



HAL
open science

Kinetic and structural characterization of therapeutic albumin chemical functionalization using complementary mass spectrometry techniques

Rabah Gahoual, Gérard Bolbach, Ikram Ould-Melha, Gilles Clodic, Yannis Francois, Daniel Scherman, Nathalie Mignet, Pascal Houzé

► To cite this version:

Rabah Gahoual, Gérard Bolbach, Ikram Ould-Melha, Gilles Clodic, Yannis Francois, et al.. Kinetic and structural characterization of therapeutic albumin chemical functionalization using complementary mass spectrometry techniques. *Journal of Pharmaceutical and Biomedical Analysis*, 2020, 185, pp.113242. 10.1016/j.jpba.2020.113242 . hal-03022806

HAL Id: hal-03022806

<https://hal.science/hal-03022806>

Submitted on 24 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Kinetic and structural characterization of therapeutic albumin chemical functionalization using complementary mass spectrometry techniques

Rabah Gahoual^{1*}, Gérard Bolbach², Ikram Ould-Melha¹, Gilles Clodic², Yannis-Nicolas François³, Daniel Scherman¹, Nathalie Mignet¹, Pascal Houzé^{1,4}

¹Université de Paris, Faculté de sciences pharmaceutiques et biologiques, Unité de Technologies Chimiques et Biologiques pour la Santé (UTCBS), CNRS UMR8258, Inserm U1022, Paris (France)

²Sorbonne Université, Institut de Biologie Paris Seine, Plateforme de spectrométrie de masse et protéomique, Paris (France)

³Laboratoire de Spectrométrie de Masse des Interactions et des Systèmes (LSMIS), CNRS – UMR7140, Université de Strasbourg, Strasbourg (France)

⁴Service de Biochimie générale, Hôpital universitaire Necker-Enfants Malades, Assistance Publique – Hôpitaux de Paris (AP-HP), Paris (France)

Corresponding author:

Dr. Rabah Gahoual

Unité de Technologies Chimiques et Biologiques pour la Santé (UTCBS)

Faculté de Pharmacie – Université de Paris

4, avenue de l'observatoire

75270 Paris Cedex 06

E-mail: rabah.gahoual@parisdescartes.fr

HIGHLIGHTS

- Chemical functionalization of HSA generating protein conjugates was investigated
- MALDI-TOF allowed estimating the number of species coupled to the peptide backbone
- MALDI-TOF provided the global reaction kinetic over the intact protein
- LC-MS/MS enabled to perform a precisely localized the modified lysine residues
- Data crossing established link between accessibility and reactivity of the residues

ABSTRACT: Protein conjugates such as antibody-drugs conjugates (ADCs) represents the next generation of therapeutic proteins. They allow to combine the biological properties of the protein format with the characteristics of the conjugated ligands. The reaction implemented to couple ligands to the peptide backbone represents a crucial aspect of the production of protein conjugates, influencing the nature and the heterogeneity of the conjugates obtained. Here, we report the concomitant use of MALDI-TOF MS and LC-MS/MS analysis to investigate the chemical functionalization of human serum albumin (HSA) by the intermediate of lysine residues, previously used to generate biopharmaceutical agents for medical imaging. A kinetic was performed by collecting samples after different reaction times and analyzing them using the two techniques. MALDI-TOF MS analyses allowed estimating the number of conjugated ligands in a robust manner and assess the global functionalization kinetic on the intact protein level. Results demonstrated a maximum of 38 modified residues out of the 59 lysines available showing the limitation of the chemical functionalization. Consequently, LC-MS/MS analysis provided a site-specific characterization of the residues undergoing chemical modification. Data exhibited unique properties due to the presence of the ligands which allowed to identify without ambiguity the residues exhibiting different modification rate and enabled the identification of the unmodified lysine. Results were compared to the structure of HSA described from crystallography data. The comparison strongly suggested that accessibility is influencing the residues respective reactivity. The relevant complementarity of the different techniques could be emphasized in order to perform an extensive characterization concerning the evolution of the primary structure of the protein during the chemical reaction, providing an improved insight on the conjugation process and offering the potentiality to tune the reaction.

INTRODUCTION

The usage of protein as therapeutic drugs has demonstrated a major success in a variety of therapeutic fields like oncology and immune disorder for instance [1, 2]. As an illustration, the number of proteins approved for therapeutic use is experiencing an exponential progress and the technologies involved for their production appear to be in constant evolution [3]. Thus, this success has allowed to change in a couple of decades the paradigm in the industry, leading to the emergence of the biopharmaceutical industry which currently represents a significant portion of pharmaceutical product development [4]. Biopharmaceutical products therefore include a wide variety of proteins like insulin, monoclonal antibodies (mAbs), interferon- γ , erythropoietin or albumin [5]. Regarding the implementation of therapeutic proteins, one of the most advanced application is represented by the chemical modification of proteins. Protein chemical modification has been successfully employed to produce antibody-drug conjugates (ADCs). ADCs represents an emerging type of therapeutic proteins based on mAbs which incorporate a highly potent cytotoxic drug covalently bonded to the peptide backbone through chemical linkage [6]. Currently, the US Food and Drug Administration (FDA) has approved a total of six ADCs which include trastuzumab emtansine (Kadcyla, Genentech/Roche), brentuximab vedotin (Adcentris, Seattle Genetics) and inotuzumab ozogamicin (Pfizer/Wyeth) and more than 50 ADCs are in clinical trials demonstrating the interest of the biopharmaceutical industry for this type of proteins [7]. Similarly, chemical functionalization of proteins could also be demonstrated using human serum albumin (HSA). For instance, HSA was chemically modified in order to incorporate simultaneously a significant number of carbohydrates, like galactose (GSA) or lactose (LSA), in addition to chelation moieties [8, 9]. In this case, the functionalization of HSA provides the ability to the protein to chelate technetium 99m (^{99m}Tc) in order to produce a radiopharmaceutical agent which can be detected using scintigraphy for medical imaging [10]. Concomitantly, the incorporation of carbohydrates offers the possibility to assess the quantity of functional hepatocytes which are targeted by the intermediate of the asialoglycoprotein receptor (ASGPR) [11]. The use of this type of bioimaging agent has demonstrated to be particularly relevant to assess the liver function in different contexts such as liver cirrhosis, hepatocarcinoma, liver failure detection in addition to provide a suitable tool to help decide upon performing hepatectomy procedure. Regarding the clinical practice, ^{99m}Tc -GSA is currently commercially available in Japan as a labelling kit for liver function assessment [12]. Therefore, the conjugation of ligands to therapeutic proteins can be envisaged to extend drastically the application range of biopharmaceutical products by providing the opportunity to enrich and/or tune the characteristics of the agent while benefiting from the specificities of the protein format like suitable serum half-life for example.

Concerning the production of conjugated protein products, the ligands can be coupled to the peptide backbone of the protein by the implementation of different chemical reactions mainly depending on the type of amino acid residues involved [13, 14]. Therefore, drug coupling by the intermediate of cysteine's thiol moiety has been described for the preparation of brentuximab vendotin [15]. Similarly, the preparation of protein conjugates can be achieved by the intermediate of the amino group located on lysine's lateral chain [16]. Commonly, conjugation reaction conducts to the formation of a mixture of variants exhibiting different numbers of ligands effectively coupled to the proteins peptide backbone [17]. For a given variant, conjugated ligands can be spatially distributed on different amino acids [18]. As a consequence, protein conjugates exhibit an extra layer of heterogeneity generated from different

number and localization of ligands coupled during production which is adding to the inherent structural complexity of the initial protein [19]. The heterogeneity of protein conjugates may be originating from differences in the kinetic of the conjugation reaction depending on the position of the amino acid residues on the peptide backbone. Therefore, it is of utmost importance to provide analytical methodologies which allows to perform a comprehensive structural characterization of protein conjugates, in conjunction with detailed kinetic information concerning the conjugation reaction. The possibility to correlate the structural characterization regarding conjugated ligands and kinetic information could be extremely valuable in order to investigate in a detailed manner the overall kinetic of the conjugation reaction.

An extended panel of analytical techniques is currently used in order to perform a complete characterization of biotherapeutic proteins over the different levels defining their structure [20]. Mass spectrometry (MS) has gradually taken a pivotal role in structural characterization of biotherapeutic proteins including protein conjugates [19]. This is explained by the outstanding specificity as well as sensitivity provided by MS instrumentation and the possibility using dedicated methodologies to obtain highly detailed structural information [21]. For example, MS could be implemented in order to identify the ADCs species exhibiting different payloads [22]. MS often needs to be used in conjunction with a separation technique like liquid chromatography [23] or capillary electrophoresis [24, 25]. Recently, we described a methodology for the production of lactosylated HSA conjugated with diethylenetriaminepentaacetic acid (DTPA) [10]. This protein conjugate could be successfully use to form chelate with Europium (Eu^{2+}) in order to be used with fluorescence detection. The implementation of the Eu^{2+} lanthanide provides an interesting platform for the early development of bioconjugate aimed to be used as radiopharmaceutical agents. The influence of the incorporation of lactose moieties *in vitro* demonstrated the possibility to enhance hepatocyte targeting by the conjugation of lactose ligands [26].

