Published in final edited form as:
J Proteome Res. 2007 June ; 6(6): 2323–2330. doi:10.1021/pr070112q.

Enrichment and Analysis of Non-enzymatically Glycated Peptides: Boronate Affinity Chromatography Coupled with Electron Transfer Dissociation Mass Spectrometry

Qibin Zhang[¶], **Ning Tang**[§], **Jonathan W. C. Brock**[⊥,‡], **Heather M. Mottaz**[¶], **Jennifer M. Ames**[∉], **John W. Baynes**[⊥], **Richard D. Smith**[¶], and **Thomas O. Metz**[¶,*]

[¶] *Biological Sciences Division and Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352* [§] *Life Science & Chemical Analysis, Agilent Technologies, Santa Clara, CA 95052* [⊥] *Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208* [∉] *School of Biological Sciences, Queen's University Belfast, Belfast BT9 5AG, Northern Ireland*

Abstract

Non-enzymatic glycation of peptides and proteins by D-glucose has important implications in the pathogenesis of diabetes mellitus, particularly in the development of diabetic complications. However, no effective high-throughput methods exist for identifying proteins containing this low abundance post-translational modification in bottom-up proteomic studies. In this report, phenylboronate affinity chromatography was used in a two-step enrichment scheme to selectively isolate first glycated proteins and then glycated, tryptic peptides from human serum glycated *in vitro*. Enriched peptides were subsequently analyzed by alternating electron transfer dissociation (ETD) and collision induced dissociation (CID) tandem mass spectrometry. ETD fragmentation mode permitted identification of a significantly higher number of glycated peptides (87.6% of all identified peptides) versus CID mode (17.0% of all identified peptides), when utilizing enrichment on first the protein and then the peptide level. This study illustrates that phenylboronate affinity chromatography coupled with LC-MS/MS and using ETD as the fragmentation mode is an efficient approach for analysis of glycated proteins and may have broad application in studies of diabetes mellitus.

Keywords

non-enzymatic glycation; boronate affinity enrichment; electron transfer dissociation; collision-induced dissociation; post-translational modification; liquid chromatography; mass spectrometry

Introduction

Non-enzymatic glycosylation, or glycation, is the covalent binding of single reducing sugars (glucose, fructose, ribose, etc.) to primary amino groups in proteins[1] and is distinct from glycosyl transferase-catalyzed glycosylation. The latter plays important roles in immunology and cell biology,[2] where oligosaccharides are attached to specific protein side chains, such as

* To whom correspondence should be addressed. Dr. Thomas O. Metz, Biological Sciences Division, Pacific Northwest National Laboratory, P.O. Box 999, MSIN: K8-98, Richland, WA 99352. Phone: (509) 376-8333 Fax: (509) 376-2303. E-mail: thomas.metz@pnl.gov.
‡Current address: Department of Biochemistry, Vanderbilt University, Nashville, TN 37232

asparagine (N-linked), serine and threonine (O-linked), and the C-termini of cell surface proteins (GPI lipid anchored). The initial product of a non-enzymatic glycation reaction is a labile Schiff base intermediate, which slowly isomerizes to form a stable ketoamine, called the Amadori compound (AC) (Figure 1a).[1] The AC can then undergo further oxidation and rearrangement reactions to form a series of more reactive, colored and fluorescent compounds, termed advanced glycation end-products (AGEs),[3] which are believed to play a pathogenic role in the development of diabetic complications. The Maillard hypothesis of diabetic complications proposes that chronic, cumulative chemical modification of proteins by glycation and AGEs alters their structure, function, and turnover.[4] Indeed, tissue AGE levels increase in long-lived proteins of diabetic patients[5, 6] versus control individuals and correlate with the severity of both diabetic nephropathy and retinopathy.[7] AGEs are also thought to play a role in other age-related chronic diseases, such as atherosclerosis,[8] through effects on vascular elasticity and thickening of vascular walls. They have also been proposed to play a role in abnormal amyloid aggregation in age-related neurodegenerative disorders such as Alzheimer's[9] and Parkinson's[10] diseases.

Glycated hemoglobin (HbA1c) has been used as a clinical diagnostic marker of relatively long-term (4–6 weeks) glucose control in diabetic patients since the 1970s.[11] In order to find more sensitive and informative protein biomarkers for monitoring of glycemic control and to more fully understand the role of glycation and AGEs in the development of diabetic complications and other disorders, comprehensive proteomic studies are required to identify those glycated proteins whose altered structure may contribute to pathology. Previous assays for protein glycation include various chemical reactions on *cis*-diol functional groups of the AC and subsequent dye labeling methods for spectroscopic detection of ACs on intact proteins or in protein hydrolysates;[12, 13] however the specificity of these methods remains poor, and they can only provide information on the overall level of glycation. Important glycation site information is not obtainable from intact proteins and is lost during sample processing of protein hydrolysates. Because the total level of glycated proteins *in vivo* is only a few percent,[14] it is necessary to have an enrichment step to enable efficient identification of these low level post-translational modifications (PTMs) prior to MS analysis in bottom-up proteomics. Enrichment methods for glycosylated proteins and peptides have employed affinity enrichment with lectin columns.[15] Alternatively, a hydrazide affinity matrix may be used to capture glycosylated peptides following periodate oxidation; captured peptides can then be released using glycosidase treatment, e.g. with PNGase F.[16, 17] However, these methods are not applicable to the AC modification, because the binding selectivity of lectins relies on specific conformations of oligosaccharides, and periodate fully oxidizes the secondary hydroxyl groups of the AC to carboxylic acids. Alternatively, incomplete oxidation of the AC moiety can lead to aldehyde formation under mild conditions. The aldehydes may then efficiently couple to the hydrazide resin; however, no suitable enzymes or chemical methods are available to release the attached peptides for subsequent MS analysis. Fortunately, separation-based assays at the intact protein level have been developed for enrichment of glycated proteins,[13] primarily by phenylboronate affinity chromatography (BAC), which relies on the strong coordination between *cis*-diols and boronate ion.[14, 18–20] Yet, until recently, it has not been evaluated for enrichment of glycated proteins in complex samples[21] and has been rarely used in the enrichment of glycated peptides.[22]

An additional challenge in the analysis of glycated proteins is identification of the glycation site by mass spectrometry. Peptide mapping methods,[23–25] which use the mass of enzymatically digested peptides coupled with *in silico* digestion of proteins, cannot be applied to complex proteome-level glycation analyses because identifications of the glycation sites are ambiguous and based solely on monosaccharide modification-induced mass increases of the parent peptide. Although tandem mass spectrometry with collision-induced dissociation (CID) has been applied to glycated peptide sequencing,[26] high abundance ions corresponding to

various degrees of neutral water loss dominate the MS/MS spectra and very limited and weak peptide backbone fragmentation is observed. While certain patterns of neutral losses can hint at the presence of a glycated peptide,[26, 27] information leading to a confident identification of that peptide sequence is lacking. Precursor-ion scanning methods based on the Amadori-derived lysine immonium ion at *m/z* 192.1 were recently used to map glycation sites using quadrupole-time-of-flight (Q-TOF) MS,[28] however, this method has some limitations for broad applications in bottom up proteomics, due to the low mass cut off limitation of all commercial 3-D or linear ion trap instruments.

