Free Radical Research

ISSN: 1071-5762 (Print) 1029-2470 (Online) Journal homepage: http://www.tandfonline.com/loi/ifra20

A method to produce fully characterized ubiquitin covalently modified by 4-hydroxy-nonenal, glyoxal, methylglyoxal, and malondialdehyde

Mara Colzani, Angela Criscuolo, Gaia Casali, Marina Carini & Giancarlo Aldini

To cite this article: Mara Colzani, Angela Criscuolo, Gaia Casali, Marina Carini & Giancarlo Aldini (2016) A method to produce fully characterized ubiquitin covalently modified by 4-hydroxy-nonenal, glyoxal, methylglyoxal, and malondialdehyde, Free Radical Research, 50:3, 328-336, DOI: 10.3109/10715762.2015.1118472

To link to this article: http://dx.doi.org/10.3109/10715762.2015.1118472

Accepted author version posted online: 10 Nov 2015.
Published online: 14 Jan 2016.

FREE RADICAL RESEARCH, 2016
VOL. 50, NO. 3, 328–336
http://dx.doi.org/10.3109/10715762.2015.1118472

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

A method to produce fully characterized ubiquitin covalently modified by 4-hydroxy-nonenal, glyoxal, methylglyoxal, and malondialdehyde

Mara Colzani, Angela Criscuolo, Gaia Casali, Marina Carini and Giancarlo Aldini

Department of Pharmaceutical Sciences, Università degli Studi di Milano, Milan, Italy

ABSTRACT

Reactive carbonyl species (RCS) and the corresponding protein adducts (advanced glycoxidation or lipoxidation end products, i.e. AGEs and ALEs) are now widely studied from different points of view, since they can be considered as biomarkers, pathogenic factors, toxic mediators and drug targets. One of the main limits of the research in this field is the lack of standardized and fully characterized AGEs and ALEs to be used for biological, toxicological, and analytical studies. In this work, we set up a procedure to prepare and fully characterize a set of AGEs and ALEs by incubating ubiquitin – a model protein selected as target for carbonylation – with four different RCS: 4-hydroxy-*trans*-2-nonenal (HNE), methylglyoxal (MGO), glyoxal (GO), and malondialdehyde (MDA). After 24 h of incubation, the extent of protein carbonylation was estimated using a recently developed quantitative strategy based on high-resolution mass spectrometry. The resulting AGEs and ALEs were fully characterized by both intact protein and bottom-up analyses in terms of: stoichiometry of the total amount of modified protein, elucidation of the structure of the RCS-deriving adducts, and localization of the RCS-modified amino acids. Each RCS exhibited different reactivity toward ubiquitin, as detected by quantifying the extent of protein modification. The order of reactivity was MGO > GO > HNE > MDA. A variety of reaction products was identified and mapped on lysine, arginine, and histidine residues of the protein. In summary, a highly standardized and reproducible method to prepare fully characterized AGEs/ALEs is here presented.

ARTICLE HISTORY
Received 17 July 2015
Revised 12 October 2015
Accepted 5 November 2015
Published online
8 December 2015

KEYWORDS
4-hydroxy-trans-2-nonenal; glyoxal; high-resolution mass spectrometry; malondialdehyde; methylglyoxal; reactive carbonyl species

Introduction

Reactive carbonyl species (RCS) generated by lipid and sugar oxidation are bioactive molecules involved in several toxicological mechanisms. The main damaging effects of RCS are due to their ability to covalently modify nucleophilic macromolecules such as proteins, whose covalent modifications induce cellular and tissue damages by different mechanisms, including protein function derangement, protein oligomerization, immunogenicity, and activation of specific receptors (e.g. RAGE) [1,2]. RCS and the corresponding protein adducts (advanced glycoxidation or lipoxidation end products, i.e. AGEs and ALEs) [3] are now widely studied at different levels and in particular as biomarkers, pathogenic factors, toxic mediators, and drug targets [1,4–6].

One of the main limits of research in this field is the lack of standardized and fully characterized AGEs and ALEs to be used for biological, toxicological, and analytical studies. This is due to the wide chemical diversity of RCS that can be formed endogenously or exogenously, resulting in the formation of a great variety of reaction products (AGEs/ALEs) that differ not only on the basis of the modifying moiety, but also on their ability to target different proteins on different amino acids. For instance, α,β-unsaturated aldehydes can react with histidine or cysteine residues through C3 and form the corresponding Michael adducts or can react with the primary amino group of lysine leading to the corresponding Schiff bases [7]. Di-aldehydes such as glyoxal (GO) can react with the side chain of lysine to yield N_{ε}-(carboxymethyl)lysine (CML) or with the side chain of arginine forming two possible dihydroxyimidazolidine intermediates that can open to give N^7-(carboxymethyl)arginine or dehydrate to yield three different imidazolone derivatives [3].

Many studies in this field utilize AGEs/ALEs-albumin of different sources (human, rat, or bovine), prepared using different reactants, including reducing sugars (glucose, ribose, fructose) or aldehydes

CONTACT Mara Colzani mara.colzani@unimi.it Department of Pharmaceutical Sciences, Universita degli Studi di Milano, Via Mangiagalli 25, Milan 20133, Italy

Supplemental data for this article can be accessed http://dx.doi.org/10.3109/10715762.2015.1118472.

(glycolaldehyde or glyoxylic acid, GO, methylglyoxal (MGO), 4-hydroxy-*trans*-2-nonenal [HNE], acrolein), and different methods in which several variables (e.g. the buffer, the incubation temperature, and the incubation time) can vary considerably [8,9]. Moreover a detailed chemical characterization of most of the so far reported albumin reaction products is lacking, as is information on the reproducibility and standardization of the procedure. Additionally, model proteins other than albumin have been used in some studies [10,11].

