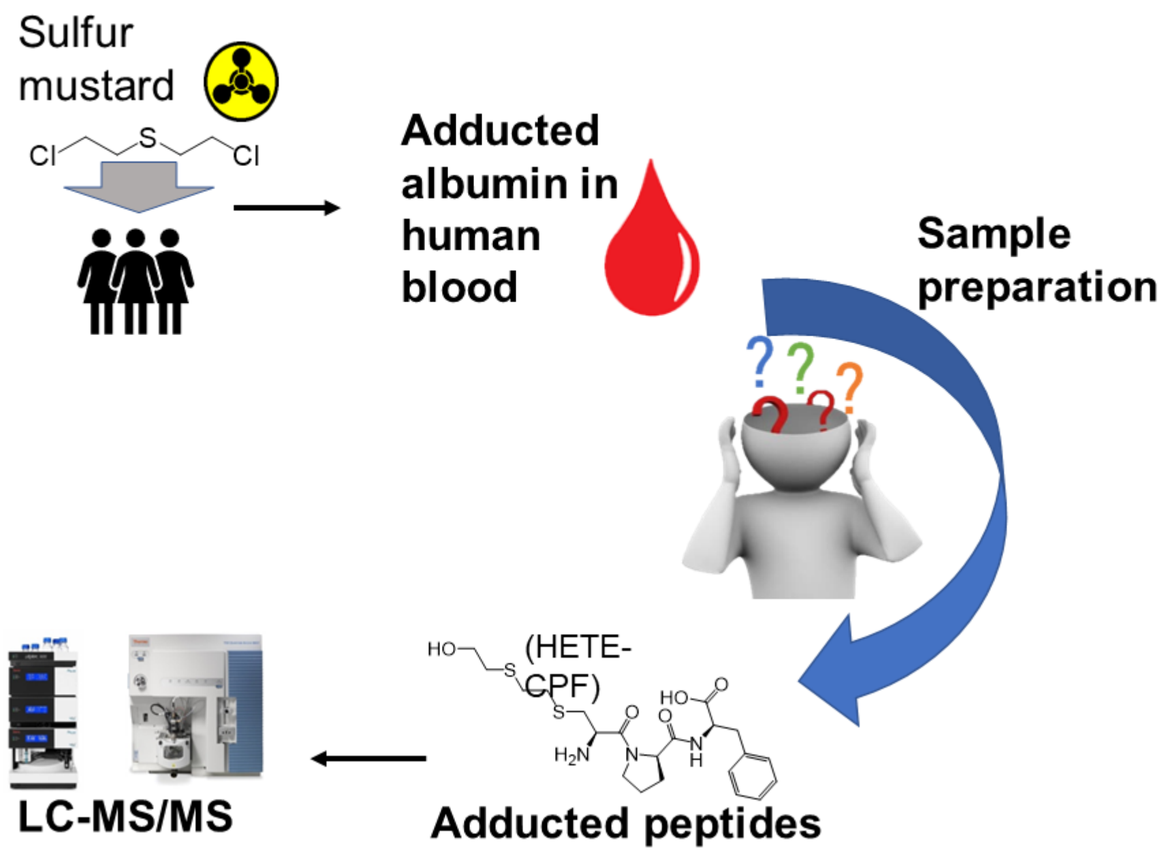
57. Piórkowska E, Musijowski J, Buś-Kwaśnik K, Rudzki PJ. Is a deuterated internal standard appropriate for the reliable determination of olmesartan in human plasma? J Chromatogr B. 2017;1040:53-9.

58. Povirk LF, Shuker DE. DNA damage and mutagenesis induced by nitrogen mustards. Mutat Res Genet Toxicol. 1994;318(3):205-26.

59. Wang QQ, Begum RA, Day VW, Bowman-James K. Sulfur, oxygen, and nitrogen mustards: stability and reactivity. Org Biomol Chem. 2012;10(44):8786.

60. Noort D, Hulst A, Jansen R. Covalent binding of nitrogen mustards to the cysteine-34 residue in human serum albumin. Arch Toxicol. 2002;76(2):83-8.

61. Steinritz D, Striepling E, Rudolf KD, Schröder-Kraft C, Püschel K, Hullard-Pulstinger A, et al. Medical documentation, bioanalytical evidence of an accidental human exposure to sulfur mustard and general therapy recommendations. Toxicol Lett. 2016;244:112-20.


Sulfur mustard
Cl
S
Cl
Adducted albumin in human blood
Sample preparation
HO
(HETE-CPF)
S
H_2N
N
O
HO
O
N
H
LC-MS/MS
Adducted peptides

Valérie Pichon is Full Professor at Sorbonne University and head of the Analytical Chemistry Department at ESPCI Paris (PSL University). Her research focuses on the development and miniaturization of antibody- or aptamer-based sorbents and imprinted polymers for trace analysis in complex samples, and their coupling with LC-MS systems. She has been awarded the Analytical Chemistry Department Prize of the French Society of Chemistry (2001), the Silver Medal of the French National Center for Scientific Research (2022) and recently the Research Prize of the French Association for Separation Sciences (2023).

Electronic Supplementary Material

Analytical methods based on liquid chromatography for the analysis of albumin adducts involved in retrospective biomonitoring of exposure to mustard agents ; Lorenzo Avigo, Florine Hallez, Audrey Combès, Charlotte Desoubries, Christine Albaret, Anne Bossée, Valérie Pichon*

Table S1 Amino acid sequences of serum albumin from human, chicken, duck and ostrich. Sequences were obtained from UniProtKB database entries: human, chicken, duck and ostrich. Amino acid numbering does not consider either the signal peptide alone (duck and ostrich) or the signal and propeptide (human and chicken). Adapted from (45).

Species	Entry (UniProt)	Sequence
Human (Homo sapiens)	P02768	AQYLQQC^{34}PFEDHV
Chicken (Gallus gallus)	P19121	AQYLQRC^{38}SYEGLS
Duck (Anas platyrhynchos, Mallard, Anas boschas)	R0M0W6	AQYLQRC^{43}SYDGLS
Ostrich (Struchio camelus australis)	A0A093A422	AQYLQRC^{43}SYEGLS

Table S2 Recovery of the adduct HETE-CPF from a test sample of human blood plasma in vitro spiked by 100 ng/mL of sulfur mustard with the use of different enzymes (n=3, P=0.95). Digestion conditions: 50°C, 90 min, pH 7.8. Data from (39).

Enzyme	Recovery of HETE-CPF, %
Proteinase K (QIAGEN)	62 ± 9
Proteinase K (Applichen Panreac)	42 ± 8
Proteinase K (SibEnzim)	30 ± 9
Pronase (Sigma Aldrich)	Not detected