Published in final edited form as:

J Proteome Res. 2014 February 7; 13(2): 777–785. doi:10.1021/pr400848n.

Accurate Identification of Deamidated Peptides in Global Proteomics using a Quadrupole Orbitrap Mass Spectrometer

Angelito I. Nepomuceno¹, Radiance J. Gibson¹, Shan M. Randall¹, and David C. Muddiman¹

¹W.M. Keck Fourier Transform Mass Spectrometry Laboratory, Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695

Abstract

Deamidation of asparagine and glutamine residues is a common post-translational modification. Researchers often rely on mass spectrometric based proteomic techniques for the identification of these post-translational sites. Mass spectral analysis of deamidated peptides is complicated and often misassigned due to overlapping ¹³C peak of the amidated form with the deamidated monoisotopic peak; these two peaks are only separated by 19.34 mDa. For proper assignment it is inherently important to use a mass spectrometer with high mass measurement accuracy and high resolving power. Herein, mouse brain tissue lysate was prepared using filter-aided sample preparation (FASP) method and Stage Tip fractionation followed by analysis on a nanoLC coupled to a quadrupole orbitrap (Q-Exactive) mass spectrometer to accurately identify more than 5400 proteins. Mass spectral data was processed using MASCOT and ProteoIQ for accurate identification of peptides and proteins. MASCOT search values for precursor and MS/MS mass tolerances were investigated, and it was determined data searched with greater than 5 ppm precursor mass tolerance resulted in the misassignment of deamidated peptides. Peptides that were identified with a mass measurement accuracy of ± 5 ppm were correctly assigned.

Introduction

Mass spectrometry continues to be a leading technology used in the field of proteomics.¹ The widely used shotgun “bottom-up proteomics” approach allows for the identification of proteins from complex mixtures. This method relies on the enzymatic digestion of proteins into peptides, followed by separation using nanoscale liquid chromatography (LC) coupled to a mass spectrometer. Data is often collected in a data dependent acquisition (DDA) mode, in which a full precursor spectrum is acquired and then followed by a series of MS/MS product ion spectra. The ions chosen for MS/MS are based on the ion abundance with a defined threshold, and the number of consecutive product ion spectra is defined by the user.

* Author for Correspondence David C. Muddiman, Ph.D., W.M. Keck Fourier Transform Mass Spectrometry Laboratory, Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695, Phone: 919-513-0084, david_muddiman@ncsu.edu.

The series of precursor ions and resulting MS/MS ion peaks can be searched using a variety of search engines. Several programs such as SEQUEST^{2, 3}, Mascot⁴ and X!Tandem⁵ are often utilized.

Database searching requires a user-defined protein database obtained for a specified organism. Each program performs an *in-silico* digestion of the protein database to produce a peptide list based on specified parameters such as digestive enzymes, number of missed cleavages, and fixed/variable modifications. A precursor *m/z* is first matched to possible candidate peptides within a defined mass tolerance. MS/MS spectra are then matched to possible theoretical sequence information from the candidate peptides and then given a confidence score. These identified peptides are then grouped into their corresponding protein(s).

Challenges in global proteomics arise with accurate identification of low-abundant peptides with insufficient precursor ion abundance. Advancements in instrument acquisition speeds, resolving power, and mass measurement accuracies have greatly improved the number of proteins identified. To achieve the maximum number of proteins identified, a quadrupole orbitrap mass spectrometer (Q-Exactive) was utilized in these experiments. The Q-Exactive offers many attributes that allow for a comprehensive analysis of proteins extracted from a complex biological matrix. Unlike hybrid instruments with dual mass analyzers, all spectra are obtained using an orbitrap mass analyzer with acquisition rates as fast as >12 Hz at 17.5 k_{FWHM} resolving power.^{6, 7} High resolving power spectra and high mass measurement accuracy can thus be obtained with the use of the orbitrap mass analyzer. Many researchers have developed strategies that account for high resolution tandem mass spectra⁸ and high precursor mass accuracy⁹.