Hence, a highly standardized and reproducible preparation of AGEs/ALEs is required together with their full characterization in terms of stoichiometry, structure elucidation of the modifying moieties, and identification of the nucleophilic sites undergoing the modifications. Moreover the method should also be able to modulate the extent of the protein modification – a parameter that should be easily determined in quantitative or semi-quantitative terms.

We believe that the availability of a well-characterized series of AGEs/ALEs is essential for different types of studies, for instance to compare their biological effects and in more detail to establish the contribution of the RCS and the degree of the modifications on the biological/toxicological properties of the protein adduct.

In this paper, we set up a procedure to prepare and fully characterize a set of AGEs and ALEs derivatives by incubating a model protein (ubiquitin) with HNE, GO, MDA, and MGO, which were selected since they represent the most studied lipid and sugar oxidation products and they belong to different chemical classes of RCS [3].

Materials and methods

Reagents

Ultrapure water was prepared with a Milli-Q H_2O purification system (Millipore, Bedford, MA). Sodium dihydrophosphate ($NaH_2PO_4{\cdot}H_2O$), disodium phosphate ($Na_2HPO_4{\cdot}2H_2O$), ammonium bicarbonate (NH_4HCO_3), formic acid (HCOOH), sodium borohydride ($NaBH_4$), LC-MS grade H_2O/0.1% formic acid, LC-MS grade acetonitrile/0.1% formic acid, lyophilized ubiquitin from bovine erythrocytes (U6253), MGO (67028), GO (50649), malondialdehyde tetrabutylammonium salt (63287), and 2,5-hexanedione (165131) were purchased from Sigma-Aldrich (Milan, Italy).

4-Hydroxy-2-*trans*-nonenal dimethylacetal was synthesized according to the literature [12] and stored at −20 °C. For each experiment, fresh HNE was prepared starting from stored HNE-dimethylacetal, which was evaporated under nitrogen stream and hydrolyzed with 10 mM HCl, pH 3 for 1 h at room temperature in the dark, to obtain HNE. The concentration of HNE was estimated by UV absorbance at 224 nm (molar extinction coefficient = 13,750 $M^{-1} \times cm^{-1}$).

In vitro generation of protein adducts by RCS

Ubiquitin was dissolved in phosphate buffer (10 mM sodium dihydrophosphate $NaH_2PO_4{\cdot}H_2O$ buffered with 200 mM disodium phosphate $Na_2HPO_4{\cdot}2H_2O$ to reach pH 7.4) to a final concentration equal to 10 μM. Ubiquitin was incubated at 37 °C with HNE, MGO, GO, malondialdehyde (MDA) or hexane-2,5-dione at increasing concentrations (50 μM, 100 μM, 200 μM, 500 μM, 1 mM, 2 mM, and 5 mM) corresponding to protein:RCS molar ratios equal to 1:5, 1:10, 1:20, 1:50, 1:100, 1:200, and 1:500. About 10 μM ubiquitin was incubated without RCS, as control sample. The reactions were stopped after 24 h by diluting the excess of RCS by ultrafiltration, using Amicon YM3 filter units (Millipore, Milan, Italy). To do so, 70 μl of solution was diluted with 430 μl of water, loaded on the filter units, and centrifuged at 14,000 × *g* for 15 min. The eluates containing the excess reagents were discarded, while the ubiquitin retained by the filter (about 50 μl of solution) was rinsed with 450 μl of water. The eluate was discarded and the washing procedure was repeated twice. After the last centrifugation, the eluates were discarded and the filters were flipped to recover the ubiquitin solution (about 70 μl). If necessary, the sample was brought to the initial volume (70 μl) with water. Ubiquitin was either enzymatically digested or directly injected into the mass spectrometer for intact protein analysis (see below).

Intact protein analysis by microflow automated loop injection ESI-MS

The analysis of ubiquitin upon incubation with RCS in presence or absence of carbonyl quenchers was performed accordingly to a strategy already described for HNE-modified ubiquitin [13]. Briefly, 40 μl of ubiquitin solution recovered from filter units were mixed with 40 μl of denaturing solution (H_2O/CH_3CN/HCOOH; 40/60/0.2; v/v/v). About 5 μl aliquots of the samples were injected into a LTQ-Orbitrap XL mass spectrometer using an ESI source (Thermo Scientific, Milan, Italy). Samples were automatically injected by an Ultimate 3000 RSLCnano system that pumped an isocratic mobile phase (H_2O/CH_3CN/HCOOH; 70/30/0.1; v/v/v) at a constant flow rate of 10 μl/min. Sample injection and spectra acquisition were fully automated and controlled by the software Xcalibur (version 2.0.7, Thermo Scientific) and Chromeleon Xpress (Dionex, version 6.80). Source parameters were set as follows: spray voltage 1.8 kV,

capillary temperature 220 °C, capillary voltage 35 V, tube lens offset 120 V. A list of 20 background ions [14] was used as lock mass values for real-time mass calibration. Mass spectra were acquired by the Orbitrap analyzer in positive ion mode using the following settings: profile mode, scan range *m/z* 110-2000, AGC target 5×10^5, maximum inject time 500 ms, resolving power 100,000 (FWHM at *m/z* 400). A dedicated Xcalibur processing method was set to quantify the area under the 11 + multicharged peaks, localized in the *m/z* range 779.00–783.50 for unmodified ubiquitin, 793.00–797.50 for HNE-modified ubiquitin, 784.25–785.75 + 789.50–791.00 + 794.75–796.25 for GO-modified ubiquitin, 784.00–787.00 for MGO-modified ubiquitin, and 784.00–785.50 + 788.10–789.60 for MDA-modified ubiquitin.