Recently, electron transfer dissociation (ETD) fragmentation using a modified linear ion-trap was developed by Hunt and co-workers.[29] This technique is analogous to electron capture dissociation (ECD) implemented on Fourier transform ion cyclotron resonance mass spectrometers[30] but uses aromatic anions as an electron source. During ETD fragmentation, bond dissociation occurs immediately after electron transfer, providing more extensive sequence information while labile modifications remain intact. ETD is particularly well-suited for the characterization of peptides containing PTMs[29, 31] and has been successfully applied in phosphopeptide analyses;[29] abundant peptide backbone c and z type ions were detected, resulting in almost complete sequence coverage in the ETD fragmentation spectra. Recently, we have used ETD in analyses of glycated peptides,[32] demonstrating that ETD is superior to CID for glycated peptide analysis; almost complete sequence coverage was observed in ETD MS/MS spectra, while various neutral loss peaks dominated the CID MS/MS spectra without appreciable sequence information.[32]

In this work, we report the first use of phenylboronate affinity chromatography to enrich glycated proteins and peptides from a complex protein sample, followed by analysis of the enriched glycated, tryptic peptides utilizing data-dependent tandem mass spectrometry with alternating ETD and CID MS/MS.

EXPERIMENTAL

Chemicals and materials

All chemicals, standard glycosylated proteins, *in vitro* glycated human serum, ribonuclease A (RNase; type II-A), and glucose assay kits were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise stated. Micro-BCA protein assay kits and ICON™ 9K concentrators were purchased from Pierce (Rockford, IL). Sequencing-grade trypsin was purchased from Promega (Madison, WI). A Tricorn™ high performance liquid chromatography column (5 mm ×100 mm) was purchased from GE Healthcare; Glycogel™ II boronate affinity gel (Pierce, Rockford, IL) was a gift from Dr. Bart Haigh of the Institute for Bioanalytics (Branford, CT), and Affi-gel™ 601 boronate affinity gel was purchased from Bio-Rad Laboratories (Hercules, CA).

Non-enzymatic glycation of RNase

RNase was glycated with D-glucose under anaerobic conditions as previously described.[23] Briefly, RNase (13.7 mg, 1 μmol) was dissolved in 1 mL of a solution of glucose (0.4 M) in phosphate buffer (0.2 M, pH 7.4) made in 0.2 μm filtered deionized water. The solution was then divided into two aliquots and one drop of toluene was added to maintain sterile conditions; anaerobic conditions were maintained through the addition of 1 mM diethylenetriaminepentaacetic acid and nitrogen atmosphere. The solutions were then incubated at 37°C for 3 and 14 days. Incubations were terminated by storage at −20°C.

Boronate affinity chromatography

The empty Tricorn chromatography column (5 mm ×100 mm) was packed with a slurry of either Glycogel™ II or Affi-gel™ 601 boronate affinity gel under gravity flow. Prior to affinity fractionation, free glucose remaining in standard protein and glycated human serum samples was removed with ICON concentrators to avoid competition for boronate binding sites; the completeness of glucose removal was determined by micro-glucose assay. Multiple levels of boronate affinity enrichment were evaluated: (1) enrichment of glycated proteins only, (2) enrichment of glycated peptides only, and (3) enrichment of first glycated proteins followed by enzymatic digestion and subsequent enrichment of glycated peptides. Control samples consisted of protein samples that were not subjected to enrichment on any level. Glycated proteins or glycated peptides (up to 0.5 mg) were dissolved in 500 μl buffer A and injected onto the packed boronate affinity column using an Agilent (Agilent Technologies, Santa Clara, CA) 1100 series LC system equipped with a fraction collector. The LC solvents were (A) 50 mM $MgCl_2$, 250 mM NH_4OAc, pH 8.1 and (B) 0.1 M HOAc, and a gradient (0% B for 10 min; 0–100% B in 0.1 min; 100% B for 10 min; 100–0%B in 0.1 min; 0%B for 10 min) was used to separate glycated from non-glycated proteins/peptides at a flow rate of 1.0 ml/min. In addition, we evaluated the effect of reducing the LC flow rate (from 1 ml/min to 0.15 ml/min during the washing step) on the binding of glycated peptides; similarly, the effect of completely stopping the LC flow for 5–20 minutes after sample injection was evaluated. The LC effluent was monitored at 280 nm with a UV detector. The percentage of glycated proteins or peptides was calculated as the peak area corresponding to bound proteins or peptides divided by the sum of the peak areas corresponding to bound and unbound proteins. Proteins from both the bound and flow-through fractions were collected, concentrated, and desalted with ICON concentrators prior to tryptic digestion, while peptides from both the bound and flow-through fractions were collected, concentrated by speed-vac (Thermo Savant, Milford, MA), and subsequently desalted with C_{18} SPE cartridges.

Protein Digestion

Proteins (10 mg/ml) collected from the bound and flow-through fractions of boronate affinity enrichment or proteins not subjected to enrichment were dissolved in 8 M urea/100 mM NH_4HCO_3 (pH 8.2) and reduced with 5 mM dithiothreitol (DTT) for 1 hr at 37°C; free sulfhydryl groups were alkylated with 20 mM iodoacetamide at room temperature for 1 hr in the dark. Samples were then diluted with 50 mM NH_4HCO_3 (pH 8.0) to reduce the urea concentration to below 1 M, and $CaCl_2$ was added to a final concentration of 1.5 mM prior to the addition of sequencing-grade trypsin at a ratio of 1:40 (w/w, enzyme: protein). Samples were digested at 37°C for 9 h. The final digestion mixture was passed through C_{18} SPE cartridges for desalting and eluted peptide solutions were concentrated by speed-vac before further analysis.

LC- MS/MS analyses of peptides

For LC-MS/MS analyses, an Agilent 6340 Ion Trap LC/MS system equipped with ETD capability was coupled to an Agilent HPLC-Chip Cube MS interface, well plate sampler, nanoflow pump, and capillary pump.[33] Fluoranthene anions from a small negative chemical ionization source were used to transfer electrons in the ETD process. To identify eluting peptides, the MS was operated in a data-dependent MS/MS mode (300–2200 *m/z*), in which a full MS scan was followed by three CID and ETD scans each (performed alternately). A dynamic exclusion window of 0.5 min was used to discriminate against ions previously selected for fragmentation in two sequential scans. For CID fragmentation, the smart fragmentation feature was enabled, and the fragmentation amplitude was 1.3 V with a start amplitude of 30% and end amplitude of 200%; for ETD fragmentation, reactant accumulation time was 40 ms and reaction time was 80 ms. The HPLC separation was carried out on the Chip with an

integrated separation column (Zorbax 300SB-C18, 75 nm ×150 mm, 5 μm) and an enrichment column (Zorbax 300SB-C18, 40 nL, 5 μm). The voltage applied to the HPLC-Chip capillary was 1725 V, and the LC solvents were (A) 0.1% formic acid in H_2O and (B) 0.1% formic acid in 90% CH_3CN/10% H_2O. For analysis of RNase samples, 2 μL (0.8 μg) of each sample were injected, and gradient separation (3–43% B in 20 min; 43–90% B in 5 min; hold at 90% B for 5 min) was performed at a flow rate of 300 nL/min; for analysis of serum samples, 2 μL (2.0 μg) of each sample was injected, and gradient separation (3–43% B in 80 min; 43–90% B in 15 min; hold at 90% B for 5 min) was performed at the same flow rate.