Researchers commonly process their protein samples first by using reducing agents such as dithiothreitol to break any intramolecular and intermolecular disulfide bridges. The addition of iodoacetamide (IAM) is often followed to covalently bond with the free thiol groups to prevent any cysteine disulfide formations from reoccurring. Typical trypsin digestion protocols use buffers with a pH value between 7–8 for 1–24 hour duration at 37 °C. During the peptide process, peptides are frequently searched with post-translation modifications. Peptides are searched with a fixed carbamidomethyl modification on cysteine residues that accounts for the iodoacetamide (+ 57.02 Da). However, data is often searched with variable modification such as oxidation of methionine (+ 15.99 Da) and deamidation of asparagines (N) and glutamines (Q) (+ 0.98 Da). Researchers have optimized trypsin proteolysis in order to maximize detection of targeted peptides.^{10–12}

Oxidation of methionine residues is an important post-translational modification; it has been related to oxidative stress and aging.¹³ However, oxidation can be induced during the electrospray process.¹⁴ This is the reason for searching oxidation on methionine residues as a variable modification. Differentiating between the oxidized and unmodified peptides can easily be distinguished due to the + 15.99 Da mass shift. Deamidation of asparagine and glutamine on the other hand results in only a + 0.98 Da mass shift. Deamidation has been related to protein degradation and is thought to play an important part of aging studies.^{15, 16} Deamidation has also resulted in loss of protein structure, solubility and function.¹⁷ Further, accurate identification of deamidated peptides impacts the assignment of *N*-glycosylated sites.

The deamidation of asparagine to aspartic acid and glutamine to glutamic acid is a non-enzymatic process, the half-life can be between 1–500 days.¹⁸ Current protocols for tryptic digestion are favorable for deamidation due to the basic pH condition and can cause the misrepresentation of deamidation of native proteins. To circumvent this challenge, Li *et al.*

has implemented an O¹⁸ labeling strategy to monitor deamidation during sample preparation.¹⁹ MALDI and HPLC-MALDI studies resulted in approximately 70–80% of –Asn-Gly- sites as being deamidated after a standard 12 hr tryptic digestion at 37 °C.²⁰ Hao and coworkers have shown that the ¹³C peaks of amidated peptides can be misassigned as monoisotopic peaks of the corresponding deamidated peptides using a linear quadrupole ion trap fourier transform mass spectrometer, and they discuss an arbitrary method for the determination of deamidated or amidated peptides.¹⁰

Search parameters were optimized in order to increase the number of identified proteins for a global proteome study on mouse brain tissue lysate. Increasing the value on the precursor mass tolerance beyond 5 ppm resulted in an increase in proteins identified. It was also determined that a MS/MS tolerance of 0.02 Da was optimal for the number of proteins identified. Further investigation of the increase in number of proteins identified revealed a large contingency of peptides that were inaccurately assigned as deamidated. These misassigned deamidated peptides increased the number of proteins identified, resulting in falsely identified proteins. It is inherently important not only to use a mass spectrometer with high mass measurement accuracy and high resolving power but to also carefully consider search parameters in the bioinformatics software in order to avoid false positive protein identifications.

Materials and Methods

Animals

Animal testing was performed under Institutional Care and Use Committee regulations and approval at North Carolina State University. Mice were housed at Laboratory Animal Research facilities at the College of Veterinary Medicine. Mice were sacrificed at 5 days of age, and the brains were removed. Following the removal of the brain, the brains were washed with 0.1 M phosphate buffered saline (PBS) and were snap frozen in liquid nitrogen, weighed and then stored at –80°C.

