analytical chemistry

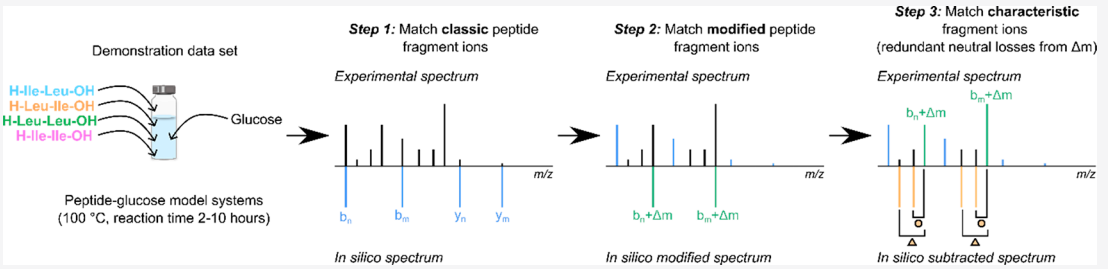
pubs.acs.org/ac Article

Open Search of Peptide Glycation Products from Tandem Mass Spectra

Michelle T. Berger,* Daniel Hemmler, Philippe Diederich, Michael Rychlik, James W. Marshall, and Philippe Schmitt-Kopplin*

Cite This: *Anal. Chem.* 2022, 94, 5953−5961 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Identification of chemically modified peptides in mass spectrometry (MS)-based glycation studies is a crucial yet challenging task. There is a need to establish a mode for matching tandem mass spectrometry (MS/MS) data, allowing for both known and unknown peptide glycation modifications. We present an open search approach that uses classic and modified peptide fragment ions. The latter are shifted by the mass delta of the modification. Both provide key structural information that can be used to assess the peptide core structure of the glycation product. We also leverage redundant neutral losses from the modification side chain, introducing a third ion class for matching referred to as characteristic fragment ions. We demonstrate that peptide glycation product MS/MS spectra contain multidimensional information and that most often, more than half of the spectral information is ignored if no attempt is made to use a multi-step matching algorithm. Compared to regular and/or modified peptide ion matching, our triple-ion strategy significantly increased the median interpretable fraction of the glycation product MS/MS spectra. For reference, we apply our approach for Amadori product characterization and identify all established diagnostic ions automatically. We further show how this method effectively applies the open search concept and allows for optimized elucidation of unknown structures by presenting two hitherto undescribed peptide glycation modifications with a delta mass of 102.0311 and 268.1768 Da. We characterize their fragmentation signature by integration with isotopically labeled glycation products, which provides high validity for non-targeted structure identification.

INTRODUCTION

Non-enzymatic glycation of amino compounds, also known as the Maillard reaction (MR), is a hallmark of food quality, metabolic stress, disease, and aging.[1−4] After Amadori product formation by spontaneous attachment of a reducing sugar to amino or guanidino groups and subsequent rearrangement, a series of loosely understood condensation and rearrangement steps lead to a diverse set of modifications known as advanced glycation end products (AGEs).[5,6] The reaction conditions including the nature of the amino compound reactant greatly impact the type of reactions, intermediate and end products of glycation reactions.[7,8] This makes the MR certainly one of the most complex reaction networks producing a multitude of reaction products only from a few initial precursors.

Due to its high specificity, sensitivity, and speed, mass spectrometry (MS) has been a cornerstone of glycation product analysis for several decades.[9−12] Many methods have been developed for the analysis of glycation products.[13−15] High-resolution MS techniques enable simultaneous detection and identification of both early and advanced glycation products even in complex mixtures.[16−18] This makes high-resolution MS particularly suitable for non-targeted glycation studies.[19] The highly accurate precursor mass and fragmentation behavior are powerful features for structure elucidation and diagnostic fragment identification.[20−23] Challenges in non-targeted analysis including the lack of commercial standards and the diversity of structures can be tackled using isotope labeling.[24−26]

Yet, peptide glycation product tandem mass (MS/MS) spectra contain multidimensional structural information, and manual interpretation of MS/MS spectra is time-consuming and can be cumbersome. Peptides harboring the same type of