Data Analysis

Spectrum Mill MS Proteomics Workbench Rev A.03.03.069 (Agilent Technologies) was used for peptide identifications by searching LC-MS/MS data obtained in both CID and ETD fragmentation modes against either the combined RNase FASTA file and a yeast database with 5870 protein entries or the human International Protein Index (IPI) database (Version 3.05 consisting of 49,161 protein entries; available online at www.ebi.ac.uk/IPI) for RNase and human serum samples, respectively. The minimum matched peak intensity was set to 2%, with a precursor ion mass tolerance of ± 2.5 Da and a product ion mass tolerance of ± 0.7 Da; a sequence tag length >1 was also specified. A parameter file incorporating the following mass modifications was used in the search for glycated peptides: dynamic modification for cysteine residues (carbamidomethylation, 57.02 Da), dynamic modification for methionine residues (oxidation, 16.00 Da), and dynamic modification for lysine residues (AC, 162.05 Da). The following criteria were used to filter raw search results: score >8.0, %SPI >60, and Delta Rank1-Rank2 >2.5. Additionally, each dataset was also searched against the same FASTA or IPI database with all protein sequences reversed in order to estimate the peptide false discovery rate (FDR), calculated as the # of false peptide identifications divided by the total # of peptide identifications (from both forward and reverse protein sequence database searches).

RESULTS

BAC enrichment of glycated model proteins

Affinity chromatography is a powerful technique for the separation of biomacromolecules based on their biological activity or chemical structure. We chose to use Glycogel™ II boronate affinity gel for isolation of glycated proteins and peptides because of its high molecular weight cut off limit.[34, 35] This boronate affinity gel is comprised of *m*-aminophenylboronic acid covalently attached to an insoluble, inert agarose matrix (Figure 1b); the boronic acid moiety binds to *cis*-diol groups on glycated species under alkaline conditions to form a reversible five-member ring complex (Figure 1c). When a sample containing a mixture of glycated and non-glycated proteins is applied to the boronate affinity support, the glycated proteins are retained on the gel whereas the non-glycated species will be washed free. The bound glycated proteins can then be eluted either by lowering the buffer pH (Figure 1c) with 0.1 M acetic acid or by using a buffer containing a high concentration of sorbitol, which has higher binding affinity to boronate than glycated proteins.

RNase has long been regarded as a model protein for studies on glycation; [23, 36] it is a small, stable, monomeric protein with 10 ε-amino groups from lysine residues and 1 α-amino group. In its native form in the pancreas, there is negligible glycation on any of these primary amino group sites. Glycated RNase samples were prepared by incubating with D-glucose under anaerobic conditions for 3 and 14 days;[23] anaerobic conditions were chosen to prevent further oxidation of the AC moiety to advanced glycation end-products.[23] When the glycated RNase was fractionated via BAC (Figure 2a), RNase incubated with glucose for 14 days had a higher percentage of binding fraction (87.1%) compared with RNase incubated with glucose for only 3 days (60.0%). This is in agreement with published data on the kinetics of glycation of RNase.

[23, 36] In contrast, when native RNase (i.e. no incubation with glucose) was applied to the same column, only 0.6% of total protein (presumably glycated) was retained on the column; the majority of the protein was non-glycated and appeared in the wash buffer. While this observation is intriguing, the relatively low amount of binding for native RNase did not warrant characterization of this fraction by LC-MS/MS. Importantly, boronate affinity enrichment on the protein level exhibited high reproducibility, with a CV of 0.26% observed for the percentage of binding fraction during replicate enrichment (n = 4) of the same sample.

In principle, every compound containing *cis*-diol groups can bind to the boronate column. Therefore, we further evaluated the binding specificity of the boronate column using two standard enzymatically glycosylated proteins, α-1-antichymotrypsin and α-2-hs-glycoprotein, which contain 6 N-linked glycans and 2 N-linked and 3 O-linked glycans, respectively. These two proteins have been previously used as standards for enrichment of glycosylated proteins with periodate/hydrazide chemistry.[16] As shown in Figure 2a, only a small fraction (1.5% and 1.2% respectively) of the glycoproteins were retained on the boronate column. This may be due to the presence of sialic acids at the terminus of the oligosaccharides, which may not contain *cis*-diols due to O-acetylation. However, the sialic acids also carry a negative charge which, under alkaline conditions, may lead to electrostatic repulsion from the boronate resin. Although mannose and galactose residues in the middle of the oligosaccharide chain contain some *cis*-diol groups, combined electrostatic repulsion and steric hindrance likely prevent the binding of these oligosaccharide chains to the boronate resin. The separation of enzymatically and nonenzymatically glycosylated proteins by boronate affinity resin has also been observed by Gould *et al.*[18, 19]

BAC enrichment on peptide level

Although glycated proteins can be selectively enriched by boronate affinity chromatography, the binding efficiency is typically less than 100%, as with any affinity chromatography method; the binding fraction also inevitably contains a small fraction of non-glycated proteins. Furthermore, because glycation generally occurs only on N-terminal and some side chain amino groups of proteins, peptides generated by enzymatic digestion prior to MS analysis increase the sample complexity and counteract the benefits of enrichment at the protein level. For high-throughput LC-MS/MS analyses of glycated proteins, it is necessary to further enrich the glycated peptides from tryptic digests of enriched glycated proteins.

Therefore, the bound fraction corresponding to RNase incubated with glucose for 14 days was digested with trypsin, and the resulting tryptic peptide mixture was reapplied to the same Glycogel™ II column. As expected, only a portion (14%) of the total peptides was retained (Figure 2b); enrichment on the peptide level exhibited comparable reproducibility to enrichment on the protein level. We also evaluated Affi-gel™ 601 boronate affinity gel for enrichment of glycated peptides. This material has a molecular weight cut-off of 6000 Da and is reported to be selective for *cis*-diol-containing low molecular weight compounds.[37] However, we observed loss of glycated peptides prior to gradient elution, resulting in difficulties in the collection of bound fractions (data not shown). Therefore, only the Glycogel™ II column was used for enrichment of both glycated proteins and glycated peptides.

To further characterize the composition of flow-through and binding fractions, both fractions from the enrichment of RNase glycated with glucose for 14 days were collected, digested with trypsin, and subjected to LC-MS/MS analysis; similarly, both fractions from a parallel peptide level enrichment were collected and desalted with C_{18} SPE cartridges prior to LC-MS/MS analysis. All glycated peptides were identified in the binding fractions, and no glycated peptides were identified in the corresponding flow-through fractions from either the protein or the peptide level enrichments (data not shown). The unique glycated peptides identified in MS/MS analyses in ETD mode are shown in Table 1. The results are consistent with the previous

work of Brock *et al.*,[23] although some additional lysine residues, such as K^{31} and K^{91}, were found to be glycated in the current study. In addition, two glycation sites were not observed in our study, the N-terminus and K^{37}. Glycation at the N-terminus was not reported by the Spectrum Mill software, due to a limitation of the current algorithm regarding modifications on N-termini. The peptide containing glycation at K^{37} (6 residues in length) was likely not observed due to poor retention on the C_{18} SPE cartridges used for cleaning tryptic digests. The discrepancies in glycated peptide identifications between the binding fractions of protein level and peptide level enrichments may result from differences in sample complexity, i.e. the presence of non-glycated peptides after digestion of glycated proteins will lead to ionization suppression and undersampling during data-dependent MS/MS analyses. Nonetheless, the identification of glycated RNase peptides in only the binding fractions of both glycated protein and peptide enrichments clearly demonstrates the efficiency of BAC and is promising for extending its application to more complex protein mixtures.