In this work, we developed an analytical methodology based on complementary MS experiments in order to characterize from a kinetic and structural aspects in concomitance, the reaction performed for ligand coupling on lysine residues of HSA. HSA is the most abundant protein in blood plasma which represents a concentration about 35-50 mg/mL for human serum. HSA sequence is composed of 585 amino acids which lead to an average molecular mass of 66472 Da [27]. The structure of the protein is divided in three homologous domains (domain I to III), subdivided in subdomains A-B [28]. HSA was first approved for therapeutic use in 2005 by the US FDA. Concerning large scale production, it can be isolated from collected serums or produced using recombinant technologies [29]. HSA is principally indicated in the treatment of hypovolemia originating from cirrhotic ascites and hypoalbuminemia which may resulted from important burning or acute nephrosis for instance [28]. HSA incubated with 2-iminothiolane (2-IT) was sampled at different time throughout the chemical functionalization reaction. The samples were analyzed using MALDI-TOF MS analysis in order to determine the number of ligands coupled to the peptide backbone of the protein depending the incubation time used to perform the reaction. In parallel, the samples were submitted to proteolytic digestion followed by LC-MS/MS analysis in order to characterize the lysine residues undergoing chemical modification. Finally, the data obtained using LC-MS/MS analysis were compared to the tertiary structure of HSA obtained from X-ray crystallography in order to further understand the parameters influencing the occurrence of the chemical modification and give an improved insights regarding the formation of the protein conjugate.

EXPERIMENTAL SECTION

Chemicals used were systematically of analytical grade or high purity grade. Ultra-pure water used to prepare buffers and sample solutions was obtained using an Milli-Q reference A+ water purification system purchased from Merck Millipore (Billerica, MA). LC-MS grade H₂O and acetonitrile (ACN) used for UPLC-MS/MS experiments were purchased from VWR chemicals (Fontenay-sous-Bois, France) respectively. Clinical human serum albumin (HSA) was purchased from LFB laboratories (Les Ulis, France). 2-iminothiolane hydrochloride (2-IT), dithiothreitol (DTT), iodoacetamide (IAM) and endoproteinase Glu-C were purchased from Sigma-Aldrich (Breda, The Netherlands). Trypsin enzyme was purchased from promega (Madison, WI).

Human serum albumin chemical modification (HSA-2IT). Therapeutic HSA was chemically modified using 2-iminothiolane using a protocol adapted from Chaumet-Riffaud and coworkers [10]. Briefly, a volume of 25 μ L of therapeutic HSA at a concentration of 200 mg/mL was sampled and diluted to a final volume of 500 μ L using a solution of PBS/EDTA 5 mM. The sample was then transferred to a tube containing 7.5 mg of 2-IT. The mixture homogenized and incubated at room temperature under mild agitation for a total duration of 24h. In order to perform a kinetic follow-up of the reaction, 30 μ L corresponding to 300 μ g of the mixture was collected after incubation of 5 min, 30 min, 60 min, 2h, 4h and 24h. In order to stop the reaction and clean-up the sample from the remaining 2-IT, the collected samples were consequently desalted using Amicon 3kDa centrifugation membrane (Merck Millipore, Molsheim, France) by performing 5 cycles of dilution/concentration against milliQ H₂O. Samples were then stored at 5°C until further analysis.

MALDI-TOF mass spectrometry (MS) analysis. Intact has and 2IT-has samples obtained at different incubation times were characterized by mass spectrometry (MS). MS experiments were performed using a 4700 Proteomic Analyzer mass spectrometer from AB Sciex (Framingham, MA, USA). This instrument is implementing a matrix assisted laser desorption (MALDI) ionization source and a time-of-flight (TOF) mass analyzer. The matrix used was composed of α -Cyano-4-hydroxycinnamic acid (HCCA) and purchased from Sigma-Aldrich (Breda, The Netherlands). Based on bicinchoninic acid results, the samples were diluted to a final concentration of 1 μ g/ μ L in milli-Q water. A sample volume of 1 μ L was mixed with 3 μ L of HCCA (10 mg/mL in H₂O/ACN), then 1 μ L of this mixture was deposited on a stainless steel plate. The droplet was left to dry at room temperature. MS measurements were performed in positive ionization and linear mode in the same conditions and using the same laser energy. Ion transfer parameters were optimized using the clinical grade human serum albumin. MS spectra were acquired of the m/z range 8400 – 101000 Th. Each MS spectrum was obtained through the summation of 10000 laser shots. MS data treatment was performed using Data explorer software (AB Sciex, Framingham, MA, USA). The average mass, for the detected species was determined using BSA calibration standard kit (AB Sciex) considering the different charge states detected from 1 to 3.

Trypsin digestion of modified serum albumin. For each sample, a volume equivalent to a quantity of 25 μ g of protein was sampled. The sample was then diluted by adding systematically 10 μ L of ammonium bicarbonate 50 mM (pH 8.0) and the samples were incubated at 40°C for 10 min. DTT was added to the sample to obtain a final concentration of 10 mM. Samples were then heated to

90°C for 10 min. The sample was then cooled down to room temperature and IAM was added to a final concentration of 10 mM. Afterwards, the samples were incubated in the dark at room temperature for 20 min to allow alkylation of thiols residues. A volume of 1 µL of trypsin (0.5 µg/µL) was added to the mixtures which was left for incubation at room temperature for 3h. Another volume of 1 µL was added then and digestion was performed overnight at 37°C. Following digestion, a volume of 1 µL of formic acid (FA) 98% was added and the sample was left at RT for 2h. Finally, the samples were diluted to a final concentration of 0.25 µg/µL using H₂O (0.1% FA). Digested samples were stored at 5°C prior to UPLC-MS/MS analysis.

Peptide mixture characterization using liquid chromatography - tandem mass spectrometry (UPLC-MS/MS). Peptide mixtures obtained from either trypsin or Glu-C were separated by RP-UPLC (ACQUITY, Waters, Manchester, UK) using a C18 stationary phase (BEH C18 1.7 µm, 2.1 × 150 mm) purchased from Waters (St Quentin-en-Yvelines, France) directly hyphenated with a LTQ orbitrap XL mass spectrometer (Thermo Scientific, Bremen, Germany). The mobile phases were composed of 0.1% formic acid (FA) in water (mobile phase A) and 0.1% FA in acetonitrile (mobile phase B). Peptide separation was carried out using a gradient from 5 to 80% B in 38 min and maintained at 80% B for 3 min, at a flowrate of 100 µL/min. The sample volume used for UPLC-MS/MS experiments was systematically 10 µL. LTQ orbitrap XL MS was equipped with heated electrospray ionization source (HESI-II) from Thermo Scientific (Bremen, Germany). ESI source parameters were set as follow: ESI voltage -4.0 kV, sheath gas flowrate value was 40 and an auxiliary gas flowrate value of 12. ESI nebulizer temperature was set to 300°C. Capillary voltage and tube lens were set to 35V and 90V respectively. MS/MS experiments were performed in a Top5 data-dependent acquisition (DDA) composed of one full MS scan over the mass/charge (m/z) range 150-2000 followed by five sequential MS/MS realized on the five most intense ions detected at a minimum threshold of 500 counts. Full MS scans were collected in profile mode using the high resolution FTMS analyzer (R = 60,000) with a full scan AGC target of 1E6 and microscans = 1. The ion trap was used in centroid mode at normal scan rate to analyze MS/MS fragments. MSn AGC target was set to 1E4 with microscans = 3. Ions were selected for MS/MS using an isolation width of 2 Da, then fragmented by collision induced dissociation (CID) using a normalized CID energy of 35, an activation Q of 0.25 and an activation time of 30 msec. The default charge state selected was z = 2. Using these parameters, the total duty cycle was determined to be 0.65 sec. Parent ions were excluded from MS/MS experiments for 60 sec in case ion triggered an event twice in 15 sec using an exclusion mass width of ±1.5 Th. The instruments were controlled using Xcalibur 2.1.0 SP1 Build 1160 (Thermo Scientific, Bremen, Germany).

MS/MS data analysis. Data obtained from UPLC-MS/MS experiments were analyzed using Xcalibur Qual Browser 2.2 SP1.48 (Thermo Scientific, Bremen, Germany). Purely proteolytic peptides (no miscleavages or PTMs except from cysteine carbamidomethylation) were determined *in silico* considering HSA amino acid sequence described previously in the literature [27], for trypsin and endoproteinase Glu-C digestion respectively. For each proteolytic enzymes, conventional cleavage rules were applied [30] and carbamidomethylation of cysteine (+57.0215 Th) considered as a systematic modification. In order to identify conjugated lysine residues, a mass increment corresponding to the addition of 2-IT followed by carbamidomethylation of the thiol group (+158.0514 Th) was

considered. Peptide identifications were achieved manually from the conjunction of intact peptide mass measurements in full MS and MS/MS peptide fragments attribution using a mass tolerance inferior to 5 ppm in MS and 0.05 Th in MS/MS.

RESULTS AND DISCUSSION

Human serum albumin (HSA) chemical modification. Clinical grade human serum albumin (HSA), approved for therapeutic use, was used in order to prepare the protein conjugate (cf. experimental section). The implementation of this type of protein offers a cost-effective alternative and HSA is readily available. In addition, HSA exhibits a serum half-life of 19-20 days which appears particularly relevant for use as therapeutic products [31]. The relatively long serum half-life of HSA is attributed to the ability of the protein to interact the neonatal Fc receptor (FcRn) through a pH-dependent interaction which prevent the degradation of HSA in the endosome [32, 33]. HSA is also one of the main component of Abraxane[®] which is a Paclitaxel loaded onto albumin nanoparticles used for treatment of breast cancer [34]. Previously, HSA was implement in various cases in order to produce protein conjugates [35, 36].