Sample Preparation and Digestion

Lysis buffer at a ratio of 5 µL:1 mg of tissue was added to extract the proteins from the brain tissue. All reagents were purchased from Sigma Aldrich (St. Louis, MO, USA). The lysis buffer consisted of 50 mM Tris, 8 M Urea, 2 M Thiourea, 10 mM EDTA, 10 mM DTT, 0.001% sodium azide and was adjusted to a pH of 7.8 using hydrochloric acid. An OMNI TIP Homogenizing Kit (Omni International, Kennesaw, Georgia) was employed to break apart frozen tissue that was suspended in the lysis buffer. The homogenizer is a mechanical blade that was used for no longer than 1 minute to prevent the tissue from heating up. This was followed by adding sodium dodecyl sulfate (SDS) to make up a final concentration of 2% (w/v) and then placed onto the Genie Disrupter (Scientific Industries) for 1 minute followed by 5 minutes of incubation period on ice. This step was then repeated twice more. The sample was then centrifuged for 30 minutes at 14,000 × g. The soluble fraction was removed from the centrifuge tube without disrupting the cellular debris pellet or the top lipid layer.

The samples were digested following the filter-aided sample preparation (FASP) method published by Wisniewski and co-workers with minor changes.²¹ 21 µL of ~7 mM dithiothreitol (DTT) was added to 9 µL of brain tissue lysate for a final concentration of 5 mM DTT. The sample was incubated for 30 min at 56 °C to reduce the protein disulfide bonds. The sample was then mixed with 200 µL of 8 M urea in 0.1 M Tris/HCl (pH 8.5) and transferred onto a Vivacon 500 30 kDa MW cutoff filter (Vivaconproducts, Littleton, MA) and centrifuged at a constant 14,000g and 21 °C to prevent carbamylation using a

refrigerated benchtop centrifuge (Eppendorf, Hauppauge, NY) for 15 min. This step was repeated once more and then the flow-through solvent was discarded. Alkylation step was performed when 100 μ L of 0.05 M iodoacetamide was added and incubated for 20 min in the dark at room temperature. The filter was centrifuged for 10 min. Wash steps were performed with 100 μ L of 8 M urea (3 $\times$) and 100 μ L of 0.05 M ammonium bicarbonate in H₂O (3 $\times$), each step the filter was centrifuged for 10 min. The flow-through container vials were switched out with clean ones. 40 μ L of ammonium bicarbonate buffer and 2.5 μ L of 0.5 μ g/ μ L trypsin solution were added onto the filter. Trypsin digestion was performed for 16 hours at 37 °C. Following digestion, 40 μ L of 0.1% formic acid was added to the filter, and the peptides were eluted by centrifugation for 10 min at 14,000 $\times$ g. Protein concentration was determined using a NanoDrop 2000c (Thermo Scientific, Wilmington, DE) reading at 280 nm.

StageTip Anion Fractionation

A detailed stage tip procedure was used and is described by Gokce and co-workers.²² Briefly, a 200 μ L pipette tip (Eppendorf, Hauppauge, NY) was packed with six anion disks from an Empore C18 extraction disk (3M, St. Paul, MN). The peptides that were obtained from the FASP procedure were diluted to ~200 μ L by adding 100 μ L of Britton Robinson buffer pH 11 and were loaded on top of the anion stage tip. Peptides were eluted into a micro centrifuge tubes by centrifugation. Additional peptides were eluted from the anion stage tip using Britton Robinson buffers pH 8, 6, 5, 4, and 3 separately. Fractions were dried to completion and stored at -80 °C until analysis.