BAC enrichment of glycated proteins and peptides from human serum glycated *in vitro*

Human serum glycated *in vitro* (Sigma-Aldrich, St. Louis, MO) was subjected to affinity chromatography for the enrichment of glycated proteins. As shown in Figure 3a, the percentage of glycated proteins was ~28%. The obtained glycated proteins were desalted and digested with trypsin, and the resulting peptide mixture was further enriched for glycated peptides. Approximately 5.0% of total peptides were estimated to be glycated (Figure 3b). For comparison, a tryptic digest of a human serum sample glycated *in vitro* that was not enriched on the protein level was also passed through the affinity separation column. As expected, the recovery of glycated peptides (1.2%) was much less (Figure 3b) than from samples of pre-enriched glycated proteins.

Glycated peptide analysis by ETD and CID MS/MS

Several different enriched glycated peptide samples were prepared from human serum glycated *in vitro* and subjected to LC-MS/MS analysis under alternating CID/ETD MS/MS to determine which fragmentation technique is more suitable for analysis of glycated peptides using bottom-up proteomics, i.e. we wished to specifically demonstrate the utility of ETD MS/MS in the analysis of, glycated tryptic peptides, as has been recently demonstrated for model glycated peptides.[32]

The resulting MS/MS data were analyzed by Spectrum Mill software, and the raw results were filtered as described in the Experimental section. The number of spectra containing identified peptides is summarized in Table 2 (the list of 88 unique identified glycated peptides and corresponding 27 unique proteins is shown in the Supporting Information). In ETD mode, 87.6% of all peptides were identified as glycated in samples subjected to dual enrichment, i.e. enrichment on the protein level followed by enrichment on the peptide level; in comparison, 76.4% of all peptides were identified as glycated in samples enriched on only the peptide level. Only 12.3% of all peptides were identified as glycated in samples enriched on only the protein level. In digests of human serum glycated *in vitro* and without any level of enrichment, only 6.4% of all peptides were identified as glycated. This data illustrates the utility of dual enrichment on both the protein and peptide level for the efficient identification of glycated peptides. Although the same trend of a higher number of spectra containing identified glycated peptides in dual enriched samples was observed during CID MS/MS, far fewer glycated peptides were identified overall compared to ETD MS/MS as shown in Table 2. This is illustrated in Figure 4, where typical MS/MS spectra of a tryptic, glycated peptide from α-1-antitrypsin precursor obtained under both CID and ETD are shown. As expected under ETD conditions, almost a complete sequence of c and z ions are produced (top frame), and the glucose modified lysine can be unambiguously identified; however, peaks corresponding to various degrees of neutral water loss dominate the CID spectrum (bottom frame). There are

no detectable b and y ions to provide sequence information for this peptide. This observation is in line with previous results obtained for ETD analysis of model glycated peptides.[32] It is important to note that the majority of glycated peptides identified in ETD analyses of human serum glycated *in vitro* were mapped to highly abundant serum proteins, such as albumin, α-1-antitrypsin, apolipoproteins AI and AII, complement C3 precursor, serotransferrin precursor, etc. (see Supporting Information) This is not unexpected, as the 22 most abundant proteins represent ~99% of the total protein mass in serum.[38] Therefore, immunodepletion of human serum samples should be utilized prior to boronate affinity enrichment of glycated proteins and peptides, in order to increase the coverage of low-abundance glycated serum proteins that may serve as biomarkers of developing diabetic complications.

The same datasets were also searched against a reversed protein sequence database, and the raw results were filtered under the same criteria. As shown in Table 2, the false positive discovery rate (FDR) for data obtained under ETD mode is lower (median FDR of 1.9% for 7 datasets) than under CID mode (median FDR of 5.6% for 7 datasets); falsely identified glycated peptides in CID mode contribute significantly to the overall number of false identifications. The much improved FDR obtained in ETD mode for glycated peptides is due to the higher quality spectra obtained from better peptide backbone fragmentation.

The flow-through fractions from enrichment at both the protein level and the peptide level were also subjected to LC-MS/MS analysis after appropriate sample treatment. The results (Table 2) show that there are almost no glycated peptides identified in any of the flow-through fractions, especially in the case of dual enrichment and enrichment only on peptide level.

Discussion

The binding efficiency of glycated proteins to boronate columns is affected by the pH of the binding buffer and the binding buffer ionic strength. The buffers used in our separation represent the optimized buffer system as previously described,[34, 35] so no further optimization was performed in this study. However, it was also reported that the binding efficiency could be affected by the contact time between species containing *cis*-diols and the boronate column.[34] Therefore, we tested the separation by decreasing the flow rate in the washing step from 1 ml/min to 0.15 ml/min, as well as stopping the flow after sample injection for 5–20 minutes for incubation. However, there was no detectable difference in the ratio of binding/flow-through fractions, which showed that the binding between the *cis*-diol groups and the boronate anions under alkaline conditions is rapid. Therefore, increasing the contact time between the column and samples is unnecessary for enrichment of glycated species.

Our results showed that dual enrichment on both the protein and the peptide level is the most efficient method for enrichment of glycated peptides. However, the amount of glycated proteins collected during protein level enrichment from samples glycated *in vivo* might be very limited for further sample processing (eg. desalting, digestion etc.), because the *in vivo* glycation level is much lower. Thus, further sample processing will lead to unavoidable sample loss, which will result in reduced recovery during subsequent enrichment on the peptide level. Thus, when the initial sample amount is limited and only one level of glycation enrichment can be afforded, peptide level enrichment is preferred over enrichment on the protein level. This is supported by the data shown in Table 2, where a single enrichment on the peptide level resulted in 76.4% of MS/MS spectra containing glycated peptides. In contrast, a single enrichment on the protein level resulted in 12.3% of spectra containing glycated peptides, which reflects the counter enrichment effect mentioned earlier for analysis of peptides obtained from enrichment performed only on protein level.

ETD and CID MS/MS scans were performed alternately in this study, which aims to provide a fair comparison between these two fragmentation techniques; the same parent ions were used for both ETD and CID analyses. The results unequivocally demonstrate that ETD is very effective for glycated peptide analysis in bottom-up proteomic studies. However, it is generally believed that the ETD mode is less efficient than CID because the electron transfer process results in significant ion losses (observed ETD fragment ion intensities are only 20–30% compared to CID[32]). Therefore, ETD may not be as sensitive as CID in characterizing low abundance proteins. This was not an issue in our study, since ETD yielded significantly more identifications compared with CID for glycated peptides, with 156 spectra containing glycated peptides in ETD mode compared to 15 spectra containing glycated peptides in CID mode for a dual enrichment sample (Table 2). An additional advantage of ETD is the confidence of the results; in general, the scores of the identified glycated peptides were much higher than that obtained under CID. This feature is reflected in the almost complete sequence coverage in MS/MS spectra obtained under ETD. However, with respect to peptides with stable PTMs, ETD may not offer an advantage, likely due to the relatively low sensitivity of ETD versus CID.[32] This may be less of an issue when longer peptides resulting from proteolysis with enzymes having more restricted cleavage sites are analyzed.

Finally, the relatively low number of unique glycated peptides identified in our study is the result of alternating CID/ETD scans. When only ETD fragmentation mode is employed for the entire LC-MS/MS experiment, there likely will be more glycated peptides identified.