Received: January 24, 2022
Accepted: March 25, 2022
Published: April 7, 2022

https://doi.org/10.1021/acs.analchem.2c00388
Anal. Chem. 2022, 94, 5953−5961

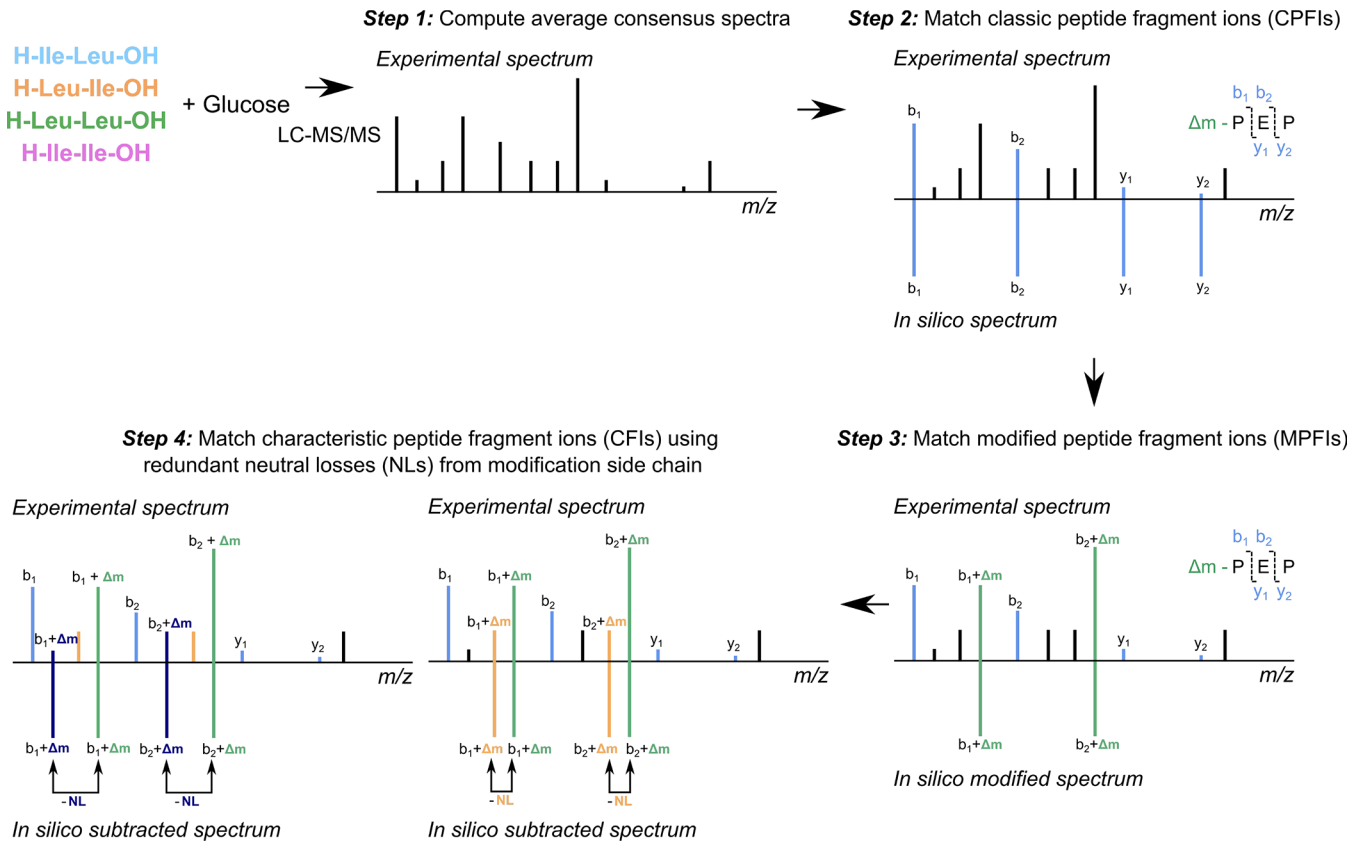


Figure 1. Open peptide glycation product search strategy step-by-step workflow. Step 1, consensus spectra computation of glycation products from four standard peptide model systems. Step 2, *in silico* digestion of input peptides provides a set of theoretical peptide fragment ions. Comparison of a CPFI catalogue (theoretical fragments and relevant amino acid fragmentation) against experimental spectra. The exact mass of the chemical modification (Δm) is calculated by subtraction of the corresponding peptide from the precursor mass ($\Delta m = \text{mass}_{\text{glycation product}} - \text{mass}_{\text{peptide matched}}$). Step 3, each theoretical fragment is subtracted from each experimental peak. Fragments yielding Δm after subtraction are matched as MPFIs. Step 4, each theoretical fragment is subtracted from each experimental peak. Fragments yielding redundant Δm_n values due to characteristic neutral losses from the modification side chain are identified as CFIs.

chemical modification do not show common absolute fragments and are eluted at different retention times. This even impedes the integration and structure alignment of well-known glycation products across peptides. It is also important to consider that glycation reaction product mixtures are complex, which makes comprehensive structure determination a challenging task. Even given substantial improvements in the quantity and quality of MS/MS data acquired on modern mass spectrometers, a vast diversity of non-enzymatic chemical modifications has remained unidentified.

In shotgun proteomics, several computational strategies have been developed to identify modified peptides, including ion indexing methods.[27−31] Non-targeted peptide glycation product search lacks computational strategies. In conventional open searching of peptide chemical modifications, regular peptide fragment ions are used for peak matching. For peptide glycation products, only a fraction of the experimental spectrum can be matched to all theoretical peptide fragment ions when considering only unmodified ions, leading to a poor correspondence between theoretical and experimental data. Matching shifted ions with modifications alongside regular peptide fragment ions in a dual indexing approach has been demonstrated to empower the open search concept in the proteomics field.[32] There is a need to also match fragments containing unknown modifications for fast identification of known and discovery of novel glycation products. Further, there is much valuable information in the neutral losses from the peptide modification side chains, and it is crucial to leverage these characteristic fragments for non-targeted glycation product analysis.

Here, we use three types of fragment ions for open search of peptide glycation products (Figure 1): (i) classic peptide fragment ions (CPFIs) to describe the theoretical peptide ion ($[\text{Pep} + \text{H}]^+$) and all theoretical N-terminal (e.g., a- or b-ions) and C-terminal ions (e.g., y-ions) without modifications, (ii) modified peptide fragment ions (MPFIs) to define peptide ions with modifications, and (iii) characteristic fragment ions (CFIs) to specify modified fragments with characteristic redundant neutral losses. We propose a computational approach that enables fast matching of experimental spectra, which allows exploration for both known and unknown peptide glycation products.

EXPERIMENTAL SECTION

Preparation of Peptide Glucose Model Systems. D-(+)-Glucose (≥99.5%) was purchased from Sigma-Aldrich (Steinheim, Germany). H-Ile-Ile-OH and H-Leu-Leu-OH were purchased from Santa Cruz Biotechnology (Texas, USA). H-Ile-Leu-OH and H-Leu-Ile-OH were purchased from Bachem (Bubendorf, Switzerland). Glucose stock solution (0.08 M) was mixed with each peptide standard stock solution (0.02 M) [both prepared in MilliQ-purified water from a Milli-Q Integral Water

https://doi.org/10.1021/acs.analchem.2c00388
Anal. Chem. 2022, 94, 5953−5961

Purification System (18.2 MΩ, Billerica, MA, USA)] 1:1 (*v*/*v*). Aqueous peptide standard solutions (0.01 M) were prepared as control samples. Model systems were heated in closed glass vials for 2 and 10 h at 100 °C according to the protocol recently described.[33] Control samples were heated, analogously. Sample preparation was performed in triplicate ($n = 3$).

LC-MS/MS. Model systems were diluted 1:5 (*v*/*v*) with an aqueous solution containing 2% acetonitrile (LC−MS grade, Merck, Darmstadt, Germany) prior to LC−MS/MS analysis. Samples were analyzed by ultrahigh-performance liquid chromatography (UHPLC) (Acquity, Waters, Milford, MA, USA) coupled to a quadrupole time-of-flight mass spectrometer (maXis, Bruker Daltonics, Bremen, Germany). For reversed phase (RP) chromatography, an ACQUITY UPLC BEH C8 column (150 × 2.1 mm, 1.7 μm, Waters, Milford, MA, USA) was used. The column temperature was maintained at 60 °C. RP separation was run in gradient mode. The RP eluent A was a composition of 5 mmol/L NH_4Ac and 0.1% acetic acid in water and RP eluent B was composed of acetonitrile. The pre-equilibration time was set to 2.5 min. Initial conditions were set to 90% eluent A and 10% eluent B. This composition was maintained until 1 min. Eluent B was increased to 22% within 2 min. Eluent B then reached 27% within 4 min and 55% after an additional 5.5 min. Eluent B was finally increased to 95% within 2.5 min and kept at 95% until the end of the run. The gradient was completed after 17 min. Samples were injected *via* full-loop injection (10 μL). Mass spectra were acquired in positive electrospray ionization (ESI) mode. Internal calibration was performed using a tuning mix solution (Agilent Technologies, Waldbronn, Germany) prior to each measurement. Parameters of the ESI source were as follows: capillary voltage 4000 V, dry gas temperature 200 °C, nebulizer pressure 2 bar, and nitrogen flow rate 10 L/min. Mass spectra were recorded with an acquisition rate of 5 Hz within a mass range of $m/z = 50-1500$. For data-dependent MS/MS acquisition, the most abundant ion of a full MS scan was subjected to MS/MS after each precursor scan. The collision energy (CE) was set to 25 eV and to change dynamically and depending on the mass of the precursor molecule. For CE optimization, fixed collision energies of 15, 20, and 25 eV were applied. Raw data were post-processed using GeneData Expressionist Refiner MS 13.5 (GeneData GmbH, Basel, Switzerland) applying chemical noise subtraction, intensity cutoff filter, calibration, chromatographic peak picking, and isotope clustering. Only features detected in all three replicates were retained in the matrix.

Analyte Classification. Analytes (retention time ≥ 2 min) were classified according to Yaylayan into two reaction pools:[34] glycation products and peptide-related analytes (e.g., degradation products). Due to poor retention on RP liquid chromatography, sugar degradation products were not considered. Ion signals reaching minimum intensity values of 1500 in all three replicates of the aqueous peptide−glucose model systems but not in control samples (peptides heated alone) were classified as peptide glycation products. Analytes reaching minimum intensity values of 1500 also in peptide blank samples were categorized as peptide-related compounds. Features found in solvent blank samples were considered as contaminants and removed by blank subtraction.

Consensus Spectra Computation. MS/MS consensus spectra computation was performed using the R package MsnBase().[35] MS/MS spectra were merged using the consensusSpectrum() function. To mitigate the abundance of noisy and non-reproducible fragments in fragmentation patterns, we retained mass peaks present in a minimal proportion of 50% of spectra in the final consensus spectra. Intensities were summed for the aggregated peaks, and the maximum m/z merge distance was set to 0.005 Da.