Quantification of adducts formation by RCS by intact protein analysis

Peak areas were automatically detected and quantified post-acquisition using Xcalibur Quan Browser. The percentage of modified ubiquitin obtained upon incubation with every concentration of the different RCS was computed as:

$$\%\ \text{modified ubiquitin} = \frac{\text{AUC modified ubiquitin}}{\text{AUC unmodified ubiquitin} + \text{AUC modified ubiquitin}} \times 100$$

Protein digestion

Ubiquitin solutions recovered from the filter units (25 µl, corresponding to about 2 µg of protein) were incubated with 500 mM sodium borohydride at 37 °C, to stabilize adducts. After 1 h, 25 µl of LDS sample buffer (2 × solution, Bio-Rad, Segrate, Italy) was added to each sample. The samples were heated for 5 min at 99 °C and loaded on a Mini-PROTEAN gel (TGX, any Kd, Bio-Rad, Segrate, Italy) for SDS-PAGE. After 20 min run, the migration was stopped and the gel was stained using Coomassie blue (Bio-Safe G-250 Stain, Bio-Rad) accordingly to the manufacturer protocol. Gel bands corresponding to ubiquitin (around 8 kDa) were cut, washed with 200 µl ammonium bicarbonate (50 mM) and digested with 1 µg of trypsin or Glu-C (both Roche Diagnostics, Monza, Italy) in 50 mM ammonium bicarbonate buffer. After overnight incubation at 37 °C, peptides were extracted by 10 min incubation with extraction solution (acetonitrile/TFA/water; 30/3/67; v/v/v) and by additional 10 min incubation with 100% acetonitrile. For each sample, collected peptides were dried using a vacuum system (Christ, Germany) and solubilized in 20 µl of 0.1% formic acid for the subsequent LC-MS/MS analysis. In-gel digestion was preferred to in-solution digestion because, under our experimental conditions, it produced a higher number of observable peptides (data not shown).

NanoLC–ESI-MS/MS analysis of peptides

Digested peptides (5 µl) diluted in 0.1% formic acid were loaded on a C18 HALO PicoFrit column (75 µm × 10 cm, 2.7 µm particle size, 100Å pore size, New Objective, Woburn, MA) for reverse phase chromatography. Samples were injected by an Ultimate 3000 RSLCnano system (Dionex, CA) controlled by the software Chromeleon Xpress 6.80 (Dionex, CA) and electrosprayed using a nanoESI source equipped with a dynamic nanospray probe (both Thermo Scientific, Milan, Italy). Peptide separation was performed by a reverse phase linear gradient from 99% buffer A (water with 0.1% formic acid), 1% buffer B (acetonitrile with 0.1% formic acid) to 60% buffer A, 40% buffer B over 60 min, at flow rate of 300 nl/min. The instrument was an LTQ-Orbitrap XL (Thermo Scientific, Waltham, MA), controlled by the software Xcalibur 2.0.7 and operated in data-dependent mode to acquire both full MS and MS/MS spectra. Full MS spectra were acquired in profile mode by the Orbitrap analyzer in a scan range equal to *m/z* 300-1500, using capillary temperature 220 °C, AGC scan 5×10^5 and resolution 60,000 FWHM at *m/z* 400. Tandem mass spectra were acquired by the LTQ analyzer for the 2 most intense ions exceeding 1×10^4 counts. MS/MS spectra acquisition was set as follows: centroid mode, resolution 15,000, precursor ions isolation width = *m/z* 2, AGC target 1×10^4 and normalized collision energy 30 eV. Dynamic exclusion was set as follows: 3 repeat counts, 30 s repeat duration, 45 s of exclusion duration. Monoisotopic precursor selection was enabled; singly and unassigned charged ions were excluded from fragmentation.

Identification and localization of protein adducts

The software Proteome Discoverer (version 1.3.0.339, Thermo Scientific, Waltham, MA) was used to extract peaks from spectra and to match them to the primary sequence of bovine ubiquitin (Uniprot entry sp|P0CG53|1-76, i.e. corresponding to amino acids 1-76, forming the ubiquitin monomer). Accordingly to the enzyme used for the digestion, trypsin or Glu-C was selected as proteases, allowing a maximum of three missed cleavages. Peptide and fragment ion tolerances were set to 10 ppm and 0.5 Da, respectively. Allowed variable modifications were methionine oxidation in

combination with RCS-specific modification, as follow: HNE-deriving Michael adducts (+156.11502 and +158.13067 for its reduced form) and Schiff bases (+140.12011, reduced form) on Lys, Arg, and His; GO-deriving carboxymethyl-Arg and Lys (+58.00548) and hydro-imidazolone on Arg (+39.99492); MGO-deriving methyl-hydro-imidazolone (+54.01056), tetra-hydro-pyrimidine (+144.04226), and argpyrimidine (+80.02622) on Arg residues; MGO-deriving carboxyethyl-Arg and Lys (+72.02113); or MDA-deriving N-(2-propenal)lysine (+54.01056). As quality filter, only peptide matches with "high" confidence (FDR ≤ 0.01) were considered as genuine identifications. The MS/MS spectra of modified peptides were requested to match the expected ions (*b* and/or *y*) neighboring the modified amino acid both at the N- and C-termini, to obtain a confident mapping of the modification site.