CONCLUSIONS

We have demonstrated that for bottom-up proteomics, dual enrichment with boronate affinity chromatography on both the protein and peptide level is the most efficient method for enriching glycated peptides. When limited sample amounts are available, enrichment on only the peptide level is more efficient than on the protein level. A complete comparison of ETD and CID fragmentation modes in a bottom-up proteomic study clearly demonstrated that ETD can identify more glycated peptides with higher confidence and can significantly improve the data-dependent tandem mass spectrometric analysis of glycated peptides. Boronate affinity chromatography and ETD MS/MS analyses are currently being applied in our laboratory in studies of glycated proteins and peptides from plasma and erythrocytes of individuals with pre-diabetes and type 2 diabetes mellitus.

Acknowledgements

The authors would like to thank Dr. Bart Haigh of the Institute for Bioanalytics for kindly providing the Glycogel™ II boronate affinity gel. In addition, the authors thank Drs. Konstantinos Petritis and Matthew E. Monroe of Pacific Northwest National Laboratory (PNNL) and David Horn of Agilent Technologies for assisting with Spectrum Mill Software. This research was supported by NIH grant 5R21DK071283 to R.D.S. (PI) and T.O.M. (co-PI), NIH grant DK-19971 to J.W.B, and The Wellcome Trust via a Short-Term Travel Grant No. 06564 to J.M.A.; portions of this research were also supported through the National Center for Research Resources (RR018522) and performed at the Environmental Molecular Science Laboratory, a national scientific user facility located at PNNL and sponsored by the U. S. Department of Energy (DOE) Office of Biological and Environmental Research. PNNL is operated by Battelle for the DOE under Contract No. DE-AC06-76RLO-1830.

References

1. Baynes JW, Watkins NG, Fisher CI, Hull CJ, Patrick JS, Ahmed MU, Dunn JA, Thorpe SR. The Amadori product on protein: structure and reactions. Prog Clin Biol Res 1989;304:43–67. [PubMed: 2675036]
2. Lehle L, Strahl S, Tanner W. Protein glycosylation, conserved from yeast to man: a model organism helps elucidate congenital human diseases. Angew Chem Int Ed Engl 2006;45(41):6802–18. [PubMed: 17024709]

3. Ahmed N, Thornalley PJ. Quantitative screening of protein biomarkers of early glycation, advanced glycation, oxidation and nitrosation in cellular and extracellular proteins by tandem mass spectrometry multiple reaction monitoring. Biochem Soc Trans 2003;31(Pt 6):1417–22. [PubMed: 14641078]
4. Thorpe SR, Baynes JW. Role of the Maillard reaction in diabetes mellitus and diseases of aging. Drugs Aging 1996;9(2):69–77. [PubMed: 8820792]
5. Mosier MA, Occhipinti JR, Burstein NL. Autofluorescence of the crystalline lens in diabetes. Arch Ophthalmol 1986;104(9):1340–3. [PubMed: 3753285]
6. Lyons TJ, Silvestri G, Dunn JA, Dyer DG, Baynes JW. Role of glycation in modification of lens crystallins in diabetic and nondiabetic senile cataracts. Diabetes 1991;40(8):1010–1015. [PubMed: 1907246]
7. McCance DR, Dyer DG, Dunn JA, Bailie KE, Thorpe SR, Baynes JW, Lyons TJ. Maillard reaction products and their relation to complications in insulin-dependent diabetes mellitus. J Clin Invest 1993;91(6):2470–2478. [PubMed: 8514859]
8. Baynes JW. The role of AGEs in aging: causation or correlation. Exp Gerontol 2001;36(9):1527–37. [PubMed: 11525875]
9. Nicolls MR. The clinical and biological relationship between Type II diabetes mellitus and Alzheimer's disease. Curr Alzheimer Res 2004;1(1):47–54. [PubMed: 15975085]
10. Munch G, Gerlach M, Sian J, Wong A, Riederer P. Advanced glycation end products in neurodegeneration: more than early markers of oxidative stress? Ann Neurol 1998;44(3 Suppl 1):S85–8. [PubMed: 9749578]
11. Krishnamurti U, Steffes MW. Glycohemoglobin: a primary predictor of the development or reversal of complications of diabetes mellitus. Clin Chem 2001;47(7):1157–65. [PubMed: 11427445]
12. Armbruster DA. Fructosamine: structure, analysis, and clinical usefulness. Clin Chem 1987;33(12): 2153–63. [PubMed: 3319287]
13. Furth AJ. Methods for assaying nonenzymatic glycosylation. Anal Biochem 1988;175(2):347–60. [PubMed: 3071181]
14. Danze PM, Tarjoman A, Rousseaux J, Fossati P, Dautrevaux M. Evidence for an increased glycation of IgG in diabetic patients. Clin Chim Acta 1987;166(2–3):143–53. [PubMed: 3621595]
15. Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, Hirabayashi J, Kasai K, Takahashi N, Isobe T. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. Nat Biotechnol 2003;21(6):667–72. [PubMed: 12754521]
16. Zhang H, Li XJ, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat Biotechnol 2003;21 (6):660–6. [PubMed: 12754519]
17. Sun B, Ranish JA, Utleg AG, White JT, Yan X, Lin B, Hood L. Shotgun glycopeptide capture approach coupled with mass spectrometry for comprehensive glycoproteomics. Mol Cell Proteomics 2007;6 (1):141–9. [PubMed: 17074749]
18. Gould BJ, Hall PM. m-Aminophenylboronate affinity ligands distinguish between nonenzymically glycosylated proteins and glycoproteins. Clin Chim Acta 1987;163(2):225–30. [PubMed: 3568426]
19. Gould BJ, Hall PM, Cook JG. A sensitive method for the measurement of glycosylated plasma proteins using affinity chromatography. Ann Clin Biochem 1984;21(Pt 1):16–21. [PubMed: 6712132]
20. Takahashi M, Lu YB, Myint T, Fujii J, Wada Y, Taniguchi N. In vivo glycation of aldehyde reductase, a major 3-deoxyglucosone reducing enzyme: identification of glycation sites. Biochemistry 1995;34 (4):1433–8. [PubMed: 7827091]
21. Gontarev S, Shmanai V, Frey SK, Kvach M, Schweigert FJ. Application of phenylboronic acid modified hydrogel affinity chips for high-throughput mass spectrometric analysis of glycated proteins. Rapid Commun Mass Spectrom 2007;21(1):1–6. [PubMed: 17125156]
22. Adachi T, Ohta H, Hayashi K, Hirano K, Marklund SL. The site of nonenzymic glycation of human extracellular-superoxide dismutase in vitro. Free Radic Biol Med 1992;13(3):205–10. [PubMed: 1505778]
23. Brock JW, Hinton DJ, Cotham WE, Metz TO, Thorpe SR, Baynes JW, Ames JM. Proteomic analysis of the site specificity of glycation and carboxymethylation of ribonuclease. J Proteome Res 2003;2 (5):506–13. [PubMed: 14582647]