Statistics. To statistically assess the improvement of glycation product MS/MS matching by the use of different ion types, the non-parametric Wilcoxon test was applied for pairwise comparison using the wilcox.test() function of the stats() R package.

Nuclear Magnetic Resonance Spectroscopy. All samples were diluted 1:2 with a 1:1 H_2O/D_2O containing sodium 3-(trimethylsilyl)propionate-d_4 (TSP, 0.9 mM) as a chemical shift reagent and di-sodium hydrogen phosphate (0.75 M, pH 7) to buffer the sample at pH 7. The samples were measured in triplicates. Experiments were carried out using an 800 MHz Bruker AVANCE III spectrometer equipped with a 5 mm QCI-probehead at 300 K. One dimensional (1D) ^1H spectra were recorded using a pulse program constituted of presaturation during the relaxation delay followed by a 90 ° hard pulse. The overall recycling delay of each scan was set to 12 s after T1 relaxation for the peaks under investigation was determined. A 12.5 s hard pulse was used. 64 scans were acquired for every spectrum. The acquisition time was set to 4 s with a spectral width of 12,820 Hz. The assignment of the observed signals was carried out based on comparing the spectra of the unreacted control samples to the spectrum of the mixture. Relative quantification was done by integration of all isolated peaks of the peptide under investigation. The obtained areas were used to calculate the corresponding relative concentration by comparison with the TSP area.

RESULTS AND DISCUSSION

Collision Energy Optimization. As the detected ion profile of MS/MS spectra and the sensitivity are determined by the CE, this parameter was first investigated. Incubation of short-chain peptides (0.01 M) with glucose (0.04 M) at 100 °C for up to 10 h was followed by RP UHPLC-MS/MS analysis. The results were obtained under positive ESI [ESI(+)]. To obtain maximum information about the CPFI, and thus, the amino compound-derived core structure of glycation products, we applied dynamic voltages in the range 15−25 with 5 eV increments. After LC−MS/MS analysis, we computed MS/MS consensus spectra (Figure 1). It has been previously shown that the consensus spectrum is a superior representation compared to the best replicate, especially when measuring a small number of replicates and none of them being of particularly good quality.[36,37] Combining mediocre replicates to form a consensus spectrum allows to remove noise and to average experimental variation. Consensus spectra computation represents a much more robust approach compared to selecting any of the replicates for MS/MS matching, and its benefit for proteomics applications was previously demonstrated.[38]

Importantly, chemical modification of peptides can significantly affect the mass distribution of fragment ions generated by collisional activation.[39−41] For example, b_1 ions appear as dominant features in MS/MS spectra of amidinated peptides, while usually not observed for unmodified peptides due to their instability.[42] Equally, sufficient peptide backbone fragmentation may be observed for the unmodified peptide but cannot be achieved for peptide glycation products using the same CID voltage. CE dependency and variance in the relative abundance of the CPFI for fragmentation of the starter peptide H-Ile-Leu-OH and the corresponding Amadori compound $C_6H_{10}O_5$-Ile-

https://doi.org/10.1021/acs.analchem.2c00388
Anal. Chem. 2022, 94, 5953−5961

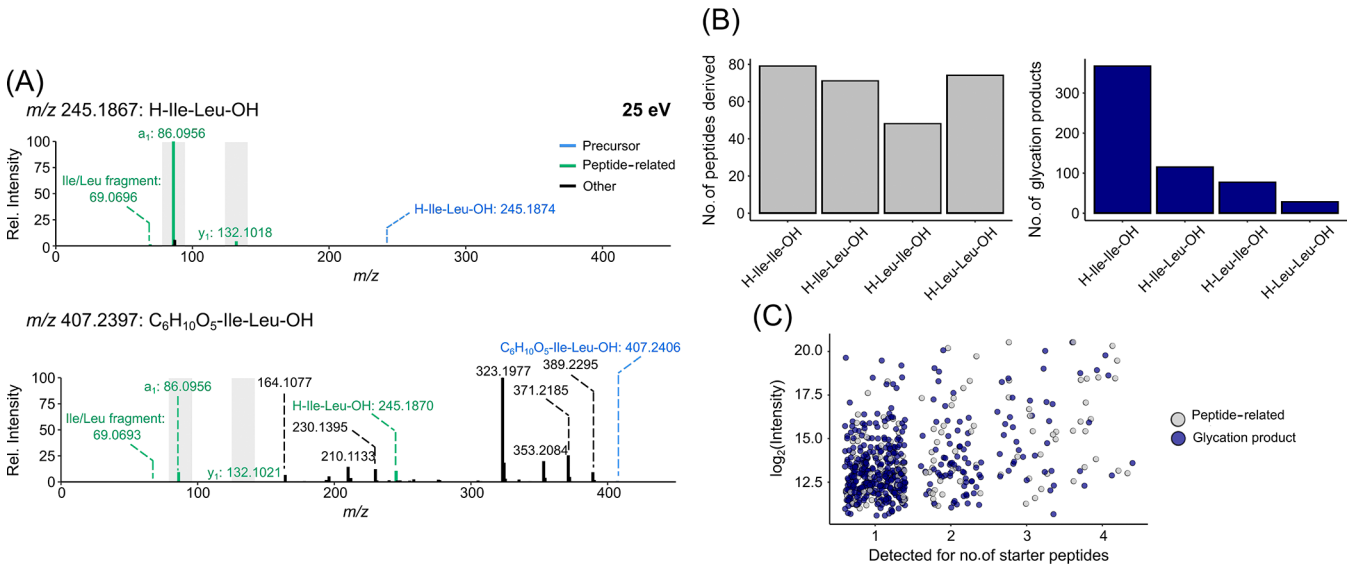


Figure 2. Characterization of short-chain peptide model systems. (A) Consensus MS/MS spectra (25 eV) of the standard peptide H-Ile-Leu-OH (top) and the corresponding Amadori product $C_6H_{10}O_5$-Ile-Leu-OH (bottom). The base peak is assigned to an abundance of 100%. Relative abundance of N-terminal a- and b-ions and C-terminal y-ions is marked by gray boxes. The precursor ion is highlighted in blue. (B) Bar plots of the abundance of peptide-derived analytes (left: gray) and glycation products (right: dark blue) in short-chain peptide-glucose model systems after heat treatment. (C) Maximum intensity as $\log_2$ of analytes across starter peptides. In total, model systems from four different starter peptides were analyzed. Legend: gray, peptide-derived analytes; dark blue, glycation products.