The protocol used for the chemical functionalization of HSA was adapted from the work of Chaumet-Riffaud *et al* which allows to incorporate ligands on the lateral chain of lysine amino acids. In order to perform the conjugation process, lysine residues are activated using the primary amino group as emphasized in Figure 1A by incubation of the protein with 2-iminothiolane (2-IT), in order to generate sulfhydryl groups. Consequently to this step, ligands can be added through conventional thiol-maleimide Michael addition [10]. While “click chemistry” thiol-maleimide reaction is well described and demonstrated excellent reaction yield [37, 38], the primary step of the reaction is particularly crucial regarding the protein conjugate generated, and conditions the number and position of ligands effectively conjugated to the lysine residues. Using this approach, ligands can be coupled only by the intermediate of modified lysine residues and the number of charges can be conserved. Indeed, except from Cys³⁴ all cysteine residues present potentially providing thiol moieties are involved in disulfide bridges [39]. In order to perform a detailed structural characterization of the protein, therapeutic HSA was incubated with 2-IT and the reaction mixture was sampled at different times throughout the reaction process with a total incubation time of 24h. For each sample, buffer exchange treatment using milli-Q H₂O was performed in order to rapidly eliminate the 2-IT compound remaining. Following sample cleanup, the modified albumin HSA-2IT was characterized in parallel using MALDI-TOF MS and LC-MS/MS analysis.

Kinetic characterization of HSA chemical modification using MALDI-TOF MS. The different samples collected from the reaction media after various incubation times were characterized using MALDI-TOF MS. Using this analytical technique provides the analysis of the intact conjugated HSA in order to identify the overall number of 2-IT compounds effectively coupled to the peptide backbone of the protein. First, the analysis was performed on a sample of the clinical grade HSA product used as a starting material in order to provide a control analysis. The MS spectrum exhibited a signal corresponding to the protein having an average m/z ratio of 66412 (Figure 2A). Considering the performances of the instrument, the value obtained is in agreement with the average molecular mass of HSA. Note, that the MS peak width obtained is relatively narrow which shows the variants of HSA potentially composing

the sample have similar masses. For the HSA-2IT sample collected after an incubation time of 2 min (exposition), the obtained MS spectrum demonstrated a signal which exhibited an average m/z ratio of 68079. Therefore, a mass increase of 1667 Da could be observed compared to the unmodified HSA which corresponds to an average addition of 16 moieties (Figure 2B). The MS spectrum did not show any signal corresponding to the unmodified HSA which demonstrates that the sample content is modified in a homogeneous manner. However, the peak width corresponding to HSA-2IT is slightly broader as compared to the unmodified HSA. Therefore, this result shows that the occurrence of the chemical functionalization generates a narrow and nearly Gaussian distribution of HSA-2IT exhibiting different numbers of conjugated species. Similarly, MALDI-TOF MS analysis was performed for HSA-2IT sample incubated for 30 min (data not shown). The MS spectra allowed to measure an average m/z value of 68931 which represented a mass difference of 2519 Da corresponding to the coupling of 25 chemical compounds to the initial HSA. After an incubation time of 1h, MS spectra obtained from the analysis of HSA-2IT sample enabled to measure an average m/z value of 69154 (Figure 2C) that was found to correspond to the coupling of 27 species to the lysine residues of HSA. For the analysis of the sample incubated for 2h, the m/z value measured was 69676 therefore corresponding to the effective coupling of 32 species (Figure 2D). The MS analysis of the sample incubated with 2-IT for 4h enabled the identification of HSA-2IT with an average m/z value of 70229 for the single charged ion. Compared to the initial HSA sample, the reaction induced a mass increase of 3817 Da which could be attributed to the coupling of 38 compounds to the peptide backbone of the protein (Figure 3). It is important to note that for this sample, the peak width demonstrated to be significantly increased compared to previous samples. The MS spectra therefore shows that the distribution regarding the number of 2-IT compounds coupled to the protein tends to be broader when incubation time is longer while the average value continues to increase. Finally, the sample was characterized using MALDI-TOF MS consequently to a 24h incubation of HSA with 2-IT. MS data allowed to obtain an average m/z ratio equal to 70246 which corresponded to the addition of 38 compounds similarly to the 4h incubation sample. However, the MS spectra exhibited poor signal intensity compared to the previous HSA-2IT samples which was attributed to the occurrence of side reaction due to the important reaction time. In addition, the peak width demonstrated to be significantly larger than the previous samples characterized.

MALDI-TOF MS data from the different experiments were compiled in Figure 3 in order to estimate the number of 2-IT coupled to HSA depending on the reaction time. The results achieved using the MS intact protein analysis demonstrated to be particularly consistent which strongly suggests that the sample clean-up applied is relevant in order to quickly stop the 2-IT coupling reaction. As expected, the reactivity between the amino group of HSA lysine residues and 2-IT compounds is particularly intense at first because only a short exposition of the protein led to the modification of approximately 16 residues. This result is explained because in the early stage of the reaction, lysine residues easily accessible are reacting simultaneously with 2-IT therefore demonstrating a brutal increase of the number of coupled species. The compiled results show that when the reaction time was increased, the number of 2-IT compounds effectively coupled to the peptide backbone increased from 16 to 38 compounds. However, the rate of 2-IT coupling is gradually decreasing between the initial exposition (2 min) and a reaction time of 4h (Figure 3). This evolution regarding the coupling rate of 2-IT is explained by the respective kinetics of the different lysine residues which implies that a portion of the accessible lysine

residues are modified rapidly while a significant number of the residues require a longer time in order to be completely modified. When the reaction is further prolonged to 24h, results of MS analysis did not exhibit any additional increase regarding the number of conjugated compounds as emphasized in Figure 3. This observation is particularly interesting and demonstrates that a significant number of lysine residues (21 in total) cannot be modified from the chemical reaction with 2-IT compounds. The follow-up of the chemical functionalization proved in this case the functionalization kinetic and more importantly the limitation of the chemical reaction process. Thereby, the results achieved from MALDI-TOF MS analysis suggests that these lysine residues may not be accessible to the chemical compound allowing them to be preserved from undergoing the chemical modification. It therefore shows indubitably that 2-IT coupling represents the pivotal step in the formation of the protein conjugate, conditioning the nature of the product obtained, using this type of conjugation process. To illustrate the impact of such limitation, for instance in the case of HSA conjugates implemented as a medical imaging agent, the number of chelating moieties effectively coupled will be limited by the lysine residues activated. However, the number of chelating moieties will influence the intensity and the sensitivity of the signal provided by the HSA conjugate in scintigraphy, because it will directly impact the number of lanthanides (^{99m}Tc) the protein conjugate can form chelates with.

In order to go further in the characterization of the coupling process, MALDI-TOF MS analysis was used. The analysis allowed characterizing the number of 2-IT species effectively coupled to the peptide backbone of the conjugated protein and following the evolution of the modification process. This analytical method is particularly straightforward, moreover it allowed to estimate the number of 2-IT species coupled in a consistent manner using a limited quantity of sample. The results obtained demonstrated an important number of coupled species during the first minutes of the reaction which was followed by a gradual, however steeper, increase of the number of 2-IT added to the protein. MS data indicated that 85% of the available lysine residues were modified after 2h. The results of MALDI-TOF MS experiments allowed to determine an optimal reaction time of 1h in order to provide a significant number of modified lysine residues while preventing the occurrence of side reaction due to the presence of an important number of free thiols groups. In addition, the MS data suggested the possibility to guide to some extent the localization of the conjugated species. This could be potentially achieved by adjusting the reaction time in order to prevent the chemical modification of certain residues.

Analysis of HSA-2IT tryptic digest using reverse phase LC-MS/MS. In parallel to the intact protein analysis performed using MALDI-TOF MS, the samples of HSA-2IT collected after various reaction times were characterized in order to attempt to precisely locate the coupled compounds over the peptide backbone of the protein. To achieve the characterization, the samples were characterized using a methodology derived from proteomic analysis. Thus, HSA-2IT samples were undergoing trypsin proteolytic digestion. Consequently, the peptide mixture generated from the sample preparation was then separated and characterized by UPLC-MS/MS. In a preliminary experiment, a sample of clinical grade HSA was characterized in order to assess the performance of the LC-MS/MS method and provide a control experiment. Results achieved from the analysis of the unmodified HSA allowed to characterize nearly

the entire amino acid sequence (Table S-1). The extended sequence coverage could be achieved by performing the manual identification of the digested peptides. For each peptide, the identification was obtained from the high resolution measurement of the m/z ratio of the complete peptide provided in MS scans concomitantly to the identification of the fragment ions generated from the CID fragmentation during MS/MS events. Note, a few peptides demonstrated poor retention on the stationary phase leading to their elution practically in the dead volume. Because the complexity of the initial sample is fairly limited, the ion selection/exclusion procedure of the MS/MS method could be adjusted in order to still provide MS/MS data even in the case of poorly retained peptides partially coeluting. Thereby, these characteristics explain the outstanding sequence coverage achieved. The analysis of the HSA control sample demonstrated that the method is compatible with the comprehensive characterization of the primary structure of the protein.