24. Lapolla A, Fedele D, Reitano R, Bonfante L, Guizzo M, Seraglia R, Tubaro M, Traldi P. Mass spectrometric study of in vivo production of advanced glycation end-products/peptides. J Mass Spectrom 2005;40(7):969–72. [PubMed: 15768376]
25. Ames JM. Application of semiquantitative proteomics techniques to the maillard reaction. Ann N Y Acad Sci 2005;1043:225–35. [PubMed: 16037243]
26. Lapolla A, Fedele D, Reitano R, Arico NC, Seraglia R, Traldi P, Marotta E, Tonani R. Enzymatic digestion and mass spectrometry in the study of advanced glycation end products/peptides. J Am Soc Mass Spectrom 2004;15(4):496–509. [PubMed: 15047055]
27. Mennella C, Visciano M, Napolitano A, Del Castillo MD, Fogliano V. Glycation of lysine-containing dipeptides. J Pept Sci 2006;12(4):291–6. [PubMed: 16180244]
28. Frolov A, Hoffmann P, Hoffmann R. Fragmentation behavior of glycated peptides derived from D-glucose, D-fructose and D-ribose in tandem mass spectrometry. J Mass Spectrom 2006;41(11):1459–69. [PubMed: 17063450]
29. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci U S A 2004;101(26):9528–33. [PubMed: 15210983]
30. Zubarev RA, Kelleher N, McLafferty FW. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. J Am Chem Soc 1998;120(13):3265–3266.
31. Hogan JM, Pitteri SJ, Chrisman PA, McLuckey SA. Complementary structural information from a tryptic N-linked glycopeptide via electron transfer ion/ion reactions and collision-induced dissociation. J Proteome Res 2005;4(2):628–32. [PubMed: 15822944]
32. Zhang Q, Frolov A, Tang N, Hoffmann R, Van de Goor T, Metz TO, Smith RD. Application of electron transfer dissociation mass spectrometry in analyses of non-enzymatically glycated peptides. Rapid Commun Mass Spectrom 2007;21:661–666. [PubMed: 17279487]
33. Tang N, Goodley P. Characterization of Protein Phosphorlation Using HPLC-Chip Electron Transfer Dissociation Ion Trap Mass Spectrometry. Agilent Application Notes. 2006
34. Glycogel II Boronate Affinity Gel. Pierce Biotechnologies Inc Instructions, 20201.
35. Litchfield J, Letsinger J, Haigh B, Guralski D, Carrington AL, Cannan S, Sredy J. Non-enzymatic glycation of plasma and cellular proteins in a short term STZ-diabetic rat model. Diabetologia 1998;41(S1):A301.
36. Watkins NG, Thorpe SR, Baynes JW. Glycation of amino groups in protein. Studies on the specificity of modification of RNase by glucose. J Biol Chem 1985;260(19):10629–36. [PubMed: 4030761]
37. Bio-Rad, Afffi-Gel 601 Boronate Affinity Gel. US/EG Bulletin 1066.
38. Qian WJ, Jacobs JM, Liu T, Camp DG 2nd, Smith RD. Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. Mol Cell Proteomics 2006;5(10):1727–44. [PubMed: 16887931]

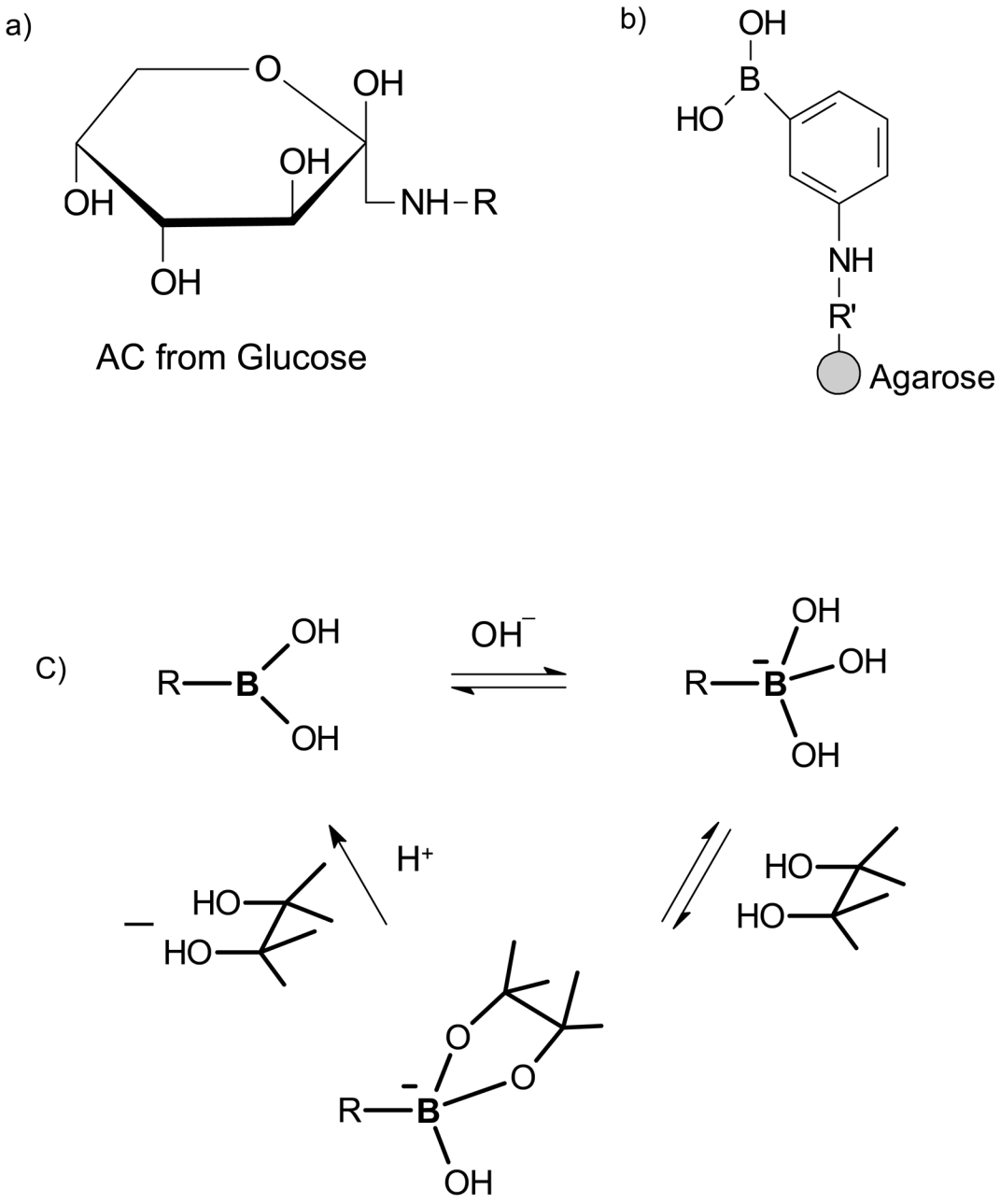

Figure 1.

(a) Structure of the Amadori compound (AC) resulting from reaction of primary amino groups with D-glucose; (b) structure of *m*-aminophenylboronic affinity gel, which is covalently bound to agarose as the solid matrix; and (c) the equilibria between boronic acids and *cis*-diol-containing compounds. Affinity attachment, elution and regeneration of boronic acid are shown.