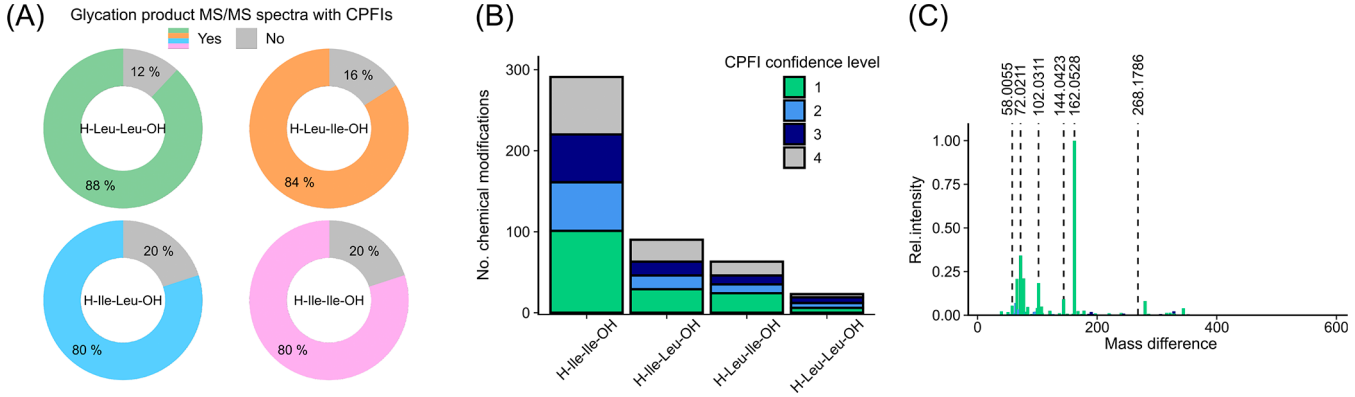


Figure 3. Matching of CPFIs. (A) Percent of the glycation products matched to CPFIs. (B) Distributions of CPFI matching confidence levels versus model system starter peptides. (C) Relative maximum intensity of glycation chemical modifications, shown as mass differences (Δm = $\text{mass}_{\text{glycation product}} - \text{mass}_{\text{peptide matched}}$) of minimum +1.0078 for confidence levels 1−3.

Leu-OH, a well-studied product of early glycation reactions, can be seen in Figure S1. At 15 eV, formation of CPFI appeared to be highest for the unmodified peptide (Figure S1A, left), while the Amadori product showed the highest relative CPFI intensity at 25 eV (Figure S1C, right). For the Amadori product, fragment diversity notably increased at the high voltages (Figure S1, right), suggesting that the formation of MPFIs and CFIs requires excess of energy. Considering the highest abundance of all three ion types, namely CPFIs, MPFIs, and CFIs, the CE was set to 25 eV (Figure 2A), changing dynamically with the precursor mass. This ensures improved certainty for full backbone coverage in peptide glycation product MS/MS spectra and facilitates proper fragmentation of higher molecular mass modifications.

Short-Chain Peptide Demonstration Data Sets. A full comprehensive study of glycation product fragmentation behavior was performed to showcase the algorithm and contribute to the dark matter of non-enzymatic chemical modifications. We chose glucose model systems from four (iso)leucine peptides, namely H-Ile-Ile-OH, H-Ile-Leu-OH, H-Leu-Ile-OH, and H-Leu-Leu-OH, as demonstration data sets. It is important to note that the short-chain peptides investigated herein were selected to keep fragmentation patterns simple and to display poor comparability of glycation reactions even among peptides with highly similar amino acid sequences. To visualize the linkage between reagents and products, we show the number of peptide-derived analytes (e.g., degradation products) and glycation products for these four standard peptides (Figure 2B). Classification of the reaction pool into peptide degradation products and reaction products was inspired by Yaylayan's approach[34] (for details, see Methods: Analyte Classification). For peptide-derived signals, similar numbers were detected independent of the starter peptide (Figure 2B, left). Substitution of leucine for isoleucine, especially at the N-terminal sequence position, caused pronounced differences in glycation product diversity (Figure 2B, right). We performed nuclear magnetic resonance spectroscopy to quantify the peptide consumed by glycation reactions and to get additional insights into the peptide reaction behavior *via* an orthogonal measurement technology.

https://doi.org/10.1021/acs.analchem.2c00388
Anal. Chem. 2022, 94, 5953−5961

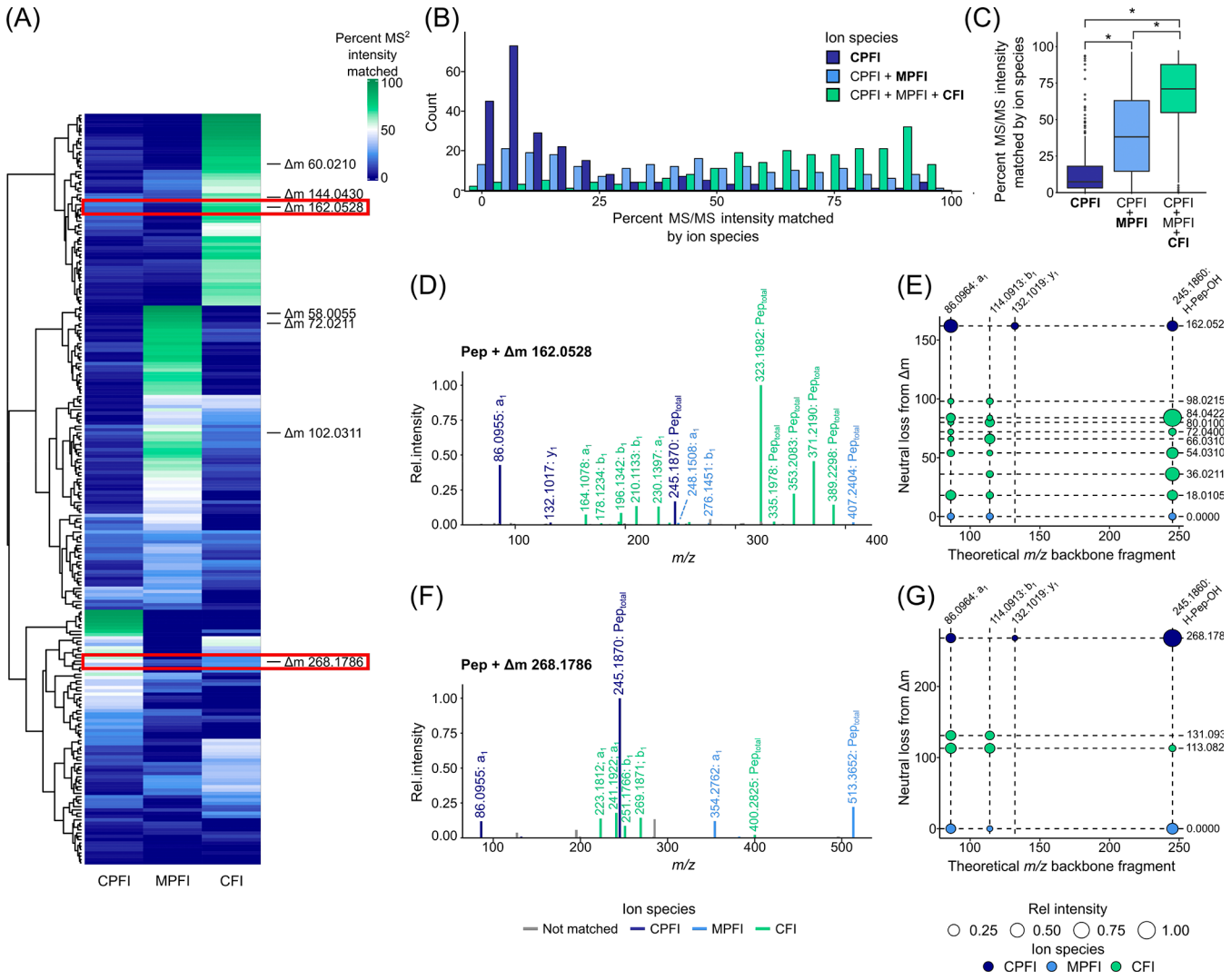


Figure 4. Contribution of CPFIs, MPFIs and CFIs to peptide glycation product MS/MS spectra. (A) Relative MS/MS intensity of each ion type visualized as a heatmap for peptide glycation products. For each mass difference, the analyte with the highest relative MS/MS intensity matched is shown. Row clustering is based on the results from hclust analysis ("ward.D2" method). Mass differences illustrated in (D−G) are highlighted by red rectangles. (B) The distribution shows the high or low coverage of glycation product MS/MS spectra upon use of CPFIs, MPFIs, and CFIs. (C) The boxplot shows glycation product MS/MS coverage for the different ion species. The box represents the interquartile range, the horizontal line in the box is the median, and the whiskers represent 1.5 times the interquartile range. Asterisks indicate significant changes (Wilcoxon test, $p < 0.001$). As an example, representative MS/MS consensus spectra are illustrated for two selected chemical modifications with a mass delta of 162.0528 Da (D) and 268.1786 Da (F) color-coded by ion type. Redundant neutral losses are shown for CFI characterization depending on corresponding peptide backbone ion m/z for the glycation modifications 162.0528 Da (E) and 268.1768 Da (G).