The sample of HSA-2IT collected after a brief exposition (2 min) to 2-IT compound was prepared in a similar fashion and analyzed using LC-MS/MS. Conventionally, the protein is submitted to reduction followed by alkylation using iodoacetamide (IAM) prior to the proteolytic digestion in order to maximize the efficiency of the digestion. As a consequence, in addition to the cysteine residues IAM is supposed to react with the free thiols generated from the coupling of 2IT to the protein like represented in Figure 1B. The total ion chromatogram (TIC) obtained from the LC-MS/MS experiment exhibited a similar intensity compared to the initial protein demonstrating a decent separation of the peptides composing the mixture. In order to identify the peptides incorporating 2-IT compounds, theoretical values corresponding to the modified peptides were extracted from MS data. Considering purely tryptic peptides without missed cleavage did not enable to identify the addition of 2-IT compounds. Indeed, as emphasized in Figure 4 the peptide T08 could only be identified free from 2-IT. However, by considering the peptide T08_09 which corresponds to the occurrence of a missed cleavage on Lys⁷⁴, it was possible to identify the m/z ratio corresponding to the unmodified peptides in addition to the homologous peptide exhibiting the coupling of 2-IT. In this case, because Lys⁷⁴ is the only residue on the peptide which may experience the reaction with 2-IT, it was possible to conclude that Lys⁷⁴ was successfully coupled to the chemical compound. Also, as expected a mass difference of +158.0514 Da was observed between the unmodified peptide and the conjugated counterpart (Table S-2). This mass difference appeared to correspond to the addition of 2-IT (+101.0299 Da) to the lysine residue concerned followed by the carbamidomethylation (+57.0215 Da) of the thiol moiety during the reduction/alkylation. The unmodified peptide T08_09 exhibited a retention time of 14.55 min whereas the equivalent peptide bearing a coupled 2-IT (T08*09) was eluted at a retention time of 14.85 min which is presented in Figure 4. The difference in retention times appears in agreement with the presence of the additional alkylated 2-IT to the peptide leading to a slightly longer retention time.

Based on this observation, the consideration of missed cleavage peptides enabled the identification of different lysine residues exhibiting the coupling of 2-IT compound (Table S-2). Systematically, the amino acids coupled with 2-IT compound could be identified only from missed cleavage peptides where the functionalized residue was not experiencing the enzymatic cleavage. Therefore, this effect could be attributed to the presence of 2-IT which inhibited the activity of trypsin on the modified lysine residues. In addition, the different peptides exhibiting the addition of the chemical compound consistently demonstrated an increased retention time of approximately 0.3 min compared to the unmodified homologous (Table S-2). As a consequence, the difference of retention time was

taken into account for the identification of coupled 2-IT compounds in order to provide an improved confidence in the result and prevent the occurrence of false positive.

As part of the LC-MS/MS analysis, the separated peptides were selected for fragmentation followed by MS/MS characterization. Figure 5 represents MS/MS spectra obtained for two different peptides which showed in each case the presence of coupled 2-IT compounds. Thus, the attribution of the fragment ions allowed principally the identification of y- and b-ions. Regarding the conjugated lysine residues, the fragments observed in MS/MS confirmed a mass difference corresponding to the residue bearing an alkylated 2-IT moiety (Figure 1B). Rather than being problematic, the characteristic mass difference of carbamidomethylated 2-IT (+158.0415 Da) was found to be useful especially in order to prevent confusion between the presence of coupled 2-IT (+101.0299 Da) and the mass corresponding to a threonine residue (+101.0476 Da). Note that due to the sample treatment applied, only alkylated 2-IT could be strictly identified which underlined the high reactivity of the free thiols groups coupled to lysine residues. As it is illustrated on the annotation obtained in Figure 5, MS/MS spectra enabled to attribute without any ambiguity the presence of coupled 2-IT to the different lysine residues of HSA. In addition, MS/MS spectra allowed to attribute precisely the amino acid residue experiencing 2-IT coupling like illustrated in Figure 5 even in the case when the peptide included two lysine residues, further proving the inhibition of the activity of trypsin.

The data achieved from LC-MS/MS analysis of HSA-2IT generated from the exposition of the protein (2 min) to the chemical compound were used in order to determine the localization of the coupled 2-IT compounds. As emphasized in Figure 6, LC-MS/MS data allowed to identify a total of 19 lysine residues experiencing 2-IT conjugation. Similarly to the analysis of the initial clinical grade HSA, the characterization enabled to achieve nearly complete sequence coverage except from single amino acid residues. The minor difference regarding to the number of conjugated lysine residues compared to MALDI-TOF MS experiments was attributed because intact protein MS analysis provides a global value whereas LC-MS/MS enables to characterize independently every conjugated lysine which still remains a probabilistic event. The inhibition of trypsin activity for lysine residues coupled to 2-IT has revealed to be particularly interesting for the site-specific characterization. Indeed, the consistent occurrence of missed cleavage on conjugated lysine residues enabled in various cases to distinguish the modification even in the case of close by residues. Also, the peptide diversity generated from the proteolytic digestion allowed to cover in a robust manner each potential conjugated residue. Therefore, the relevant properties of the LC-MS/MS characterization developed enabled to identify the occurrence of the reaction independently even for consecutive lysine residues like Lys¹³⁶-2IT / Lys¹³⁷-2IT or Lys⁵²⁴ / Lys⁵²⁵-2IT (Figure 6A). The structural characterization achieved using LC-MS/MS analysis shows that significant portion of the peptide backbone is not experiencing 2-IT coupling. For instance, no coupling of 2-IT could be observed over the amino acid portions between Lys¹⁸¹ and Lys²²⁵ or between Lys²⁸¹ and Lys³⁵¹ even though they contain 5 and 4 lysine residues respectively. This observation was strongly suggesting that portion of the peptide backbone may not be accessible to 2-IT compounds solubilized in the buffer.

A similar experiment was performed for HSA-2IT sample incubated for 30 min. In this case as well, LC-MS/MS data enabled to characterize in details the 2-IT coupled to the lysine residues. The identified residues exhibiting 2-IT coupling were consistent with

the one observed for the previous analysis, with the identification of the chemical specie on the same position. In addition, several lysine residues demonstrated the coupling of 2-IT (Lys⁶⁴, Lys⁹³, Lys²⁴⁰, Lys²⁶², Lys⁴³² and Lys⁵⁴¹). Therefore, results showed conjugation of the chemical compound for 25 lysine residues out of the 59 composing the amino acid sequence. The number of residues functionalized determined from LC-MS/MS analysis was in correlation with the value estimated from MALDI-TOF MS analysis (Figure 3).

Regarding the analysis of HSA-2IT obtained consequently to an incubation of 1h, the results represented on Figure 6B emphasizes the characterization of 26 lysine residues demonstrating functionalization with 2-IT specie. Compared to the analysis of HSA-2IT sample incubated for 30 min, the LC-MS/MS characterization of the sample demonstrated that an incubation of 1h enabled to distinguish the additional functionalization of Lys²⁷⁶ (Figure 6B). The analytical methodology developed allowed in this particular case to demonstrated the faster kinetic of Lys²⁷⁴ coupling which can be identified in the early samples while the conjugation of Lys²⁷⁶ is characterized solely after an incubation of 1h even though these amino acids are particularly close in the sequence. The results obtained therefore accentuate the precision of the structural characterization achieved regarding the localization of the conjugated residues. Concerning LC-MS/MS analysis of the sample collected after an incubation of 2h, the results provided the identification of 2-IT coupling on exactly the same lysine residues as the previous experiments. In addition, MS/MS data enabled to identify 2-IT conjugation to the residues Lys⁴⁴⁴ and with significantly lower intensities the residues Lys¹⁶², Lys¹⁹⁰, Lys¹⁹⁵, Lys⁴⁰² and Lys⁴³⁹. Similarly to the previous samples, the site-specific characterization of 2-IT conjugation demonstrated to be in agreement with the value estimated using MALDI-TOF MS analysis.

The analytical methodology developed enabled to characterize and locate in a precise manner the lysine residues conjugated with 2-IT chemical compound. The consideration of MS spectra in concomitance with MS/MS data allowed to attribute without any ambiguity the different amino acids modified along the sequence of the protein. Therefore, LC-MS/MS data revealed that the presence of conjugated 2-IT was inhibiting the activity of trypsin and that the conjugated thiol groups were experiencing complete carbamidomethylation by IAM during the sample treatment. Rather than hampering the ability to perform a comprehensive characterization, these characteristics demonstrated to be useful to the analysis. Especially, they enabled to unambiguously distinguish 2-IT coupling occurring on lysine residues closely positioned on the peptide backbone. The characterization achieved using LC-MS/MS demonstrated to be in correlation with MALDI-TOF MS analysis. The crossing of the data generated from the different techniques allowed from one aspect to locate precisely the lysine residues exhibiting significantly lower conjugation rate which resulted in progressive conjugation over time on the intact protein level. From another aspect, crossing the data obtained from the different experiments have allowed to identify the lysine residues which were completely excluded from 2-IT coupling. As a consequence, the analytical strategy developed allowed to give precious information regarding the chemical reaction by providing the amino acid affected and an additional insight regarding their respective kinetics. To the best of our knowledge, such level of characterization regarding the chemical functionalization is quite unique and was not previously described. The chemical functionalization of HSA is performed in pseudo-physiological conditions (cf. experimental section) in order to maintain as much as possible the globular structure of the protein.