Relative quantification of adduct's abundance

The signal deriving from RCS-modified peptides was relatively quantified, as compared to their non-modified counterpart, by calculating the AUC (area under the curve) of the extracted ion chromatograms (XIC) for the adducted and native forms of the same peptide sequence, using *m/z* values corresponding to the either $[M+2H]^{2+}$ or $[M+3H]^{3+}$, accordingly to the charge state reported by the Proteome Discoverer-based identification. The extent of peptide modification was expressed as percentage using the following formula:

$$\%\ \text{modified peptide} = \frac{\text{AUC modified peptide}}{\text{AUC unmodified peptide} + \text{AUC modified peptide}} \times 100$$

When multiple peaks were present in the XIC, the AUC was calculated for the peak at the retention time reported by Proteome Discoverer for peptide identification.

Results and discussion

An *in vitro* method to generate fully characterized AGEs and ALEs was set up, which was also able to modulate the extent of protein carbonylation. Ubiquitin was selected as model protein for several reasons: (i) it is commercially available in a pure form and at a low price; (ii) its molecular weight allows it to be analyzed in intact form by using the most common MS analyzers, thus permitting the determination of the relative content of the modified isoforms; (iii) the absence of cysteine residues in its sequence allows the observation of a variety of RCS-deriving modifications on other nucleophilic residues (i.e. lysine, arginine, and histidine), which would otherwise be masked by the high reactivity of cysteine [15]. Moreover ubiquitin also represents an interesting target of RCS-deriving modification, due to the involvement of lysine residues (e.g. Lys 48) in mono- and poly-ubiquitination, which regulate important cellular events such as protein degradation, trafficking, inflammation, and translation.

Four RCS belonging to different chemical classes were selected as modifying agents: the α,β-unsaturated-aldehyde HNE, the keto-aldehyde MGO, and the two di-aldehydes GO and MDA. These RCS were selected since they are the most abundant and studied endogenous and exogenous RCS acting as precursor of AGEs and ALEs.

Ubiquitin carbonylation by RCS: quantification of the modified forms of the protein

The first step of the work was to monitor the RCS-mediated ubiquitin carbonylation and to identify the main protein adducts by intact protein analysis. For this purpose, ubiquitin samples were incubated in the absence and presence of RCS, filtered, diluted in denaturing conditions, and analyzed by MS as intact protein. The mass spectrum of control, non-modified ubiquitin (5 μM concentration) is characterized by a typical multi-charged ion pattern, with charge states spanning from +7 to +13, with the +11 peak at *m/z* 779.61064 representing the most abundant signal, as shown in Supplementary Figure 1A. After deconvolution of the spectrum, the experimental average mass of unmodified ubiquitin is 8564.64 Da (Supplementary Figure 1B), in agreement with the expected average mass (8564.74 Da) with an experimental error of 11.6 ppm on the deconvoluted mass value.

After incubating ubiquitin for 24 h with RCS, the corresponding MS spectra presented additional multi-charged ion patterns attributed to the RCS-modified ubiquitin forms. RCS induced a dose-dependent increase of ubiquitin modification: the intensity and the number of the additional peaks increased at increasing concentrations of RCS (Figure 1 and Supplementary Figures 2–4), indicating the presence of different protein variants generated by protein carbonylation, especially at the highest concentrations of RCS.

The percentage of RCS-modified ubiquitin was determined by analysing the +11 charged ions (Supplementary Figure 1C). Figure 2 and Supplementary Table I report the percentage of modified ubiquitin obtained upon 24 h incubations with the different RCS. MGO was the most reactive RCS since it was able to modify about

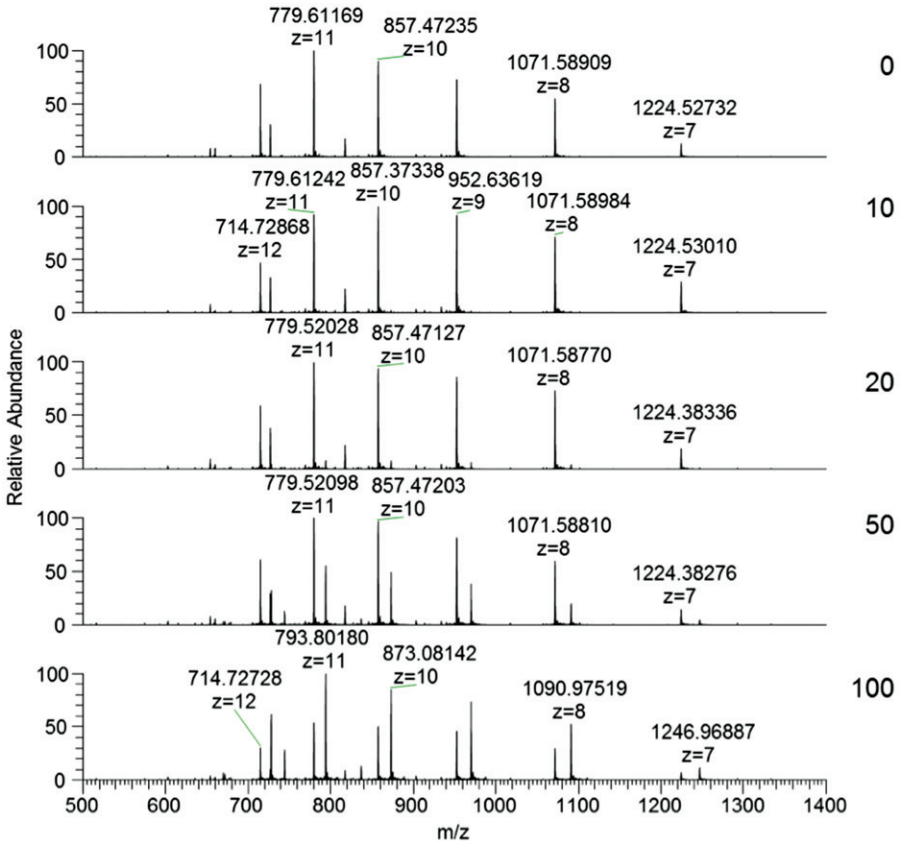

Figure 1. Mass spectra of intact ubiquitin acquired after 24 h incubation with increasing concentration of HNE. The molar excess of HNE in respect to ubiquitin concentration is indicated on the right side of the figure. Excess HNE was removed after the incubation by ultrafiltration.