nano-flow Reversed Phase Chromatography

A Thermo Scientific EASY nLC II (Thermo Scientific, San Jose, CA) was coupled to a quadrupole Orbitrap benchtop mass spectrometer (Q-Exactive, San Jose, CA) using a vented column configuration.²³ The vented column consisted of a 75 μ m $\times$ 5 cm trap (IntegraFrit, New Objectives, Woburn, MA) coupled to a 75 μ m $\times$ 15 cm column (PicoFrit, New Objectives, Woburn, MA). Both the trap and column were packed in-house with Magic C18AQ stationary phase (5 μ m particle, 200 Å pore, Auburn, CA). Mobile phase A and B were composed of water/acetonitrile/formic acid (98/2/0.2% and 2/98/0.2% respectively). The dried samples were reconstituted to 0.08 μ g/ μ L concentration. 15 μ L of the sample was loaded onto the trap at 4.5 μ L/min with 100% mobile phase A. Flow was then diverted onto the column at the start of the gradient: 2% B (0–5 min), 2%–5% B (5–7 min), 5%–40% B (7–208 min), 40%–95% (208–218 min), 95% (218–228 min), 95%–2% (228–230 min), 2% (230–240 min).

Mass spectrometric analysis was performed on a Q-Exactive with optimized global proteomics parameters.²⁴ To summarize, MS transients in the orbitrap were acquired with 70 k_{FWHM} resolving power at m/z = 200. The automatic gain control (AGC) target for MS acquisitions was set to 1E6 with a maximum ion injection time of 30 ms. The scan range was set from 400 to 1600 m/z . Microscans were set to 1 for both the MS and MS/MS. Data dependent acquisition was set for 12 MS/MS spectra and the dynamic exclusion was set to 30 sec. The MS/MS resolving power was set to 17.5 k_{FWHM} at m/z = 200. MS/MS AGC target was set to 2E5 with a maximum ion injection time of 250 ms. All MS and MS/MS spectra were obtained in profile mode.

Data Analysis

Raw LC-MS/MS data files were processed into peak lists in a .MGF format using Proteome Discoverer. The resulting .MGF files were searched using MASCOT⁴ (Matrix Science, Boston, MA) against a concatenated target-reverse mus musculus database (SwissProt, Feb 2013). Carbamidomethyl (C) was set as a fixed modification, and oxidation (M) were set as

variable modifications and included a maximum of 2 missed cleavages. Data were also searched with and without deamidation (N and Q) as a variable modification. The precursor ion search tolerance were searched from 1 to 100 ppm, and the fragment ion tolerance was set to ± 0.02 Da, ± 0.6 Da, ± 0.005 Da. Statistical filtering using a 1% false discovery rate of the identified proteins were accomplished using ProteoIQ.^{25–27} Proteins that were identified with only 1 peptide were filtered from the total protein list.

Results and Discussions

To maximize the number of proteins identified, the mass tolerance for the precursor mass varied from 1 to 100 ppm, while MS/MS data was searched at ± 0.02 Da, 0.6 Da and 0.005 Da. Figure 1a depicts the overall trend of increasing proteins identified as the precursor ion mass tolerance reached 20 ppm. Subsequently, the data indicate that searching with ± 0.02 Da resulted in more identified proteins than 0.6 and 0.005 Da. All global proteomics experiments prior to this study were performed on an LTQ-Orbitrap or LTQ-FT system. Typical search criteria in our lab consisted of ± 5 ppm for precursor mass tolerance and ± 0.6 Da MS/MS tolerance when using LTQ detection for MS/MS sequencing. Since all MS and MS/MS spectra are obtained using the orbitrap on the Q Exactive, MS/MS data were searched with a more stringent tolerance of 0.02 Da (20 ppm at 1000 Da).

The observations from Figure 1a indicate that higher MS precursor mass tolerances results in more identified proteins. Data were further investigated by examining the identified peptides to determine the cause of the increase in identified proteins. Figure 1b illustrates the frequency of the mass accuracies of all identified peptides using ± 100 ppm precursor mass tolerance and 0.02 Da as the MS/MS tolerance. Figure 1b distinctly reveals there are two separate distributions from the 59,055 total peptides identified. One distribution of peptides has a precursor mass measurement accuracy (MMA) that falls between -5 and 5 ppm, and another between 5 to 20 ppm. The latter distribution had experimental measured masses that were systematically above the theoretical mass resulting in a higher MMA according to Equation 1.