Thereby, the absence of reactivity from a significant number of lysine residues could be potentially due to their position in the tertiary structure of the protein making them inaccessible to 2-IT solubilized in the buffer.

Comparison between LC-MS/MS characterization and crystallographic data. The characterization performed using the LC-MS/MS method demonstrated the possibility to locate in a precise manner the lysine residues effectively coupled with 2-IT compounds. Thereby, the limitations of the reaction emphasized in MALDI-TOF MS analysis by the occurrence of a plateau concerning the number of 2-IT compounds coupled (Figure 3) could be linked to the lysine which did not experienced modification based on LC-MS/MS data. Consequently, the characterization obtained from LC-MS/MS analysis was compared to the crystallographic data of HSA in order to further investigate the causal connection between the localization of the lysine residue and the occurrence of the chemical functionalization. The crystal structure of HSA obtained using X-ray diffraction has been described previously in the literature and extensively studied since [40]. The structure of HSA available in the “RCSB protein data bank” was used for the comparison [41].

The lysine residues present on the peptide backbone were highlighted differently on the tridimensional representation of HSA structure depending on their chemical coupling with 2-IT which is represented in Figure 7. As expected, the modelling of the structure demonstrated the presence of several lysine residues exposed to their external environment which is illustrated in Figure 7A for Lys²³³ and Lys²⁴⁰. Therefore, the residues exposed like Lys²³³ and Lys²⁴⁰ demonstrated to experience rapid coupling with 2-IT which could be deduced from their identification in LC-MS/MS data originating from the exposition sample (Figure 6A). That type of positioning of the amino acid could be also observed for instance in the case of Lys¹², Lys³⁷⁸, Lys⁵⁷³ and Lys⁵⁷⁴ which were identified in the same manner as being coupled to 2-IT early in the reaction process (Figure S-1). As a consequence, the comparison of the site-specific characterization provided from LC-MS/MS data is favorable with a potential influence of the accessibility of the residues and the reactivity with 2-IT. In addition to consider lysine residues demonstrating 2-IT conjugation, the structure of the protein was investigated for the residues which did not exhibited chemical modification. Figure 7B shows domain IIIa-helix 2 and domain IIIb-helix 2 of the HSA tertiary structure. In that case, it was possible to observe that the two helixes are positioned in parallel planes. This conformation of the protein has for consequence that the direct environment of Lys⁴¹³ and Lys⁵³⁴ is surrounded by other part of the peptide backbone of the protein. In addition, based on the structural representation obtained from crystallography, both amino acids are close in space and partially facing each other. Therefore, the absence of 2-IT coupling for the two residues could be attributed to their inaccessibility caused from steric hindrance. In that sense, Figure 7B shows that Lys⁴¹⁴ which is positioned in a different direction due to the shape of the helix (Domain IIIa, helix 2) escapes from the steric effect and was found out to undergo chemical modification (Figure 6).

In a similar manner, different lysine residues which could not be coupled with 2-IT compound demonstrated to be buried deeper into the tertiary structure of the protein like Lys²¹², Lys²⁸⁶ or Lys¹⁷⁴ (Figure S-2). However, several lysine residues which considering their position should be accessible to 2-IT compounds were identified as unmodified from LC-MS/MS analysis. This observation therefore

suggests that even if the 2-IT conjugation to the lateral part of lysine residues is mainly driven by their accessibility to the solvent, additional parameters may affect the chemical reaction as well. The results obtained from the correlation of LC-MS/MS analysis and the structure of the protein resolved in a crystallography demonstrated to represent a powerful approach. Indeed, the conjunction of the characterization provided by LC-MS/MS analysis and crystallography demonstrated in this case a relevant complementarity. The analytical methodology described therefore allowed to identify without any ambiguity the lysine residues experiencing coupling with 2-IT and study their respective spatial position, in order to understand in a detailed manner, the structural parameters influencing the chemical reaction.

CONCLUSION

In the present work, we could develop an analytical methodology for the in-depth assessment of the chemical reaction concerning lysine residues used to produce certain types of protein conjugates. The analytical strategy implemented, was based on the conjunction of MALDI-TOF MS analysis, LC-MS/MS analysis and crystallography structure. It could be applied to the characterization of HSA modified through lysine residues, previously described for the production of conjugates used in medical imaging and diagnostics. MALDI-TOF MS analysis demonstrated the possibility to estimate in a straightforward manner the number of chemical compounds effectively coupled to the protein. In addition, the kinetic of the reaction was investigated on the intact protein level using this technique which allowed to identify that the coupling reaction rate on the protein level was rapidly decreasing. In addition, results showed that a significant number of lysine residues did not experienced modification therefore clearly revealing a limitation of the conjugation process. Concomitantly, the samples were submitted to proteolytic digestion followed by LC-MS/MS analysis. The characterization allowed to locate in the precise manner the compounds coupled to the peptide backbone of the protein without any ambiguity. LC-MS/MS data demonstrated to be in correlation with the analysis performed on the intact level. MS/MS results enabled to distinguish exactly the lysine residues exhibiting hasten 2-IT conjugation compared to the residues which required longer reaction time in order to be modified. Also, MS/MS data could be used to identified the amino acid residues which could remained conserved. The conjunction of the analytical techniques allowed to correlate the kinetic of the reaction over the intact protein to the evolution of its primary structure, revealing an improved insights regarding the reaction process. Finally, LC-MS/MS results were compared to the structure of the protein using crystallographic data. The comparison enabled to identify that lysine residues easily accessible to the solvent demonstrated to be modified early during the incubation. On the opposite side, a significant number of residues buried deeper inside the tertiary structure of HSA demonstrated to be totally preserved from reaction with 2-IT compound. Thus, results showed the relevant complementarity of the analytical techniques employed. It is the first time that such level of characterization could be achieved regarding the chemical reaction used in order to generate protein conjugates. The characterization strategy developed demonstrated to be particularly informative with the possibility to precisely locate the position of the ligands despite the inherent complexity of the protein, in addition to provide an excellent sensitivity. The methodology developed could be implemented to optimize the conditions of protein conjugates production for instance for a wide variety of protein format like ADCs. Also, because a

significant number of lysine residues demonstrated slower reaction rate, it is possible to potentially envisage to tune the reaction process by adjusting the reaction time in order to prevent the modification of some amino acids, like in the case of HSA fatty acids binding sites (Lys³⁵¹, Lys⁴⁷⁵) to preserve the activity of the protein while still producing the protein conjugate.

ACKNOWLEDGMENTS

The authors would like to thank Annie Leyris, Christiane Gabriel-Calixte and Oura Conteh (Faculté de Pharmacie, Université de Paris, Paris, France) for their precious support. The authors would like to express their gratitude to Dr. Marie-Claude Menet (Faculté de Pharmacie, Université de Paris, Paris, France) for interesting discussions regarding protein analysis.