As suggested by the MS results, superior reactivity of isoleucine-peptides was indicated by higher levels of peptide loss (Figure S2) and increased model system complexity (Figures S3 and S4). High specificity of glycation products for the peptide reagent is visualized in Figure 2C, showing that most glycation products were unique for one single starter peptide. This may not only be based on differences in reactivity but also on marginal deviations in the starter peptide physicochemical properties, which can cause retention time shifts for the same type of reaction product. Observation of such pronounced discrepancies for highly similar short-chain peptides demonstrates that there is a need to establish an algorithm to compare glycation modifications across peptides, comprehensively.

Fragmentation Study of Short-Chain Peptide Glycation Derivatives. An algorithm for automated search of peptide glycation products can only work well if the peptide component embedded in the cores of the reaction products is considered. Thus, we begin by matching experimental spectra against the CPFIs of the corresponding candidate peptides. For CPFI fragment matching, we perform *in silico* digestion of the four peptide standards studied in this work. This peptide fragment database was complemented by an in-house amino acid LC−MS/MS library, which is available from the Mass Bank of North America. Figure 3 shows the results of this first step of the multi-step algorithm. Searching acquired glycation product MS/MS spectra against the CPFI database, up to 88% of each reaction product pool shared peptide backbone fragments or had amino acid MS/MS fragmentations (Figure 3A). CPFI intensities may vary depending on the chemical modification, and analytes were classified into four confidence levels depending on the match quality: Level 1, complete peptide backbone detection by both N- (a- and b-ions) and C-terminal (y-) ions, respectively; detection of the ion related to the unmodified peptide (e.g., m/z 245.1860 for [H-Ile-Leu-OH +

https://doi.org/10.1021/acs.analchem.2c00388
Anal. Chem. 2022, 94, 5953−5961

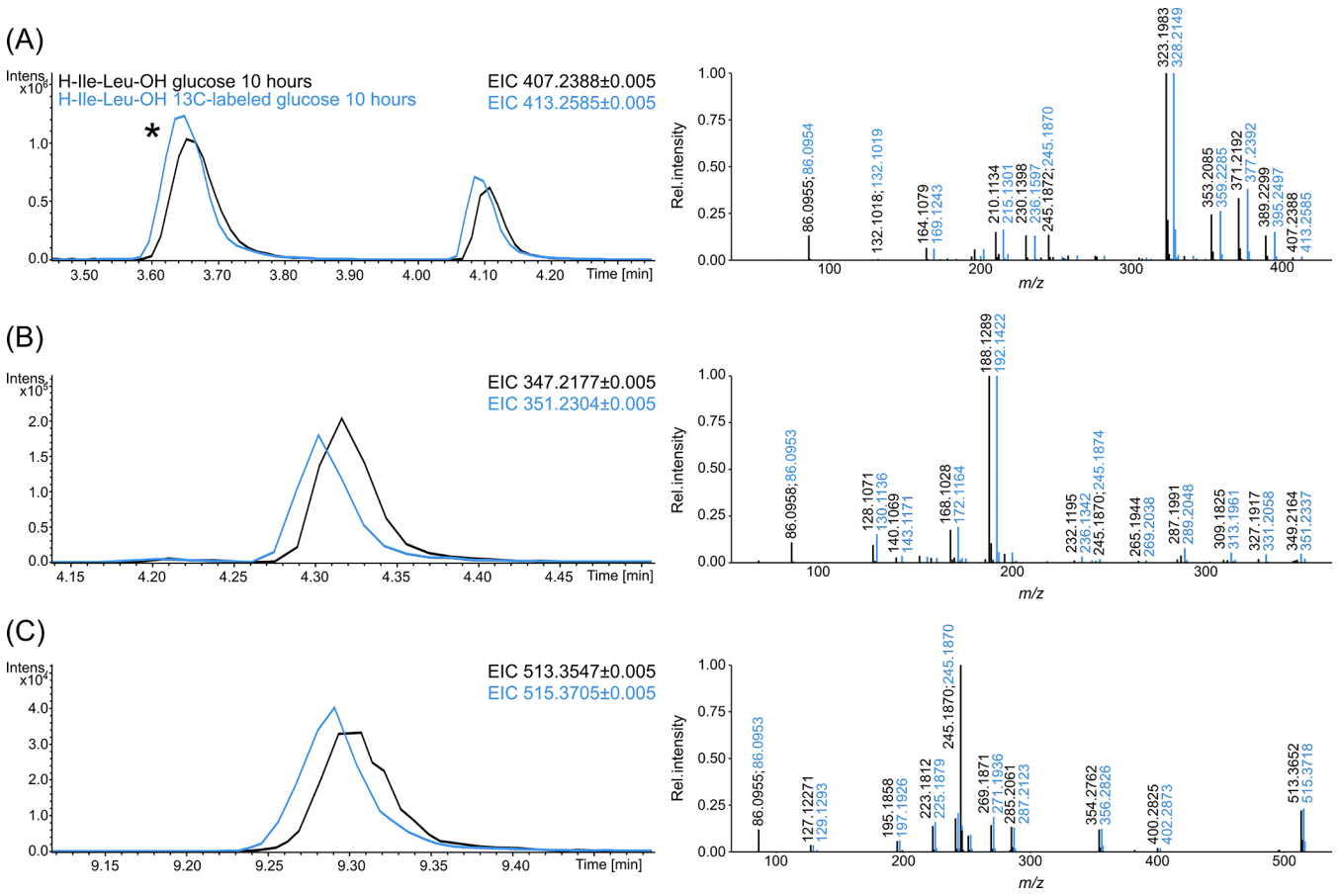


Figure 5. Chromatographic and mass spectrometric characterization of selected glycation modifications using isotope labeling. Representative extracted ion chromatograms (left) and collision-induced dissociation consensus MS/MS experiments (right) of peptide chemical modifications with a mass delta of 162.0528 Da (A), 102.0311 Da (B), and 268.1786 Da (C) (non-labeled: black, ^{13}C-labeled: blue).