$$MMA = \frac{Experimental - Theoretical}{Theoretical} \times 10^6 \quad \text{Equation 1}$$

It was initially thought that space-charge effects within the orbitrap may have been the cause for the overwhelming amount of peptides with a MMA greater than 5 ppm. Space-charge effects result from measuring too many ions in the orbitrap causing a systematic bias in mass measurements. The consequence of having too many ions in the orbitrap is ion-ion interactions, producing a shift in lower observed axial frequency (ω). The decrease in the observed axial frequency will result in an increase in observed m/z as shown in Equation 2. The increase in m/z will consequently result in the positive increase in mass error, similarly seen in Figure 1b. Space-charge effects are often compensated by using an AGC that limits the number of ions entering the orbitrap. To determine if the AGC was operating correctly, a set of deamidated peptides from the -5 to 5 ppm and 5 to 20 ppm range were chosen for further investigation to determine if the AGC was overpopulating the orbitrap. In several cases, the ionization time (IT) was maximized suggesting that the AGC limit was not reached. This suggests that the AGC was working properly by not allowing too many ions in the orbitrap thus preventing space-charge effects.

$$\omega = \sqrt{\frac{q}{m} k} \quad \text{Equation 2}$$

Upon further investigation of the 44,392 peptides between -5 and 5 ppm, 7.2% of them were deamidated (N,Q). The percent of deamidated peptides between 5 to 20 ppm was calculated to be much higher at 95.7% deamidated (N,Q). Figure 1c is the same data but searched without deamidation (N,Q) as a variable modification. Clearly the second distribution in the $5 - 20$ ppm range has noticeably disappeared. Similar distributions were identified in whole cell yeast lysates by Venable and coworkers.²⁸ They attribute a second distribution to deamidation of asparagine and glutamine residues which are identified by the $+ 0.98$ Da mass shift. Analyses of the yeast lysate were performed on a LTQ mass spectrometer measuring their mass error in Daltons. Distributions similar to this were also recognized in other datasets obtained by our laboratory (data not shown).

To further investigate the reason why there was a larger percentage of deamidated peptides found in the $5-20$ ppm range, several peptides with mass measurement accuracy (MMA) between $5-20$ ppm were examined closely. Figure 2a depicts the identification of a deamidated (left) and amidated (right) form of the same peptide within the same chromatographic run. This peptide was selected because it was a deamidated peptide that has a mass accuracy of greater than 5 ppm and the corresponding amidated form was also identified in the same chromatogram. A precursor spectrum is shown on the left hand side in which the 1060.0278 m/z was identified and selected for MS/MS. The (♦) above the peak is the m/z that was chosen for MS/MS. Below is the corresponding MS/MS spectrum for the 1060.0278 m/z . MASCOT identified the 1060.0278 m/z as the deamidated form for the peptide sequence shown in the Figure 2a. Investigation of the MS/MS spectra (below) indicates that there are residual precursor ions detected. An isotopic cluster with a monoisotopic peak at 1059.52466 m/z is clearly observed. This monoisotopic peak would correspond to the amidated form of the peptide. Also enhanced in the MS/MS spectrum is that of the y^{15} fragment (1633.790 m/z) of the peptide, which also illustrates that the fragment would be generated from the amidated form of the peptide. The overwhelming evidence from the MS/MS spectrum indicates that the peptide should be the amidated form. The mass accuracy of the deamidated precursor ion was calculated by MASCOT to be 8.75 ppm. The inability to detect the monoisotopic peak can be attributed to the fact the precursor spectrum obtained was at the beginning of the elution profile where the signal intensity is low. Based on the residual precursor ion mass from the MS spectrum, the peptide should be the deamidated form. However, the product ions observed in the MS/MS spectrum suggests the amidated form is present. Further, the mass accuracy of the monoisotopic peak for an amidated peptide was -1.95 ppm which falls within the expected mass measurement accuracy range for orbitrap analyzers.