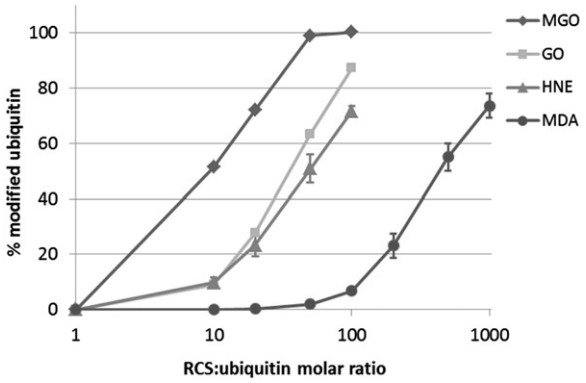

Figure 2. Percentage of modified ubiquitin obtained upon the *in vitro* incubation of the proteins (10 μM) with increasing concentration of MGO, GO, HNE, or MDA. The percentage of modified ubiquitin is plotted against the concentration of RCS (μM). Error bars represent the standard errors estimated by three independent replicate samples obtained by different incubations.

50% of ubiquitin when present at a 10-fold molar excess in respect to ubiquitin. A higher concentration (about a 50-fold excess) of HNE and GO was necessary to reach the same amount of modified ubiquitin. MDA was the least reactive RCS, since a 500-fold molar excess was necessary to obtain an appreciable amount of carbonylated protein. Hence, the order of reactivity was MGO > GO > HNE > MDA.

Identification of protein adducts by intact protein analysis

In pursuit of clarity, we focused on the +11 peak as representative of the different multi-charged ion patterns, since it represents the most intense peak in our experimental conditions, as already reported in Ref. [13]. Figure 3(A) shows the +11 peak(s) present in the mass spectra of the unmodified ubiquitin and of the samples incubated with an amount of RCS able to modify about 50% of the protein. In these mass spectra, the +11 peak corresponding to unmodified ubiquitin is detectable at nominal *m/z* 779, flanked by additional peaks at higher *m/z* values, corresponding to ubiquitin variants that underwent covalent modification by HNE, MGO, GO, and MDA. All modifications detected on ubiquitin modified by HNE, GO, MGO, or MDA are summarized in Table I.

Each RCS induced well-defined mass shifts (ΔM) in respect to native ubiquitin, as highlighted in the