REFERENCES

- [1] E.H. Romond, E.A. Perez, J. Bryant, V.J. Suman, C.E. Geyer, N.E. Davidson, E. Tan-Chiu, S. Martino, S. Paik, P.A. Kaufman, S.M. Swain, T.M. Pisansky, L. Fehrenbacher, L.A. Kutteh, V.G. Vogel, D.W. Visscher, G. Yothers, R.B. Jenkins, A.M. Brown, S.R. Dakhil, E.P. Mamounas, W.L. Lingle, P.M. Klein, J.N. Ingle, N. Wolmark, Trastuzumab plus Adjuvant Chemotherapy for Operable HER2-Positive Breast Cancer, *New England Journal of Medicine* 353(16) (2005) 1673-1684.
- [2] J.F. Colombel, W.J. Sandborn, W. Reinisch, G.J. Mantzaris, A. Kornbluth, D. Rachmilewitz, S. Lichtiger, G. D'Haens, R.H. Diamond, D.L. Broussard, K.L. Tang, C.J. van der Woude, P. Rutgeerts, Infliximab, Azathioprine, or Combination Therapy for Crohn's Disease, *New England Journal of Medicine* 362(15) (2010) 1383-1395.
- [3] H. Lagassé, A. Alexaki, V. Simhadri, N. Katagiri, W. Jankowski, Z. Sauna, C. Kimchi-Sarfaty, Recent advances in (therapeutic protein) drug development [version 1; peer review: 2 approved], *F1000Research* 6(113) (2017).
- [4] M.L. Lindsey, M. Jung, M.E. Hall, K.Y. DeLeon-Pennell, Proteomic analysis of the cardiac extracellular matrix: clinical research applications, *Expert Review of Proteomics* 15(2) (2018) 105-112.
- [5] D.S. Dimitrov, Therapeutic Proteins, in: V. Voynov, J.A. Caravella (Eds.), *Therapeutic Proteins: Methods and Protocols*, Humana Press, Totowa, NJ, 2012, pp. 1-26.
- [6] S. Ornes, Antibody–drug conjugates, *Proceedings of the National Academy of Sciences* 110(34) (2013) 13695-13695.
- [7] A. Beck, L. Goetsch, C. Dumontet, N. Corvaia, Strategies and challenges for the next generation of antibody–drug conjugates, *Nature Reviews Drug Discovery* 16 (2017) 315.
- [8] J.M. Jeong, M.K. Hong, J. Lee, M. Son, Y. So, D.S. Lee, J.-K. Chung, M.C. Lee, 99mTc-Neolactosylated Human Serum Albumin for Imaging the Hepatic Asialoglycoprotein Receptor, *Bioconjugate Chemistry* 15(4) (2004) 850-855.
- [9] K. Lemdani, H. Salmon, R. Gahoual, M. Bessodes, D. Scherman, P. Houzé, J. Seguin, N. Mignet, Assessment of the targeting specificity of a fluorescent albumin conceived as a preclinical agent of the liver function, *Nanoscale* 10(45) (2018) 21151-21160.
- [10] P. Chaumet-Riffaud, I. Martinez-Duncker, A.-L. Marty, C. Richard, A. Prigent, F. Moati, L. Sarda-Mantel, D. Scherman, M. Bessodes, N. Mignet, Synthesis and Application of Lactosylated, 99mTc Chelating Albumin for Measurement of Liver Function, *Bioconjugate Chemistry* 21(4) (2010) 589-596.
- [11] W. de Graaf, R.J. Bennink, R. Veteläinen, T.M. van Gulik, Nuclear Imaging Techniques for the Assessment of Hepatic Function in Liver Surgery and Transplantation, *Journal of Nuclear Medicine* 51(5) (2010) 742-752.
- [12] N. Kokudo, D.R. Vera, M. Makuuchi, Clinical application of TcGSA, *Nuclear Medicine and Biology* 30(8) (2003) 845-849.
- [13] A. Beck, J.F. Haeuw, T. Wurch, L. Goetsch, C. Bailly, N. Corvaia, The next generation of antibody–drug conjugates comes of age, *Discovery medicine* 10(53) (2010) 329-39.
- [14] P. Akkapeddi, S.-A. Azizi, A.M. Freedy, P.M.S.D. Cal, P.M.P. Gois, G.J.L. Bernardes, Construction of homogeneous antibody–drug conjugates using site-selective protein chemistry, *Chemical Science* 7(5) (2016) 2954-2963.
- [15] A. Younes, N.L. Bartlett, J.P. Leonard, D.A. Kennedy, C.M. Lynch, E.L. Sievers, A. Forero-Torres, Brentuximab Vedotin (SGN-35) for Relapsed CD30-Positive Lymphomas, *New England Journal of Medicine* 363(19) (2010) 1812-1821.
- [16] S. Verma, D. Miles, L. Gianni, I.E. Krop, M. Welslau, J. Baselga, M. Pegram, D.-Y. Oh, V. Diéras, E. Guardino, L. Fang, M.W. Lu, S. Olsen, K. Blackwell, Trastuzumab Emtansine for HER2-Positive Advanced Breast Cancer, *New England Journal of Medicine* 367(19) (2012) 1783-1791.
- [17] T. Botzanowski, S. Erb, O. Hernandez-Alba, A. Ehkirch, O. Colas, E. Wagner-Rousset, D. Rabuka, A. Beck, P.M. Drake, S. Cianférani, Insights from native mass spectrometry approaches for top- and middle- level characterization of site-specific antibody–drug conjugates, *mAbs* 9(5) (2017) 801-811.

- [18] N. Said, R. Gahoual, L. Kuhn, A. Beck, Y.-N. François, E. Leize-Wagner, Structural characterization of antibody drug conjugate by a combination of intact, middle-up and bottom-up techniques using sheathless capillary electrophoresis – Tandem mass spectrometry as nanoESI infusion platform and separation method, *Analytica Chimica Acta* 918 (2016) 50-59.
- [19] A. Beck, V. D’Atri, A. Ehkirch, S. Fekete, O. Hernandez-Alba, R. Gahoual, E. Leize-Wagner, Y. François, D. Guillarme, S. Cianférani, Cutting-edge multi-level analytical and structural characterization of antibody-drug conjugates: present and future, *Expert Review of Proteomics* 16(4) (2019) 337-362.
- [20] S. Fekete, D. Guillarme, P. Sandra, K. Sandra, Chromatographic, Electrophoretic, and Mass Spectrometric Methods for the Analytical Characterization of Protein Biopharmaceuticals, *Analytical Chemistry* 88(1) (2016) 480-507.
- [21] R. Gahoual, J.-M. Busnel, A. Beck, Y.-N. François, E. Leize-Wagner, Full Antibody Primary Structure and Microvariant Characterization in a Single Injection Using Transient Isotachopheresis and Sheathless Capillary Electrophoresis–Tandem Mass Spectrometry, *Analytical Chemistry* 86(18) (2014) 9074-9081.
- [22] J. Marcoux, T. Champion, O. Colas, E. Wagner-Rousset, N. Corvaia, A. Van Dorsselaer, A. Beck, S. Cianférani, Native mass spectrometry and ion mobility characterization of trastuzumab emtansine, a lysine-linked antibody drug conjugate, *Protein Science* 24(8) (2015) 1210-1223.
- [23] V. D’Atri, S. Fekete, A. Clarke, J.-L. Veuthey, D. Guillarme, Recent Advances in Chromatography for Pharmaceutical Analysis, *Analytical Chemistry* 91(1) (2019) 210-239.
- [24] R. Gahoual, A. Beck, E. Leize-Wagner, Y.-N. François, Cutting-edge capillary electrophoresis characterization of monoclonal antibodies and related products, *Journal of Chromatography B* 1032 (2016) 61-78.
- [25] A. Lechner, J. Giorgetti, R. Gahoual, A. Beck, E. Leize-Wagner, Y.-N. François, Insights from capillary electrophoresis approaches for characterization of monoclonal antibodies and antibody drug conjugates in the period 2016–2018, *Journal of Chromatography B* 1122-1123 (2019) 1-17.
- [26] H. Salmon, R. Gahoual, P. Houzé, T. Ibrahim, M. Bessodes, D. Scherman, J. Seguin, N. Mignet, Europium labeled lactosylated albumin as a model workflow for the development of biotherapeutics, *Nanomedicine: Nanotechnology, Biology and Medicine* (2019).
- [27] B. Meloun, L. Morávek, V. Kostka, Complete amino acid sequence of human serum albumin, *FEBS Letters* 58(1) (1975) 134-137.
- [28] A.M. Merlot, D.S. Kalinowski, D.R. Richardson, Unraveling the mysteries of serum albumin—more than just a serum protein, *Frontiers in Physiology* 5(299) (2014).
- [29] Y. He, T. Ning, T. Xie, Q. Qiu, L. Zhang, Y. Sun, D. Jiang, K. Fu, F. Yin, W. Zhang, L. Shen, H. Wang, J. Li, Q. Lin, Y. Sun, H. Li, Y. Zhu, D. Yang, Large-scale production of functional human serum albumin from transgenic rice seeds, *Proceedings of the National Academy of Sciences* 108(47) (2011) 19078-19083.
- [30] P. Giansanti, L. Tsiatsiani, T.Y. Low, A.J.R. Heck, Six alternative proteases for mass spectrometry–based proteomics beyond trypsin, *Nature Protocols* 11 (2016) 993.
- [31] D. Sleep, J. Cameron, L.R. Evans, Albumin as a versatile platform for drug half-life extension, *Biochimica et Biophysica Acta (BBA) - General Subjects* 1830(12) (2013) 5526-5534.
- [32] C. Chaudhury, S. Mehnaz, J.M. Robinson, W.L. Hayton, D.K. Pearl, D.C. Roopenian, C.L. Anderson, The Major Histocompatibility Complex–related Fc Receptor for IgG (FcRn) Binds Albumin and Prolongs Its Lifespan, *The Journal of Experimental Medicine* 197(3) (2003) 315-322.
- [33] D.C. Roopenian, S. Akilesh, FcRn: the neonatal Fc receptor comes of age, *Nature Reviews Immunology* 7(9) (2007) 715-725.
- [34] W.J. Gradishar, Albumin-bound paclitaxel: a next-generation taxane, *Expert Opinion on Pharmacotherapy* 7(8) (2006) 1041-1053.
- [35] F. Esmaeili, R. Dinarvand, M.H. Ghahremani, M. Amini, H. Rouhani, N. Sepehri, S.N. Ostad, F. Atyabi, Docetaxel–albumin conjugates: Preparation, in vitro evaluation and biodistribution studies, *Journal of Pharmaceutical Sciences* 98(8) (2009) 2718-2730.

- [36] F. Kratz, U. Beyer, T. Roth, M.T. Schütte, A. Unold, H.H. Fiebig, C. Unger, Albumin Conjugates of the Anticancer Drug Chlorambucil: Synthesis, Characterization, and In Vitro Efficacy, *Archiv der Pharmazie* 331(2) (1998) 47-53.
- [37] B.H. Northrop, S.H. Frayne, U. Choudhary, Thiol–maleimide “click” chemistry: evaluating the influence of solvent, initiator, and thiol on the reaction mechanism, kinetics, and selectivity, *Polymer Chemistry* 6(18) (2015) 3415-3430.
- [38] J.M.J.M. Ravasco, H. Faustino, A. Trindade, P.M.P. Gois, Bioconjugation with Maleimides: A Useful Tool for Chemical Biology, *Chemistry – A European Journal* 25(1) (2019) 43-59.
- [39] J.L. Beck, S. Ambahera, S.R. Yong, M.M. Sheil, J.d. Jersey, S.F. Ralph, Direct observation of covalent adducts with Cys34 of human serum albumin using mass spectrometry, *Analytical Biochemistry* 325(2) (2004) 326-336.
- [40] S. Sugio, A. Kashima, S. Mochizuki, M. Noda, K. Kobayashi, Crystal structure of human serum albumin at 2.5 Å resolution, *Protein Engineering, Design and Selection* 12(6) (1999) 439-446.
- [41] F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer, M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, The protein data bank: A computer-based archival file for macromolecular structures, *Journal of Molecular Biology* 112(3) (1977) 535-542.