The right column in Figure 2a confirms the amidated peptide from both the MS and MS/MS spectra used for the identification, the precursor mass accuracy was calculated to be -0.11 ppm. In this example, the monoisotopic peak that was detected and selected for MS/MS (♦) corresponds to that of the unmodified form. The intact precursor ion agrees with the unmodified form. The mass accuracy of the precursor obtained from the MS/MS spectra is calculated to be -0.57 ppm. It can be concluded that the deamidated peptide was inaccurately assigned, and the 1060.0278 peak is truly the $M+1$ peak of the amidated form. Figure 2b are the selected ion chromatograms obtained for the monoisotopic peak of the amidated peptide (1059.5266) and the monoisotopic peak (1060.0278) of the “deamidated” peptide which show a lack of any retention time shift. Both chromatograms are similar; yet researchers have shown that deamidated and amidated peptides have different elution profiles.^{10, 17, 29} This suggests that the “deamidated” and amidated peptide is actually the same peptide.

Figure 3 is another example of the same peptide identified as both deamidated and amidated. Figure 3a corresponds to the identification of the deamidated form of the peptide and the

precursor mass spectrum. Enhanced is the isotopic cluster in which the 1082.03308 (♦) peak was chosen as the monoisotopic peak and selected for MS/MS. Since the 1082.03308 peak was chosen as the “monoisotopic mass” during data acquisition, it was then associated as the deamidated peptide resulting in a 10.9 ppm mass error because this was within the specified mass error tolerance of 100 ppm when searching the data against the protein database. Based on the isotopic distribution, the 1081.53149 should have been assigned as the monoisotopic peak. Similar to the previous example, the MS/MS spectrum (not shown) had fragment ions associated with the amidated version. Additionally, there are no fragments ions that would suggest that the deamidated form is present. Figure 3b is of the precursor that was used to identify the amidated form of the peptide. The mass accuracy calculated for the monoisotopic peak of the cluster above is 0.36 ppm. In the enhanced view of the isotopic cluster, the monoisotopic peak 1081.52966 was the correct peak to be selected for MS/MS. Fragmentation spectrum indicates fragments corresponding to the amidated form. And similar to the previous example, the elution profiles (bottom) are overlapping. Since the wrong precursor peak was chosen as the monoisotopic peak and a wide mass accuracy tolerance was used, MASCOT was able to assign it as the deamidated peptide despite the fact that the fragmentation spectrum will have all amidated fragments. However, since the M +1 peaks of all the amidated fragments will correlate to the theoretical monoisotopic deamidated fragment peaks, falsely identified deamidated peptides can result if using a relatively wide mass error tolerance. Taken together, the MS, MS/MS spectra and the elution profile of the peptide conclude the peptide as amidated in the above examples.

Deamidated peptides that fell in the range of -5 to 5 ppm were also inspected. Figure 4 illustrates a deamidated peptide correctly identified from MASCOT. The precursor spectrum is shown on the left with an enhanced view of the isotopic cluster for the peptide. From the isotopic cluster it was determined that the 652.29657 peak was the monoisotopic peak. The monoisotopic peak corresponded to that of the deamidated peptide (SMWSVNGDSISK) with a mass measurement accuracy of -2.04 ppm. MS/MS was performed on the 652.29657 peak (♦) and the corresponding MS/MS spectrum is shown on the right. The inset of the MS/MS spectrum is the isotopic cluster corresponding to the y₇ fragment ion. Based on the isotopic cluster of the y₇ fragment ion, it can be determined that this is a deamidated form of the peptide. The mass accuracy of the fragment ion monoisotopic peak is 2.67 ppm. Since the amidated counterpart was not detected, elution profiles could not be compared. Evidence obtained from the MS and MS/MS data as well as a mass measurement error <5 ppm increases the confidence that a deamidated peptide was accurately identified.