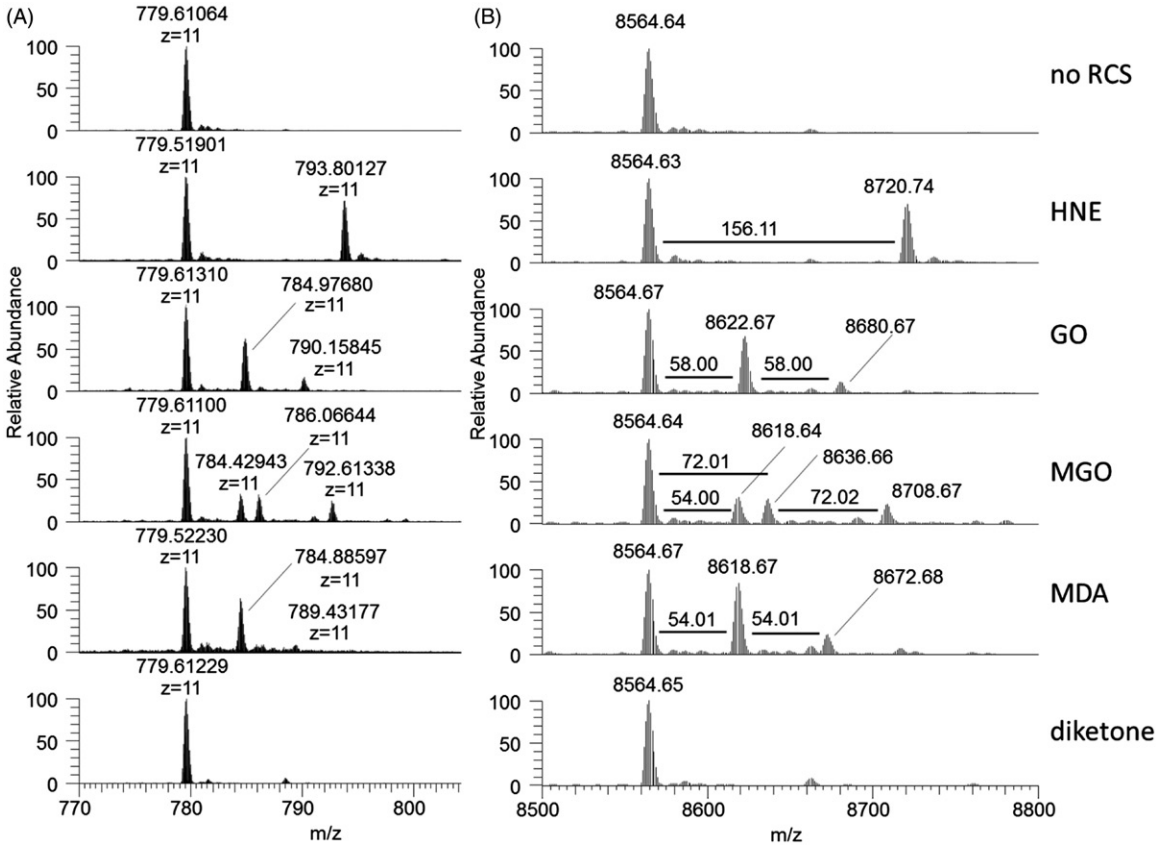

Figure 3. Mass spectra obtained upon *in vitro* incubation of ubiquitin with: phosphate buffer (control sample), glyoxal (GO), methylglyoxal (MGO), 2-hydroxy-*trans*-nonenal (HNE), malondialdehyde (MDA), and the unreactive diketone 2-5-exanedione. Spectra are reported for selected RCS concentrations that generated about 50% of modified ubiquitin. (A) Focus on the z11 peaks present in multi-charged mass spectra. (B) Deconvoluted mass spectra, showing the uncharged mass of ubiquitin in its unmodified and RCS-modified forms.

Table I. Protein adducts corresponding to the mass shifts observed on ubiquitin by MS analysis of the intact protein.

RCS	Theoretical ΔM monoisotopic	Adduct type	Residues	Ref.
HNE	156.11502	Michael adducts	KRH	[7]
GO	58.00548	Carboxymethyl	KRC	[16,17]
MGO	54.01056	Methyl-hydro-imidazolone	R	[18]
	72.02113	Carboxyethyl	KRC	[19]
	144.04226	Tetrahydropyrimidine	R	[20]
MDA	54.01056	N-propenal	K	[21]

Theoretical mass shifts and known target amino acids (described by the literature) are reported. Reactive carbonyl species legend: HNE, 4-hydroxy-*trans*-nonenal; GO, glyoxal; MGO, methylglyoxal; MDA, malondialdehyde.

deconvoluted spectra shown in Figure 3(B) and as described in the following paragraphs.

HNE-induced modifications – After incubating ubiquitin with a 50-fold molar excess of HNE, the deconvoluted mass spectrum presented a new peak at 8720.74 Da, characterized by a ΔM of 156.11 Da in respect to the native form of the protein. Consistently, the +11 peak in the original mass spectrum is shifted by 14.2 *m/z*. The new peak was attributed to the Michael adduct of HNE [7].

GO-induced modifications – The incubation of ubiquitin with 50-fold molar excess of GO induced the appearance of two new peaks characterized by ΔM of 58.00 and 116.00 Da, indicating the presence of one and two carboxymethyl modifications on the protein [16].

MGO-induced modifications – 10-fold molar excess of MGO generated three additional main peaks, showing ΔM equal to 54.00, 72.01, and 144.03 Da. These values correspond to the known adducts methyl-

Table II. Amino acids targeted by HNE, MGO, GO, and MDA on ubiquitin sequence, as detected by the analysis of peptides obtained from the digestion of the ubiquitin samples shown in Figure 3.

RCS	Adduct type	Targeted amino acids
HNE	Michael adducts	K6, K11, K33, K48, K63, H68
	Schiff base	K6
MGO	Methyl-hydro-imidazolone	R42, R54, R72, R74
	Carboxyethyl	R54, R72
	Argpyrimidine	R72
GO	Carboxymethyl	K6, R42, R54
	Hydro-imidazolone	R72
MDA	N-propenal	K6, K11, K27, K29, K33, K48, K63

hydroimidazolone [18], carboxyethyl [19] and two carboxyethyl and/or tetrahydropyrimidine [20] (isobaric modifications), respectively.

MDA-induced modifications – MDA (500-fold molar excess) generated two peaks with a fixed mass increment equal to 54.01 Da. This value corresponds to two N-(2-propenal)lysine adducts [21].

To test whether the protein adducts were artefacts generated by the ionization process or resulted from non-covalent interactions, ubiquitin was incubated with hexane-2,5-dione, a diketone which does not act as RCS. Unlike the samples incubated with HNE, GO, MGO, or MDA, no additional peaks were observed by incubating ubiquitin with hexane-2,5-dione even at the highest concentrations (Figure 3 and Supplementary Figure 5).

Localization of RCS-deriving adducts on ubiquitin sequence

In order to confirm the presence of the modifications observed in the intact protein and to map the amino acids targeted by the different RCS, aliquots of the ubiquitin samples shown in Figure 3 were analyzed using a bottom-up approach. The samples were stabilized with sodium borohydride and split into two halves that were digested using trypsin and Glu-C, respectively. The peptides resulting from the two types of digestion were analyzed by nanoLC-MS in parallel, in order to increase the sequence coverage. The identification of the RCS-modified peptides was obtained by matching the acquired spectra to the ubiquitin sequence, allowing the modifications observed on the intact proteins (Table I). Moreover, additional adducts described in the literature (listed in the Materials and Methods section) were allowed as possible peptide modification. Quality filters were applied to obtain genuine peptide identification and to unambiguously map the modified residues on the basis of fragmentation spectra. By applying these criteria, we identified different peptides bearing a variety of modifications on lysine, arginine, or histidine residues, as reported in Table II.

In general, HNE and MDA preferentially targeted lysine residues, while MGO and GO mainly targeted arginine residues. In more detail, HNE-deriving Michael adducts were localized on lysine residues K6, K11, K33, K48, and K63, and on the histidine residue H68. The HNE-deriving Schiff base was found on K6; this modification was not detected in the intact protein, likely because it is a reversible modification that can be observed only after its stabilization.