FIGURES

Figure 1. Schematic representation of the reaction performed A) between HSA lysine residues and 2-IT chemical compound in order to generate available sulfhydryl groups used for the preparation of HSA protein conjugates (HSA-2IT) and B) concerning the reduction/alkylation of HSA-2IT using iodoacetamide (IAM) prior to proteolytic digestion.

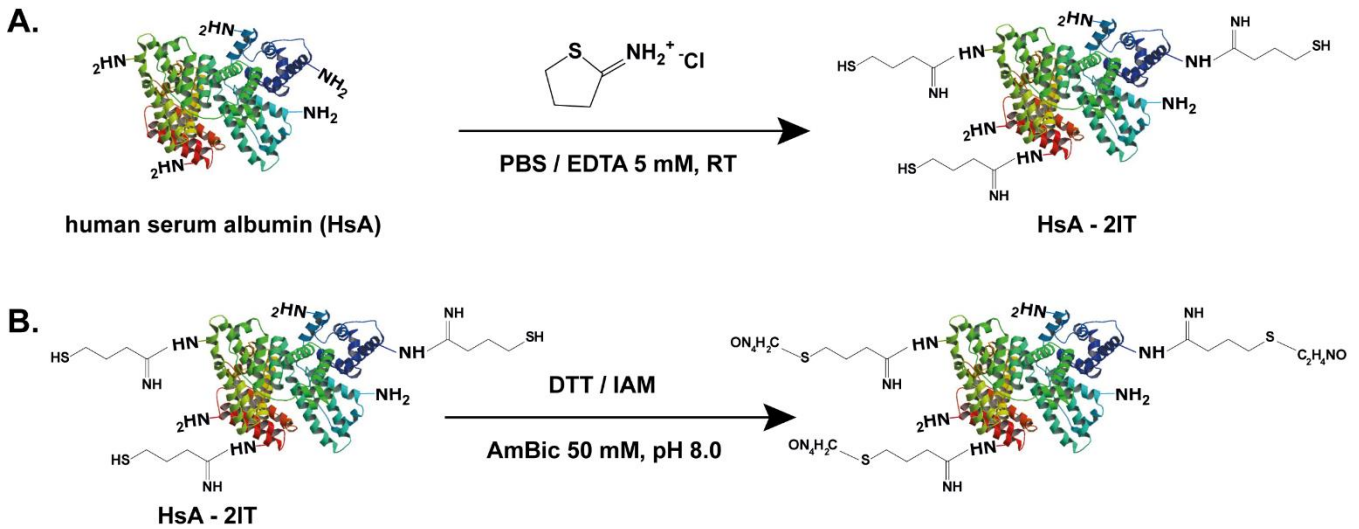


Figure 2. MALDI-TOF MS spectra corresponding to A) clinical grade HSA, B) HSA-2IT sample shortly exposed (2 min) to 2-IT, C) HSA-2IT sample incubated for 1h and D) HSA-2IT sample incubated for 2h. Average deconvoluted mass of each sample was determined considering the different charge states detected.

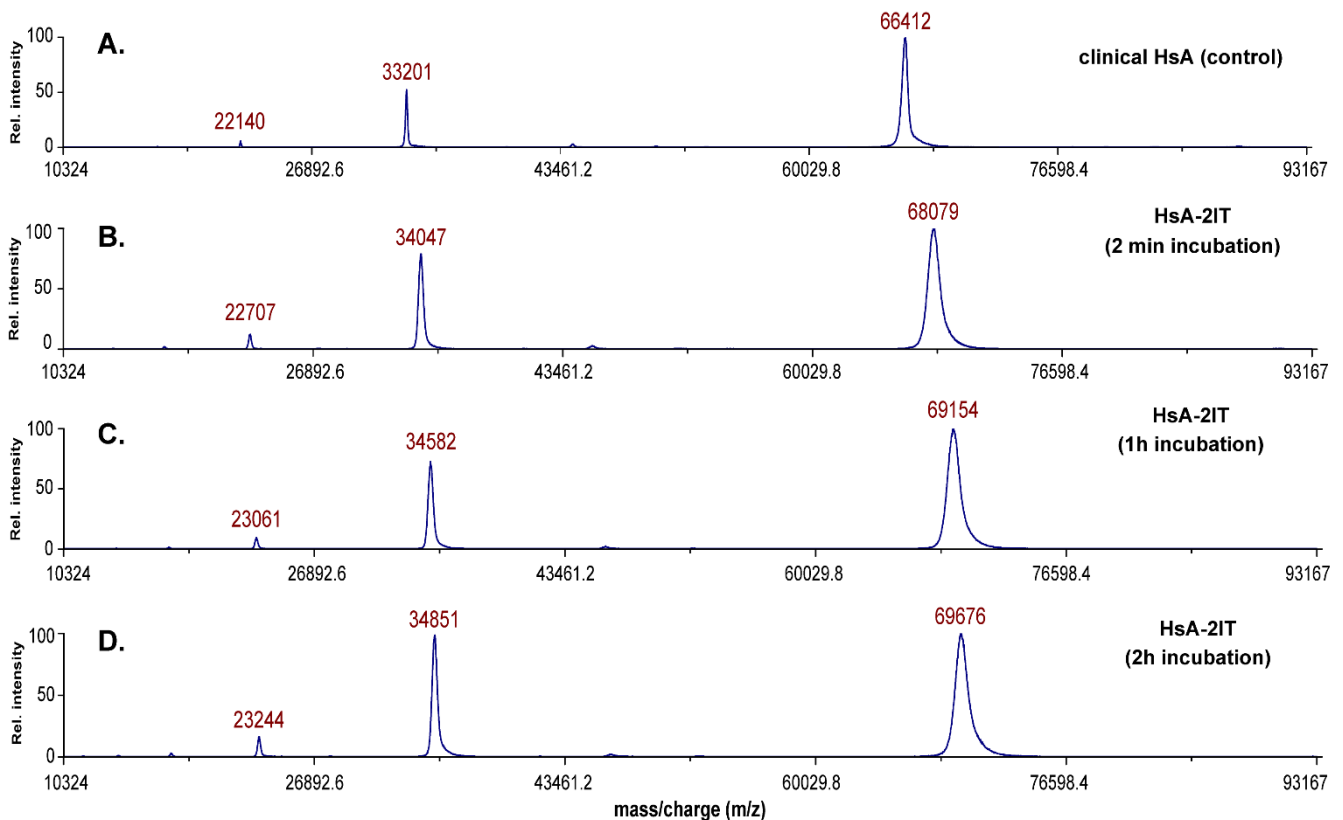


Figure 3. Graphical representation of the number of 2-IT species coupled to HSA peptide backbone depending on the reaction time determined from intact protein analysis using MALDI-TOF MS (n = 3). The values were estimated by comparing the mass shift between intact HSA and HSA-2IT samples considering 2-IT theoretical mass increment (+101.03 Da).

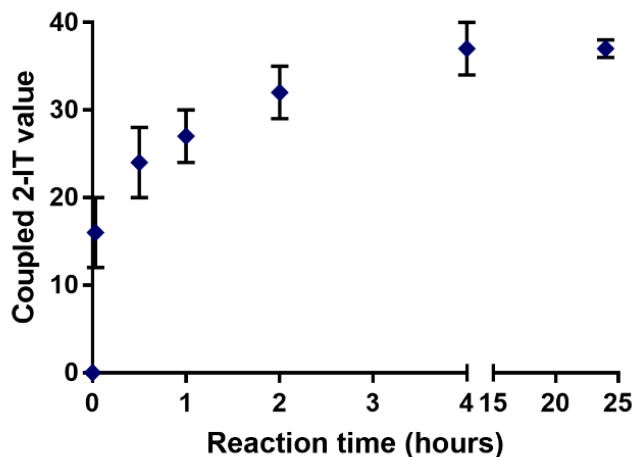


Figure 4. Extracted ion chromatogram (XIC) corresponding to m/z ratio of peptide T08 (509.2722, 2+), peptide T08_09 (483.7656, 3+) without conjugated species and peptide T08*09 (523.2783, 3+) exhibiting the coupling of 2-IT to the lysine residue (K* annotation), obtained from the analysis of the peptide mixture originating from HSA-2IT sample incubated for 2h. Peptides conjugated with 2-IT are represented using blue bars and unmodified peptides by green bars.