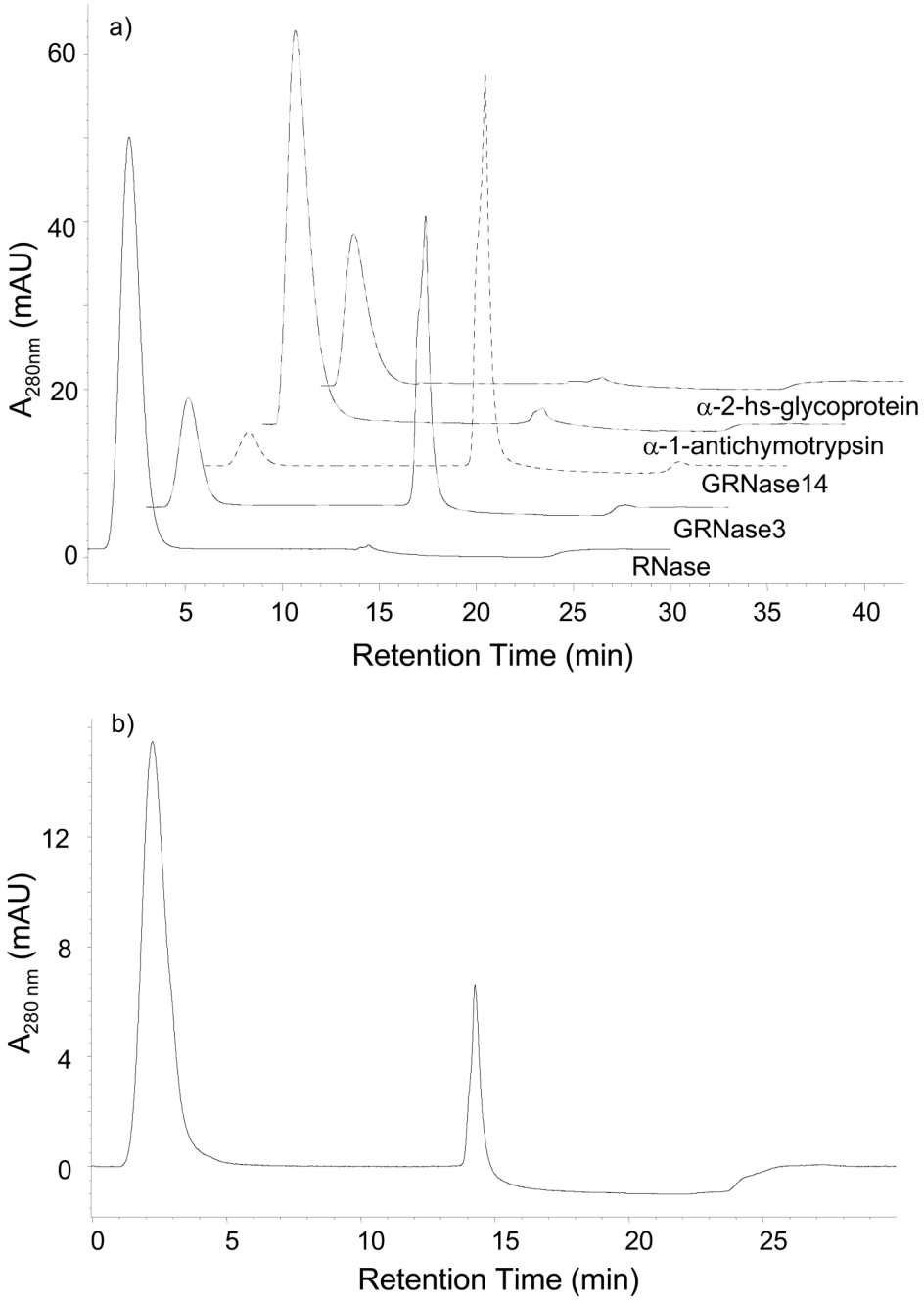

Figure 2.

(a) Chromatograms of different standard proteins passing through the Glycogel™ II boronate affinity column. The traces are labeled with the respective sample name, and the separation conditions are shown in the experimental section. Peaks at approximately 2 min are non-glycated proteins, and glycated proteins elute near 15 min; (b) chromatogram showing the enrichment of glycated peptides from a tryptic digest of enriched glycated RNase. The glycated peptides elute slightly before 15 min, and the non-glycated peptides are washed out near 2 min.

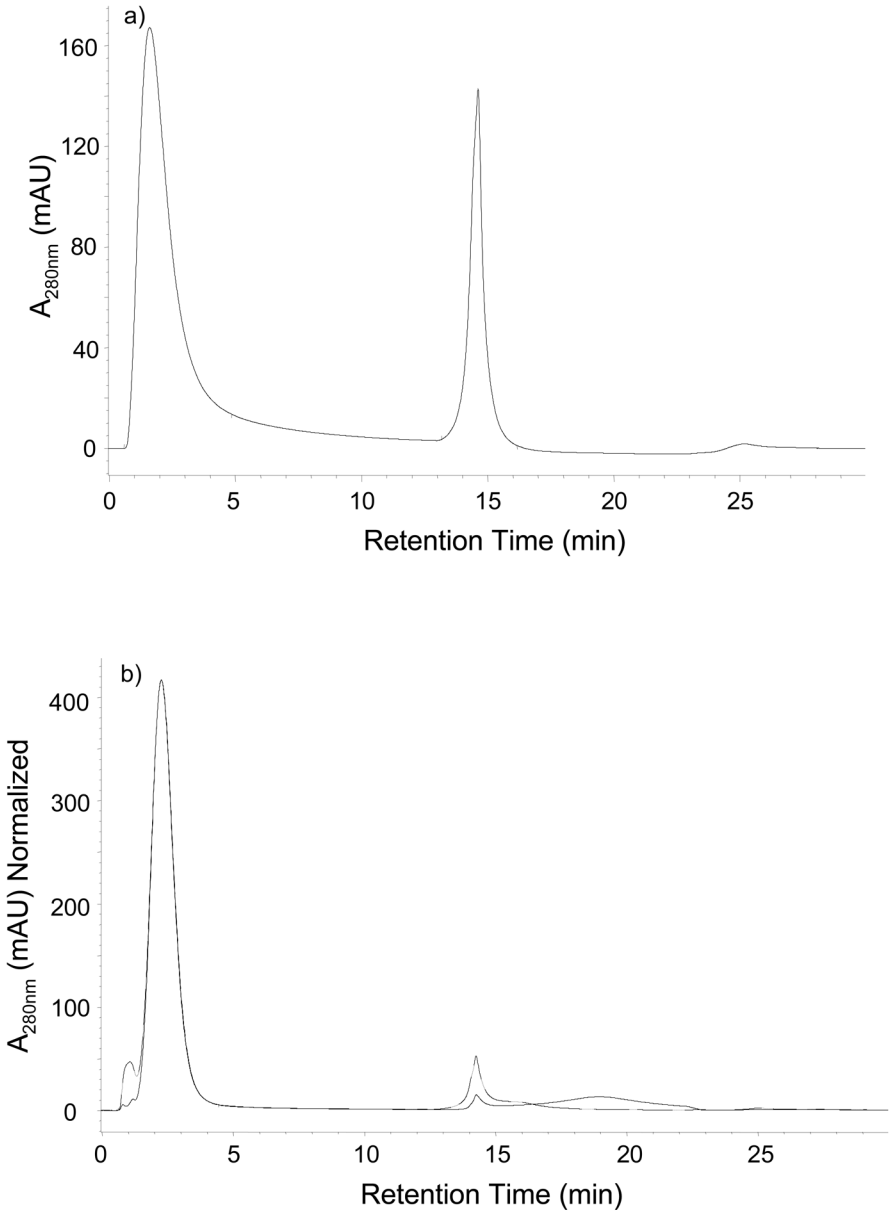

Figure 3.