The GO-deriving carboxymethyl adducts, also identified on the intact protein, were localized on the arginine residues R42 and R54 and on lysine K6. Additionally, the GO-deriving imidazolone derivative was identified on R72.

The MGO-deriving carboxyethyl (already detected on the intact protein) was localized on R54 and R72; the methyl-hydroimidazolone adduct was observed on the same amino acids and additionally on R42 and R74. Additionally, the MGO-deriving argpyrimidine was detected in R72. Based on these results we can consider that the peak at *m/z* 792 visible in Figure 3 probably corresponds to a double carboxyethyl adduct and not to the tetrahydropyrimidil-arginine, since the latter was not identified by the bottom-up analysis. In any case, it is not possible to exclude the presence of tetrahydropyrimidil-arginine, due to its isobaricity with the double carboxyethyl adduct. The MDA-deriving adduct N-propenal was detected in all seven lysine residues present in the ubiquitin sequence.

Analysis of the preferentially modified amino acids

The intensities of the signals corresponding to the unmodified and RCS-modified peptides were quantified, in order to identify the amino acids preferentially modified by each RCS. To do so, the area under the curve (AUC) of the ion chromatograms extracted for the modified and unmodified forms of the peptide (XIC) were determined and the results reported as a percentage of the modified peptide.

Supplementary Table II lists, for each RCS-modified peptide, the modification site and its relative abundance in respect to the corresponding unmodified form. Peptides originating from Glu-C digestion were preferentially considered in respect to those generated by trypsin, since RCS-modified lysine and arginine appeared to affect the efficiency of trypsin digestion on those residues. As a consequence, in the sample digested with trypsin, peptides modified on their C-terminus were not detected in both their unmodified and modified forms,

Figure 4. Localization of the amino acids preferentially modified by the four RCS. The percentage of modification for each amino acid is color-coded, accordingly to the legend. Adducts legend: MA, Michael addition; SB, Schiff base; MGH, MGO-deriving methyl-hydroimidazolone; CE, carboxyethyl; Arg-pyr, argpyrimidine; CM, carboxymethyl; GH, GO-deriving hydroimidazolone; NPK, N-propenal-lysine.

preventing relative quantification. Peptides were quantified as tryptic peptides only when the modifications involved an internal amino acid and when they were not quantified as Glu-C-deriving peptides.

The preferential sites of modification are summarized in Figure 4. H68 represents the amino acid preferentially targeted by HNE (Michael addition). GO preferentially generated the carboxymethyl adduct on R42. MGO was able to react with different amino acids and to form different adducts, but overall it preferentially generated the methyl-hydroimidazolone adducts on arginine R42, R72, and R74; R72 was also modified at relatively high amounts as carboxyethyl-arginine. MDA preferentially modified the residues K6 and K29 (N-propenal-lysine).

The adducted moieties detected in these experiments are in agreement with previous observations; all of them have already been reported as major AGEs/ALEs *in vitro* or *in vivo* [7,10,13,16,17,21–25]. No modification of lysine residues was detected for MGO, unlike previous reports [19]. Moreover, cross-linked forms of ubiquitin (e.g. MOLD, GOLD, Lys-Lys enamine) were not detected in the acquired spectra. However, it is possible that protein aggregates were present in some samples, but not detected by our instrumental setup due to their high mass values. Notably, protein precipitation was not observed in any sample and ubiquitin multimers were not observed by SDS-PAGE analysis of aliquots of different ubiquitin samples incubated with RCS (data not shown), indicating that protein aggregation was either absent or rare in our experimental conditions.

Conclusions

Several pieces of evidence indicate that both endogenous and exogenous AGEs and ALEs are involved in the pathogenic mechanisms of different diseases and that they are important modulators of toxic responses.

We believe that a great limit to research in this field is the lack of well-characterized standards allowing efficient analyses of the structure–activity relationship. For instance, the role of each specific RCS on the biological activity of the resulting AGEs/ALEs is not well defined yet, as well as the impact of the extent of protein modification on cellular and tissue toxicity. The characterization of the different types of AGEs/ALEs is critical not only to better understand the biological effects of RCS, but also to identify novel drug targets.

The method reported here can be easily adapted to other proteins to create other well-characterized AGEs and ALEs. As example of their application, standardized AGEs and ALEs can be tested as ligand of the receptor for AGEs (RAGE). This would allow comparison of the affinity of the RAGE receptor for the same protein bearing different RCS-deriving modifications, thus allowing the detection of the moiety of AGEs/ALEs responsible of the ligand:receptor interactions – an information that is currently missing. The complete structure characterization provides a solid background to study the structure–activity relationship and to evaluate the biological activity of the modified moieties arising from the condensation of the different RCS with a target protein.

In conclusion, we report an *in vitro* method for the generation of fully characterized AGEs and ALEs, where the carbonylation extent can be modulated and determined. The method is highly reproducible and represents an important tool to reach a better understanding of the structure–activity relationship of AGEs and ALEs.

Acknowledgements

G.A. acknowledges the colleagues of the COST action whose fruitful discussion led to the conception of this paper. M.C. received a post-doctoral fellowship (Dote Ricerca: FSE, Regione Lombardia) enabling her to carry the present investigation.

Declaration of interest

The authors declare no conflicts of interest.

References

[1] Dalle-Donne I, Giustarini D, Colombo R, Rossi R, Milzani A. Protein carbonylation in human diseases. Trends Mol Med 2003;9:169–176.

[2] Barlovic DP, Soro-Paavonen A, Jandeleit-Dahm KAM. RAGE biology, atherosclerosis and diabetes. Clin Sci (Lond) 2011;121:43–55.