$H]^+$). Level 2, complete peptide backbone detectability considering all types of peptide backbone ions; no requirement for detection of the unmodified peptide ion. Level 3, detection of any N- or C-terminal peptide ions. Level 4, only observation of ions related to fragmentation of relevant amino acids. Most CPFI matches correspond to confidence levels 1 and 2 for all starter peptides (Figure 3B), meaning that sufficient fragmentation of the peptide backbone was achieved for the majority of the glycation products. Chemical modification mass differences (Δm) were calculated as follows

$$\Delta m = \text{mass}_{\text{glycation product}} - \text{mass}_{\text{peptide matched}}$$

Summarized delta masses are shown in Figure 3C and Table S1. Mass differences at least observed for confidence level 3 and of minimum 1.0078 (mass of one hydrogen atom) were included. Highly abundant examples are highlighted (e.g., delta masses 162.0528, 144.0423, 72.0211, and 58.0055 Da). A mass increase of 162.0528 Da may be putatively assigned as the Amadori product (+$C_6H_{10}O_5$, 162.0528 ± 0.005 Da). Delta masses 58.0055 and 72.0211 Da may correspond to carboxymethylation (+$C_2H_2O_2$, 58.0000 ± 0.005 Da) and carboxyethylation (+$C_3H_4O_2$, 72.0211 ± 0.005 Da). Other hitherto undescribed glycation modifications with high abundance were observed (e.g., +$C_4H_6O_3$, 102.0316 ± 0.005 Da) as well as higher molecular mass shifts (e.g., +$C_{14}H_{24}N_2O_3$, 268.1787 ± 0.005 Da).

Next, experimental spectra were searched for MPFIs. MPFIs carry the glycation modification Δm and MPFI calculation is expressed as

$$m/z_{\text{MPFI}} = m/z_{\text{a-,b-ion}} + \Delta m$$

Fragment ions that remained unmatched were used for the detection of redundant neutral losses to identify CFIs. We believe that the key feature of an effective peptide glycation product search strategy is to determine not only the peptide backbone but also the ability to subtract the peptide core structure and unmask the modification structural information concealed in the MS/MS spectra. In particular, the implementation of CFIs can determine important neutral losses that do not represent core structural information and allow the comparison of glycation modifications across different peptide types.

Overall, our approach found glycation modifications predominantly producing CPFIs, MPFIs, or CFIs, respectively (Figure 4A). A rather poor contribution of CPFIs to experimental spectra was identified for most peptide glycation products. For glycation products with a low CPFI relative intensity, more than half of the useful spectral information is ignored when solely matching unmodified peptide ions. However, the amino compound core structure alone is also key information for structural identification of both known and unknown peptide glycation products. MPFIs are particularly useful for non-labile modifications, where peptide fragment ions retain the modification after dissociation of the peptide core

https://doi.org/10.1021/acs.analchem.2c00388
Anal. Chem. 2022, 94, 5953−5961

structure.[32] Other glycation products are highly prone to fragmentation of the modification side chain (e.g., Amadori products) and mainly produce CFIs by redundant neutral losses (e.g., water loss), which may give additional structural information. A high abundance of CFIs is also expected for peptides with multiple modifications (e.g., $\Delta m = \Delta m_1 + \Delta m_2$) as detectability of MPFIs may be limited in such fragmentation patterns and each individual modification Δm_n will instead be assigned to be characteristic. The fraction of interpretable fragments was remarkably boosted by refined matching using MPFIs and CFIs alongside CPFIs (Figure 4B). With MPFIs and CFIs in conjunction with regular peptide fragment ions, we significantly increased the median MS/MS intensity matched from 7 to 71% (Wilcoxon test, $p < 0.001$, Figure 4C).

We present in Figure 4D and Figure 4F two examples for matched experimental spectra of peptide glycation modifications that show different relative intensities for the three ion types (see Table S2 for detailed information on matched fragment ions). The consensus spectrum of the putatively assigned Amadori product ($+C_6H_{10}O_5$, Δm = 162.0528 ± 0.005 Da) primarily shows CFI fragments (Figure 4A, highlighted in red rectangle; Figure 4D). Here, identification sensitivity toward MPFIs is low, and such modifications do barely benefit from the inclusion of modified ions on matching. Due to its labile nature, this type of modification easily fragments. This is also true for post-translational modifications (PTMs) that are searched in the proteomics field, such as phosphorylation.[32] The main CFIs were characterized by a neutral loss of 84.0422 Da (Figure 4E), which can be annotated as $-3H_2O-CH_2O$ (84.0423 ± 0.005 Da). Besides, distinct CFI neutral losses were 18.0105 Da ($-H_2O$, 18.0105 ± 0.005 Da), 36.0211 Da ($-2H_2O$, 36.0211 ± 0.005 Da), and 54.0310 Da ($-3H_2O$, 54.0317 ± 0.005 Da). Consequently, our approach identified all Amadori diagnostic ions described in the literature so far.[22,43] We also observed that our algorithm helped to find 66.0310, 72.0400, 80.0100, and 98.0215 Da as additional redundant neutral losses, demonstrating the potential of the open search strategy.

The delta mass 268.1786 Da was annotated as a potential peptide crosslink. It may correspond to two peptide moieties that were linked *via* an ethylene bridge ($+C_{14}H_{24}N_2O_3$, 268.1787 ± 0.005 Da). The consensus MS/MS spectrum of this modification showed a large fraction of CPFIs. The relative contribution of MPFIs and CFIs was similar (Figure 4A, highlighted in red rectangle). As redundant neutral losses, 113.0820 Da ($-C_6H_{11}NO$, 113.0841 ± 0.005 Da) and 131.0934 Da ($-C_6H_{13}NO_2$, 131.0940 ± 0.005 Da) were determined (Figure 4G). These neutral losses correspond to fragmentation of the peptide core structure.

Algorithm Validation Using Stable Isotope-Labeling. Stable isotope-labeling was used to provide increased annotation confidence and to further demonstrate the utility of our open search method. Reliable identification of peptide glycation products, especially if previously undescribed, has been traditionally challenging with long data analysis times and increased false annotation rates. Applying our triple-ion search algorithm on LC−MS/MS data from both unlabeled and isotopically tagged analytes enables high-confidence annotation in an untargeted manner. For non-targeted profiling of ^{13}C-labeled glycation products, model systems were prepared from D-glucose-$^{13}C_6$. H-Ile-Leu-OH was used as the peptide reagent. Both well-known and hitherto uncharacterized glycation products were used as benchmarks. Based on chromatographic and mass spectrometric properties of the labeled reaction products, identification and characterization of selected glycation modifications were achieved. Because of their similar structures, the retention times of ^{13}C-labeled and non-labeled products were comparable (Figure 5, left). The mass-to-charge ratio differences ($\Delta m/z$) were $\Delta m/z$ = 6.0197 (Figure 5A, 6× $^{12}C \leftrightarrow {}^{13}C$ $\Delta m/z_{theoretical}$ = 6.0206), $\Delta m/z$ = 4.0127 (Figure 5B, 4× $^{12}C \leftrightarrow {}^{13}C$ $\Delta m/z_{theoretical}$ = 4.0137) and $\Delta m/z$ = 2.0158 (Figure 5C, 2× $^{12}C \leftrightarrow {}^{13}C$ $\Delta m/z_{theoretical}$ = 2.0069). These results indicate the number of glucose-derived carbon atoms that contribute to the glycation product structures (1× $^{12}C \leftrightarrow {}^{13}C$ $\Delta m/z_{theoretical}$ = 1.0034). The fragmentation patterns of labeled glycation products allow verification of known and identification of undescribed modifications. As expected, the ^{13}C-labeled and non-labeled reaction products had highly similar MS/MS spectra (Figure 5, right). Detection of common fragment ions (e.g., a_1: m/z 86.0964, b_1: m/z 132.1019, $[H\text{-Ile-Leu-OH} + H]^+$: m/z 245.1860) suggests the absence of isotopic tags. This is expected as CPFIs do not contain any modification and may be detected for all glycated peptides. In ^{13}C-labeled models, for example, the fragment ions m/z 395.2497, 377.2392, and 359.2285 (Figure 5A right, blue) of the putatively assigned Amadori product ($C_6H_{10}O_5$-Ile-Leu-OH, m/z 407.2388 ± 0.005 Da) were mass shifts $\Delta m/z$ = 6.0206 ± 0.005 to their corresponding unlabeled ions ($-H_2O$: m/z 389.2299, $-2H_2O$: m/z 371.2192, $-3H_2O$: m/z 353.2085; Figure 5A right, black). These fragments contained whole isotopic tags, which was expected as there is no ^{13}C loss from the $^{13}C_6H_{10}O_5$ side chain by any of these fragmentation pathways. The labeled ion m/z 328.2149 only retains a part of its isotopic tags and differs from its non-labeled ion m/z 323.1983 ($-3H_2O-CH_2O$) by 5× $^{12}C \leftrightarrow {}^{13}C$ ($\Delta m/z$ = 5.0172 ± 0.005). All described fragmentation patterns are consistent with the literature.