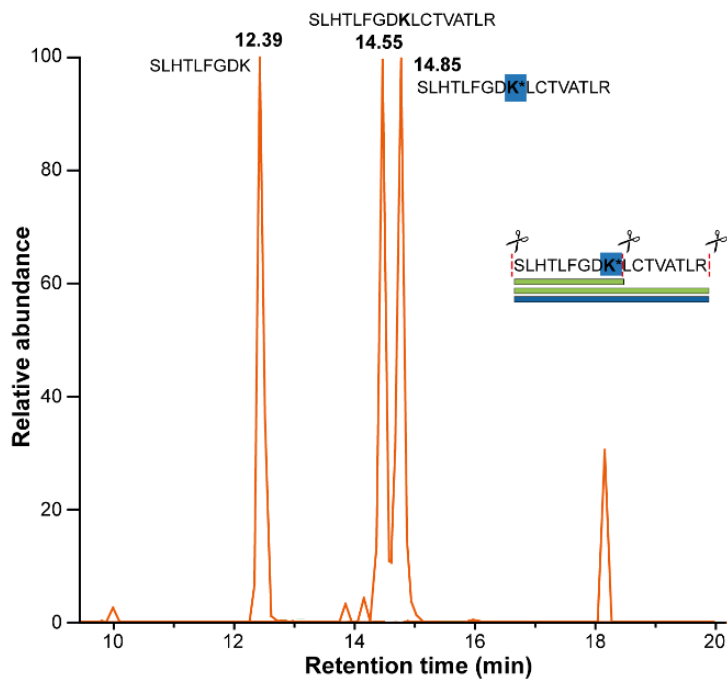


Figure 5. MS/MS spectra corresponding to A) peptide T15*16 (**K***YLYEIAR) and B) peptide T47*48 (LA**K***TYETTLEK) illustrating the fragment identification which allowed to attribute the coupling of 2-IT to lysine lateral chain. Experimental conditions: cf. experimental section. Lysine residues coupled with 2-IT compound are designated as **K***.

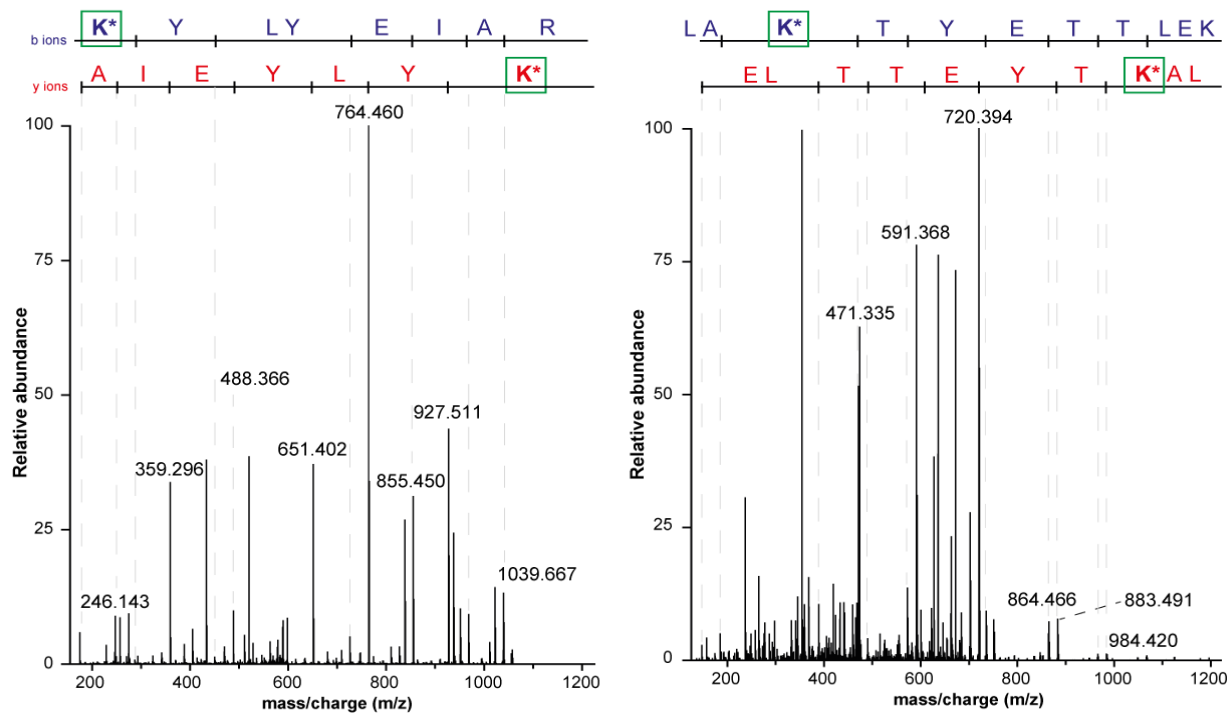


Figure 6. Sequence coverage obtained for A) HSA-2IT sample obtained through short exposition to 2-IT (left panel) and B) HSA-2IT sample obtained using an incubation time of 1h (right panel). UPLC-MS/MS spectra were recorded on LTQ orbitrap XL instrument (Thermo Scientific, Bremen, Germany). Lysine residues coupled with 2-IT are marked in blue. Peptides conjugated with 2-IT are represented using blue bars and unmodified peptides by green bars.



Figure 7. Schematic representation of the structure of HSA illustrating A) the accessibility of the residues Lys²³³, Lys²⁴⁰ and Lys²⁶² experiencing 2-IT coupling and B) the residues Lys⁴¹³ and Lys⁵³⁴ which did not react with the chemical compound. Modified residues are represented in green and conserved residues are represented in red.

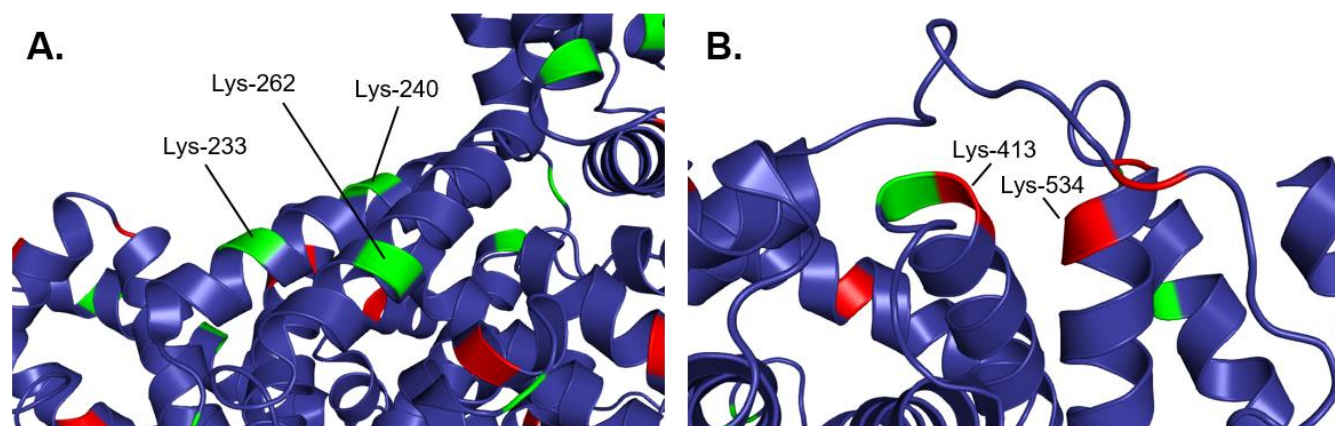


FIGURE CAPTIONS

Figure 1. Schematic representation of the reaction performed A) between HSA lysine residues and 2-IT chemical compound in order to generate available sulfhydryl groups used for the preparation of HSA protein conjugates (HSA-2IT) and B) concerning the reduction/alkylation of HSA-2IT using iodoacetamide (IAM) prior to proteolytic digestion.

Figure 2. MALDI-TOF MS spectra corresponding to A) clinical grade HSA, B) HSA-2IT sample shortly exposed (2 min) to 2-IT, C) HSA-2IT sample incubated for 1h and D) HSA-2IT sample incubated for 2h. Average deconvoluted mass of each sample was determined considering the different charge states detected.

Figure 3. Graphical representation of the number of 2-IT species coupled to HSA peptide backbone depending on the reaction time determined from intact protein analysis using MALDI-TOF MS ($n = 3$). The values were estimated by comparing the mass shift between intact HSA and HSA-2IT samples considering 2-IT theoretical mass increment (+101.03 Da).

Figure 4. Extracted ion chromatogram (XIC) corresponding to m/z ratio of peptide T08 (509.2722, 2+), peptide T08_09 (483.7656, 3+) without conjugated species and peptide T08*09 (523.2783, 3+) exhibiting the coupling of 2-IT to the lysine residue (K* annotation), obtained from the analysis of the peptide mixture originating from HSA-2IT sample incubated for 2h. Peptides conjugated with 2-IT are represented using blue bars and unmodified peptides by green bars.

Figure 5. MS/MS spectra corresponding to A) peptide T15*16 (K*YLYEIAR) and B) peptide T47*48 (LAK*TYETTLEK) illustrating the fragment identification which allowed to attribute the coupling of 2-IT to lysine lateral chain. Experimental conditions: cf. experimental section. Lysine residues coupled with 2-IT compound are designated as K*.

Figure 6. Sequence coverage obtained for A) HSA-2IT sample obtained through short exposition to 2-IT (left panel) and B) HSA-2IT sample obtained using an incubation time of 1h (right panel). UPLC-MS/MS spectra were recorded on LTQ orbitrap XL instrument (Thermo Scientific, Bremen, Germany). Lysine residues coupled with 2-IT are marked in blue. Peptides conjugated with 2-IT are represented using blue bars and unmodified peptides by green bars

Figure 7. Schematic representation of the structure of HSA illustrating A) the accessibility of the residues Lys²³³, Lys²⁴⁰ and Lys²⁶² experiencing 2-IT coupling and B) the residues Lys⁴¹³ and Lys⁵³⁴ which did not react with the chemical compound. Modified residues are represented in green and conserved residues are represented in red.