[3] Vistoli G, Maddis De D, Cipak A, Zarkovic N, Carini M, Aldini G. Advanced glycoxidation and lipoxidation end products (AGEs and ALEs): an overview of their mechanisms of formation. Free Radic Res 2013;47:3–27.

[4] Januszewski AS, Alderson NL, Metz TO, Thorpe SR, Baynes JW. Role of lipids in chemical modification of proteins and development of complications in diabetes. Biochem Soc Trans 2003;31:1413–1416.

[5] Aldini G, Dalle-Donne I, Colombo R, Maffei Facino R, Milzani A, Carini M. Lipoxidation-derived reactive carbonyl species as potential drug targets in preventing protein carbonylation and related cellular dysfunction. ChemMedChem 2006;1:1045–1058.

[6] Aldini G, Vistoli G, Stefek M, Chondrogianni N, Grune T, Sereikaite J, et al. Molecular strategies to prevent, inhibit and degrade advanced glycoxidation and advanced lipoxidation end products. Free Radic Res 2013;47:93–137.

[7] Sayre LM, Lin D, Yuan Q, Zhu X, Tang X. Protein adducts generated from products of lipid oxidation: focus on HNE and one. Drug Metab Rev 2006;38:651–675.

[8] Valencia JV, Mone M, Koehne C, Rediske J, Hughes TE. Binding of receptor for advanced glycation end products (RAGE) ligands is not sufficient to induce inflammatory signals: lack of activity of endotoxin-free albumin-derived advanced glycation end products. Diabetologia 2004;47:844–852.

[9] Ott C, Jacobs K, Haucke E, Navarrete Santos A, Grune T, Simm A. Role of advanced glycation end products in cellular signaling. Redox Biol 2014;2:411–429.

[10] Cotham WE, Metz TO, Ferguson PL, Brock JWC, Hinton DJS, Thorpe SR, et al. Proteomic analysis of arginine adducts on glyoxal-modified ribonuclease. Mol Cell Proteomics 2004;3:1145–1153.

[11] Uchida K, Stadtman ER. Modification of histidine residues in proteins by reaction with 4-hydroxynonenal. Proc Natl Acad Sci USA 1992;89:4544–4548.

[12] Rees MS, van Kuijk FJGM, Siakotos AN, Mundy BP. Improved synthesis of various isotope labeled 4-hydroxyalkenals and peroxidation intermediates. Synth Commun 1995;25:3225–3236.

[13] Colzani M, Criscuolo A, De Maddis D, Garzon D, Yeum K-J, Vistoli G, et al. A novel high resolution MS approach for the screening of 4-hydroxy-trans-2-nonenal sequestering agents. J Pharm Biomed Anal 2014;91:108–118.

[14] Keller BO, Sui J, Young AB, Whittal RM. Interferences and contaminants encountered in modern mass spectrometry. Anal Chim Acta 2008;627:71–81.

[15] Schaur R. Basic aspects of the biochemical reactivity of 4-hydroxynonenal. Mol Aspects Med 2003;24:149–159.

[16] Reddy S, Bichler J, Wells-Knecht KJ, Thorpe SR, Baynes JW. N epsilon-(carboxymethyl)lysine is a dominant advanced glycation end product (AGE) antigen in tissue proteins. Biochemistry 1995;34:10872–10878.

[17] Iijima K, Murata M, Takahara H, Irie S, Fujimoto D. Identification of N(omega)-carboxymethylarginine as a novel acid-labile advanced glycation end product in collagen. Biochem J 2000;347:23–27.

[18] Lo TWC, Westwood ME, Mclellan AC, Selwood T, Thornalleys PJ. Binding and modification of proteins by methylglyoxal under physiological conditions. A kinetic and mechanistic study with N alpha-acetylarginine, N alpha-acetylcysteine, and N alpha-acetyllysine, and bovine serum albumin. J Biol Chem 1994;269:32299–32305.

[19] Ahmed MU, Brinkmann Frye E, Degenhardt TP, Thorpe SR, Baynes JW. N-epsilon-(carboxyethyl)lysine, a product of the chemical modification of proteins by methylglyoxal, increases with age in human lens proteins. Biochem J 1997;324:565–570.

[20] Oya T, Hattori N, Mizuno Y, Miyata S, Maeda S, Osawa T, Uchida K. Methylglyoxal modification of protein: chemical and immunochemical characterization of methylglyoxal-arginine adducts. J Biol Chem 1999;274:18492–18502.

[21] Uchida K, Sakai K, Itakura K, Osawa T, Toyokuni S. Protein modification by lipid peroxidation products: formation of malondialdehyde-derived N(epsilon)-(2-propenal)lysine in proteins. Arch Biochem Biophys 1997;346:45–52.

[22] Thornalley PJ, Battah S, Ahmed N, Karachalias N, Agalou S, Babaei-Jadidi R, Dawnay A. Quantitative screening of advanced glycation endproducts in cellular and extracellular proteins by tandem mass spectrometry. Biochem J 2003;375:581–592.

[23] Henle T, Walter AW, Haeßner R, Klostermeyer H. Detection and identification of a protein-bound imidazolone resulting from the reaction of arginine residues and methylglyoxal. Z Lebensm Unters Forsch 1994;199:55–58.

[24] Klöpfer A, Spanneberg R, Glomb MA. Formation of arginine modifications in a model system of Nα-tert-butoxycarbonyl (Boc)-arginine with methylglyoxal. J Agric Food Chem 2011;59:394–401.

[25] Shipanova IN, Glomb MA, Nagaraj RH. Protein modification by methylglyoxal: chemical nature and synthetic mechanism of a major fluorescent adduct. Arch Biochem Biophys 1997;344:29–36.