We show in Figure 5B (right) that for the mass delta 102.0311 Da ($C_4H_6O_3$-Ile-Leu-OH: m/z 345.2020 ± 0.005 Da), the $\Delta m/z$ between isotopically tagged ions (e.g., m/z 331.2058 and 313.1961) and the unlabeled fragments (e.g., m/z 327.1917 and 309.1825) was 4× $^{12}C \leftrightarrow {}^{13}C$. Peptide backbone fragmentation ($-C_6H_{13}NO_2-CO$) generated the base peaks at m/z 188.1289 (black) and m/z 192.1422 (blue), leaving the labeled glycation modification intact. These results suggested that the unknown modification does indeed contain four glucose-derived carbon atoms as annotated. The isotopically labeled fragment at m/z 289.2048 was formed by a loss of $^{13}C_2H_2O_2$ from the modification side chain and corresponds to m/z 287.1991 in the unlabeled MS/MS spectrum, shifted by 2× $^{12}C \leftrightarrow {}^{13}C$. Sequential fragmentation of the peptide backbone ($-C_6H_{13}NO_2-CO$) and loss of $C_2H_4O_2$ from the modification side chain leads to m/z 128.1071 (black) shifted by 2× $^{12}C \leftrightarrow {}^{13}C$ in the labeled MS/MS spectrum (blue, m/z 130.1136). In the context of MR, $C_4H_6O_3$ was previously described as free 2-hydroxy-3-oxobutanal, which can isomerize into various tautomers with methylreductone and dicarbonyl structures.[44] To the best of our knowledge, no peptide bound glycation modification with this molecular formula has been previously described.

For the hitherto unknown cross-linked H-Ile-Leu-OH, isotopically tagged ions (e.g., m/z 402.2873, 356.2826, and 287.2123) showed a $\Delta m/z$ of 2× $^{12}C \leftrightarrow {}^{13}C$ (m/z 400.2528, 354.2762, and 285.2061; Figure 5B, right). This indicates ^{13}C-labeling of the hypothesized ethylene cross-link coming from the sugar compound. The labeled linkage may be retained, while the peptide moieties dissociate by peptide bond fragmentation

https://doi.org/10.1021/acs.analchem.2c00388
Anal. Chem. 2022, 94, 5953−5961

giving, for example, m/z 400.2825 ($-C_6H_{11}NO$) and m/z 354.2726 ($-C_6H_{13}NO_2-CO$) shifted by $\Delta m/z$ of $2\times$ $^{12}C \leftrightarrow {}^{13}C$ in the labeled MS/MS spectrum (m/z 402.2872 and 356.2826).

Integrating complementary MS/MS spectra from ^{13}C-labeled glycation products, we provide high-confidence structural information. Both known and unknown peptide glycation product annotation could be refined with orthogonal stable isotope-labeling approaches to validate the predicted structure identities.

CONCLUSIONS

In summary, understanding MS/MS spectra of glycated peptides is a crucial yet challenging task due to their complicated and peptide core structure-dependent fragmentation pattern. In this work, we revealed the labile nature of many glycation side chains, pointing out the necessity to include characteristic neutral losses for MS/MS matching. We thus developed a triple-ion strategy, also considering side-chain fragmentation, to significantly improve matching of peptide glycation product MS/MS spectra. From the study of short-chain peptide model systems, we conclude that our multi-step algorithm allows for fast, accurate, and automatic identification of well-known glycation products. By characterizing previously undescribed peptide glycation modifications, we further demonstrate that our straightforward approach is highly suitable for exploration of novel glycation products and that redundant neutral losses constitute an informative source to accurately determine their molecular identity. We also discuss the possibility to efficiently integrate isotopically labeled and label-free fragmentation patterns for increased annotation confidence. Overall, our work shows that the use of MPFIs and CFIs among conventional CPFIs provides a significant advancement in MS/MS spectra analysis, empowering the open search concept for modified peptides. We envision that our triple-ion approach, with more successful cases demonstrated, will show a great value in future non-targeted analyses of peptide modifications.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.2c00388.

Additional information; additional details about CE optimization; peptide quantification and model system characterization using ^1H NMR spectroscopy (PDF)

Glycation mass differences (XLSX)

Matched ion types for mass difference 162.0528 and 268.1786 Da (XLSX)

AUTHOR INFORMATION

Corresponding Authors

Michelle T. Berger − *Chair of Analytical Food Chemistry, Technical University Munich, 85354 Freising, Germany; Research Unit Analytical BioGeoChemistry (BGC), Helmholtz Zentrum München, 85764 Neuherberg, Germany;* orcid.org/0000-0003-1438-3009; Email: michelle.berger@tum.de

Philippe Schmitt-Kopplin − *Chair of Analytical Food Chemistry, Technical University Munich, 85354 Freising, Germany; Research Unit Analytical BioGeoChemistry (BGC), Helmholtz Zentrum München, 85764 Neuherberg, Germany;* Email: schmitt-kopplin@helmholtz-muenchen.de

Authors

Daniel Hemmler − *Chair of Analytical Food Chemistry, Technical University Munich, 85354 Freising, Germany; Research Unit Analytical BioGeoChemistry (BGC), Helmholtz Zentrum München, 85764 Neuherberg, Germany*

Philippe Diederich − *Research Unit Analytical BioGeoChemistry (BGC), Helmholtz Zentrum München, 85764 Neuherberg, Germany*

Michael Rychlik − *Chair of Analytical Food Chemistry, Technical University Munich, 85354 Freising, Germany*

James W. Marshall − *The Waltham Petcare Science Institute, Mars Petcare UK, Leicestershire LE14 4RT, United Kingdom;* orcid.org/0000-0003-3352-7629