(a) Chromatogram showing the enrichment of glycated proteins from human serum glycated *in vitro*; (b) chromatogram showing the enrichment of glycated peptides from tryptic digest of glycated proteins (dotted line) obtained in (a) and non-enriched glycated human serum (solid line). The separation conditions are shown in the experimental section. The chromatograms were normalized based on the most intense peak in each trace, i.e. the peak corresponding to the flow-through fraction.

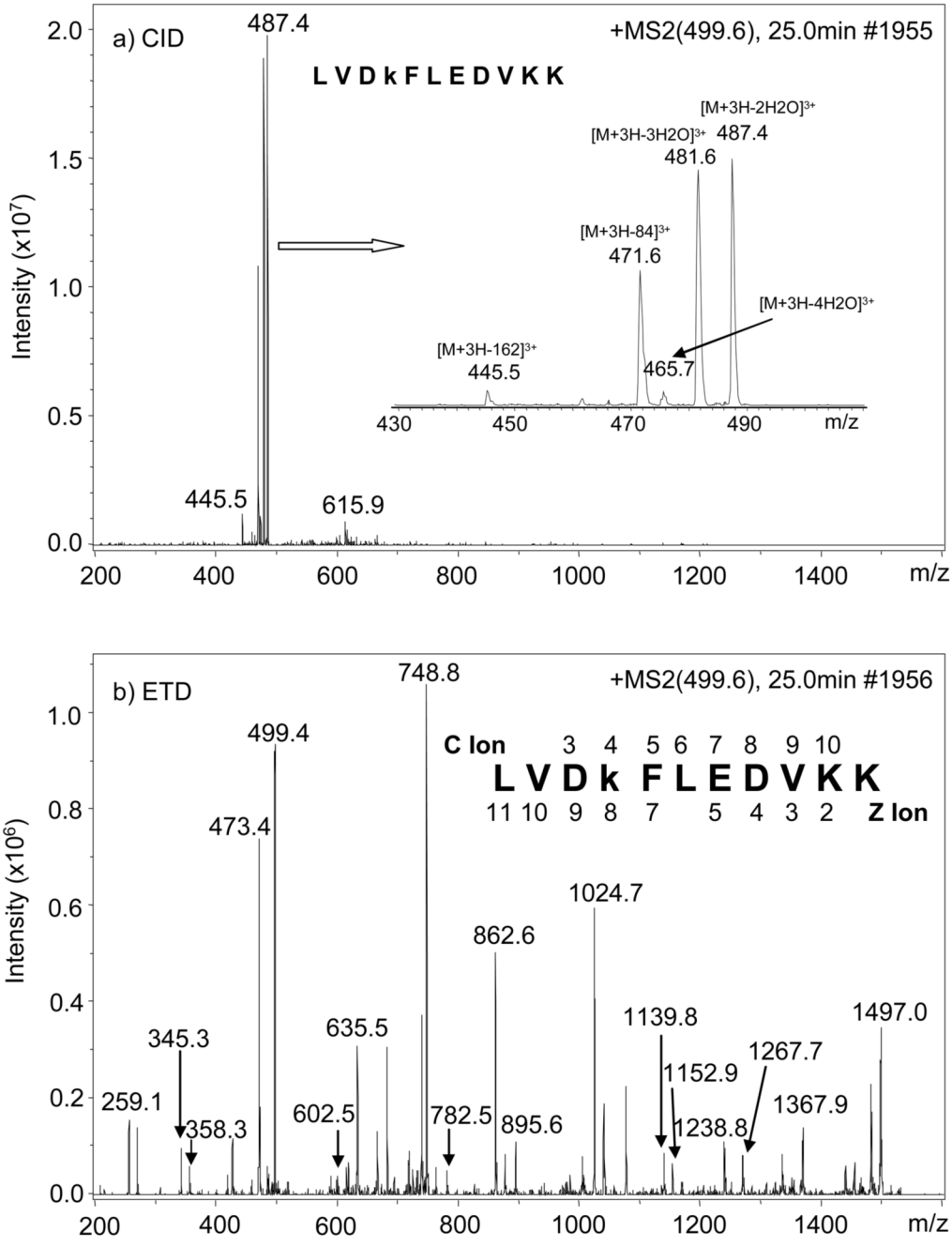

Figure 4.

MS/MS spectra obtained under CID (a) and ETD (b) fragmentation modes respectively of *m/z* 499.6, the $[M + 3H]^{3+}$ of peptide LVDkFLEDVKK from α-1-antitrypsin precursor, in which "k" represents lysine modification with glucose. Inset in (a) is the zoom in view of the ions between *m/z* 445.5 and *m/z* 487.4; the identified c and z ions were labeled above and below the sequence in (b).

Table 1

The unique glycated peptides identified from enrichment of *in-vitro* glycated RNase at protein and peptide level. The lower case letter represents the modified amino acids as stated in the Experimental section. The amino and carboxyl terminal amino acids in each peptide sequence are in parenthesis.

Peptide Sequence	Glycation Site	Identified from Protein Enrichment	Identified from Peptide Enrichment
(-)KETAAAkFER(Q)	K^{7}	Yes	Yes
(R)QHMDSSTSAASSSNYcNQMMkSR(N)	K^{31}	Yes	Yes
(R)QHmDSSTSAASSSNYcNQMMkSR(N)	K^{31}		Yes
(R)ckPVNTFVHESLADVQAVcSQK(N)	K^{41}	Yes	Yes
(R)ETGSSkYPNcAYkTTQANK(H)	K^{91}, K^{98}	Yes	Yes
(R)ETGSSKYPNcAYkTTQANK(H)	K^{98}		Yes
(K)YPNcAYkTTQANK(H)	K^{98}	Yes	Yes
(K)TTQANkHIIVAcEGNPYVPVHFDASV(-)	K^{104}	Yes	Yes

Table 2

Summary of peptide identifications with Spectrum Mill software for different human serum peptide samples after LC-MS/MS analyses under alternating CID and ETD fragmentation modes. The total number of peptide spectra identified and spectra containing glycated peptides and their ratios are listed, as well as the FDR. The filtering criteria are described under Experimental. Numbers in parentheses are the peptide spectra identified from a reversed protein sequence database search utilizing the same filter criteria as the forward database search.

Sample Name	CID total	CID-gly	CID-gly/CID total	ETD Total	ETD-gly	ETD-gly/ETD total	FDR-CID	FDR-ETD
Dual Enrichment: Binding Fraction	88 (11)	15 (3)	17.0%	178 (5)	156 (2)	87.6%	11.1%	2.7%
Peptide Enrichment: Binding Fraction	63 (3)	10 (2)	15.9%	157 (7)	120 (1)	76.4%	4.5%	4.3%
Protein Enrichment: Binding Fraction	351 (21)	8 (7)	2.3%	252 (11)	31 (5)	12.3%	5.6%	4.2%
No Enrichment	386 (23)	3 (6)	0.8%	264 (3)	17 (0)	6.4%	5.6%	1.1%
Dual Enrichment: Flow-through Fraction	450 (29)	1 (8)	0.2%	206 (4)	0 (1)	0.0%	6.1%	1.9%
Peptide Enrichment: Flow-through Fraction	396 (22)	2 (4)	0.5%	226 (3)	0 (0)	0.0%	5.3%	1.3%
Protein Enrichment: Flow-through Fraction	356 (24)	4 (2)	1.1%	253 (4)	5 (0)	2.0%	6.3%	1.6%