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.2c00388

Author Contributions

M.T.B. designed the study and performed the MS experiments, statistical and bioinformatics data analysis, and wrote the paper. D.H. supervised the data analysis. P.D. performed the nuclear magnetic resonance spectroscopy experiments and data analysis. P.S.K. provided the resources. M.T.B., D.H., P.D., J.W.M., and P.S.K. discussed and interpreted the results. All authors revised and approved the manuscript. Correspondence to M.T.B. and P.S.K.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Harmel, R.; Fiedler, D. *Nat. Chem. Biol.* **2018**, *14*, 244−252.
(2) Stadtman, E. R. *Free Radical Res.* **2006**, *40*, 1250−1258.
(3) Rabbani, N.; Thornalley, P. J. *Biochem. Biophys. Res. Commun.* **2015**, *458*, 221−226.
(4) Hellwig, M.; Henle, T. *Angew. Chem., Int. Ed.* **2014**, *53*, 10316−10329.
(5) Hodge, J. E. *J. Agric. Food Chem.* **1953**, *1*, 928−943.
(6) Hemmler, D.; Roullier-Gall, C.; Marshall, J. W.; Rychlik, M.; Taylor, A. J.; Schmitt-Kopplin, P. *Sci. Rep.* **2017**, *7*, 3227.
(7) Hemmler, D.; Roullier-Gall, C.; Marshall, J. W.; Rychlik, M.; Taylor, A. J.; Schmitt-Kopplin, P. *Sci. Rep.* **2018**, *8*, 16879.
(8) van Boekel, M. A. J. S. *Nahrung* **2001**, *45*, 150−159.
(9) Mittelmaier, S.; Pischetsrieder, M. *Anal. Chem.* **2011**, *83*, 9660−9668.
(10) Lopez-Clavijo, A. F.; Barrow, M. P.; Rabbani, N.; Thornalley, P. J.; O'Connor, P. B. *Anal. Chem.* **2012**, *84*, 10568−10575.
(11) Redman, E. A.; Ramos-Payan, M.; Mellors, J. S.; Ramsey, J. M. *Anal. Chem.* **2016**, *88*, 5324−5330.
(12) Lapolla, A.; Gerhardinger, C.; Baldo, L.; Fedele, D.; Keane, A.; Seraglia, R.; Catinella, S.; Traldi, P. *Biochim. Biophys. Acta, Mol. Basis Dis.* **1993**, *1225*, 33−38.
(13) Ahmed, N.; Thornalley, P. J. *Biochem. J.* **2002**, *364*, 15−24.
(14) Zhou, Y.; Lin, Q.; Jin, C.; Cheng, L.; Zheng, X.; Dai, M.; Zhang, Y. *J. Food Sci.* **2015**, *80*, C207−C217.
(15) Hashimoto, C.; Iwaihara, Y.; Chen, S. J.; Tanaka, M.; Watanabe, T.; Matsui, T. *Anal. Chem.* **2013**, *85*, 4289−4295.
(16) Sillner, N.; Walker, A.; Lucio, M.; Maier, T. V.; Bazanella, M.; Rychlik, M.; Haller, D.; Schmitt-Kopplin, P. *Front. Mol. Biosci.* **2021**, *8*, 660456.
(17) Schwarzenbolz, U.; Hofmann, T.; Sparmann, N.; Henle, T. *J. Agric. Food Chem.* **2016**, *64*, 5071−5078.
(18) Frolova, N.; Soboleva, A.; Nguyen, V. D.; Kim, A.; Ihling, C.; Eisenschmidt-Bönn, D.; Mamontova, T.; Herfurth, U. M.; Wessjohann, L. A.; Sinz, A.; Birkemeyer, C.; Frolov, A. *Food Chem.* **2021**, *347*, 128951.
(19) Meltretter, J.; Wüst, J.; Dittrich, D.; Lach, J.; Ludwig, J.; Eichler, J.; Pischetsrieder, M. *J. Proteome Res.* **2020**, *19*, 805−818.

(20) Xing, H.; Mossine, V. V.; Yaylayan, V. *Carbohydr. Res.* **2020**, *491*, 107985.
(21) Yuan, H.; Sun, L.; Chen, M.; Wang, J. *J. Food Sci.* **2016**, *81*, C1662−C1668.
(22) Frolov, A.; Hoffmann, P.; Hoffmann, R. *J. Mass Spectrom.* **2006**, *41*, 1459−1469.
(23) Festring, D.; Hofmann, T. *J. Agric. Food Chem.* **2010**, *58*, 10614−10622.
(24) Priego-Capote, F.; Scherl, A.; Möller, M.; Waridel, P.; Lisacek, F.; Sanchez, J.-C. *Mol. Cell. Proteomics* **2010**, *9*, 579−592.
(25) Zhang, J.; Zhang, T.; Jiang, L.; Hewitt, D.; Huang, Y.; Kao, Y.-H.; Katta, V. *Anal. Chem.* **2012**, *84*, 2313−2320.
(26) Li, S.; Gao, D.; Song, C.; Tan, C.; Jiang, Y. *Anal. Chem.* **2019**, *91*, 1701−1705.
(27) Yu, F.; Li, N.; Yu, W. *J. Proteome Res.* **2016**, *15*, 4423−4435.
(28) Chi, H.; Liu, C.; Yang, H.; Zeng, W.-F.; Wu, L.; Zhou, W.-J.; Wang, R.-M.; Niu, X.-N.; Ding, Y.-H.; Zhang, Y.; Wang, Z.-W.; Chen, Z.-L.; Sun, R.-X.; Liu, T.; Tan, G.-M.; Dong, M.-Q.; Xu, P.; Zhang, P.-H.; He, S.-M. *Nat. Biotechnol.* **2018**, *36*, 1059−1061.
(29) Devabhaktuni, A.; Lin, S.; Zhang, L.; Swaminathan, K.; Gonzalez, C. G.; Olsson, N.; Pearlman, S. M.; Rawson, K.; Elias, J. E. *Nat. Biotechnol.* **2019**, *37*, 469−479.
(30) Solntsev, S. K.; Shortreed, M. R.; Frey, B. L.; Smith, L. M. *J. Proteome Res.* **2018**, *17*, 1844−1851.
(31) Bittremieux, W.; Meysman, P.; Noble, W. S.; Laukens, K. *J. Proteome Res.* **2018**, *17*, 3463−3474.
(32) Yu, F.; Teo, G. C.; Kong, A. T.; Haynes, S. E.; Avtonomov, D. M.; Geiszler, D. J.; Nesvizhskii, A. I. *Nat. Commun.* **2020**, *11*, 4065.
(33) Berger, M. T.; Hemmler, D.; Walker, A.; Rychlik, M.; Marshall, J. W.; Schmitt-Kopplin, P. *Sci. Rep.* **2021**, *11*, 13294.
(34) Yaylayan, V. A. *Trends Food Sci. Technol.* **1997**, *8*, 13−18.
(35) Gatto, L.; Lilley, K. S. *Bioinformatics* **2012**, *28*, 288−289.
(36) Tabb, D. L.; Thompson, M. R.; Khalsa-Moyers, G.; VerBerkmoes, N. C.; McDonald, W. H. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1250−1261.
(37) Flikka, K.; Meukens, J.; Helsens, K.; Vandekerckhove, J.; Eidhammer, I.; Gevaert, K.; Martens, L. *Proteomics* **2007**, *7*, 3245−3258.
(38) Lam, H.; Deutsch, E. W.; Eddes, J. S.; Eng, J. K.; Stein, S. E.; Aebersold, R. *Nat. Methods* **2008**, *5*, 873−875.
(39) Beardsley, R. L.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 158−167.
(40) Beardsley, R. L.; Sharon, L. A.; Reilly, J. P. *Anal. Chem.* **2005**, *77*, 6300−6309.
(41) Korwar, A. M.; Vannuruswamy, G.; Jagadeeshaprasad, M. G.; Jayaramaiah, R. H.; Bhat, S.; Regin, B. S.; Ramaswamy, S.; Giri, A. P.; Mohan, V.; Balasubramanyam, M.; Kulkarni, M. J. *Mol. Cell. Proteomics* **2015**, *14*, 2150−2159.
(42) O'Hair, R. A. J.; Reid, G. E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1220−1225.
(43) Xing, H.; Yaylayan, V. *Carbohydr. Res.* **2020**, *495*, 108091.
(44) Davídek, T.; Devaud, S.; Robert, F.; Blank, I. *J. Agric. Food Chem.* **2006**, *54*, 6667−6676.

https://doi.org/10.1021/acs.analchem.2c00388
Anal. Chem. 2022, 94, 5953−5961