It has been reported that deamidated and amidated peptides will elute at different times depending on the type of chromatography used.^{17, 29} An example of a peptide that was found to be both deamidated and amidated within the ± 5 ppm range were examined. Figure 5 illustrates the selected ion chromatograms of the identified deamidated and amidated peptide DISTNYYASQKK. The top shows that the deamidated peptide elutes almost 3 minutes prior to that of the amidated form. The mass errors for both peptides were within ± 3 ppm. The combination of unique elution times and high mass measurement accuracies of both forms of the peptide, and supporting evidence from MS/MS spectra, it can be assumed that the deamidated peptide was correctly assigned.

The accurate identification of the monoisotopic peak obtained in the precursor spectrum is inherently important to the identification deamidated peptides. Proteome Discover was initially used for peak picking procedure for producing MGF files needed for MASCOT searching. As a result, several other peak picking programs were also tested such as MASCOT Distiller and Mass Matrix to evaluate if these programs corrected misidentified precursor peak values. MGF files generated from these programs were searched in MASCOT using ± 100 ppm precursor mass tolerance. The datasets showed similar

systematic mass errors for peptides ranging from 5–20 ppm (data not shown). However, it is likely that the selection of the monoisotopic peak is a result of data dependent acquisition (DDA) software. Acquiring in a DDA mode, a precursor spectrum is taken populating a candidate peak list of m/z values to be used for the following MS/MS spectra. In our study 12 sequential m/z values following the precursor spectrum were targeted for MS/MS. Once a candidate was chosen for MS/MS it was no longer a candidate for MS/MS for 30 seconds, which is often the elution time of a peptide. A consequence of using DDA mode, often a candidate is picked during the beginning of its elution profile. It then becomes a possibility to have an incomplete isotopic distribution present in the precursor spectrum shown in Figure 3. This results in the misidentification of the monoisotopic peak.

To determine if searching with or without deamidation significantly affects the number of proteins identified, the same data was searched without deamidation included. The number of total proteins identified using search parameters with and without deamidation as a variable modification is presented in Table 1, the MS tolerances are shown for 5, 20, and 100 ppm. At 5 ppm the total number of proteins identified without deamidation was 4933, including deamidation in the search parameter resulted in 4822 identified proteins. When variable modifications are searched, this further complicates the searching algorithm by increasing the potential m/z space that peptides can occupy. The search space is enlarged within the database because both modified and unmodified forms can potentially exist. Deamidation commonly occurs on 2 different residues which further enlarges the search space. However, when deamidation is included in the search parameters the number of identified proteins is greater using 20 ppm and 100 ppm as the MS tolerance. This is best explained by the misidentified deamidated peptides that fall largely in the 5 to 20 ppm range.

Label free quantitative proteomic techniques involve either the use of area under the curve³⁰ and/or spectral counting^{2, 31}. The primary method used in our laboratory for relative quantitation is spectral counting. To determine if searching without deamidation has an effect on quantitative global proteomics, we compared spectral counts from searches with and without deamidation. Average normalized spectral counts of proteins found (using 5 ppm MS tolerance) with and without deamidation were compared. No significant differences found in normalized spectral counts (NSpC) between proteins searched with or without deamidation. Searching data without deamidation will not ultimately affect the relative quantification of proteins in global proteomic datasets.

Conclusion

Deamidation is a variable modification parameter often searched because it is a nonenzymatic modification and can occur during sample preparation under basic conditions. Tryptic digestion protocols often involve a pH range between 7–8 which can induce the deamidation of asparagine and glutamine sites. Instrumentation with low resolving power and low mass measurement accuracy can be problematic in identifying the mass shift of 0.98 Da caused by deamidation. For accurate determination of deamidated sites, mass tolerance parameters must be set within 5 ppm for the precursor ion and high resolving power instruments must be used to acquire data. Using wider mass tolerances can lead to the misassignment of these peptides. Omitting deamidation as a variable modification did not significantly impact global quantification data and led to an increase in protein identifications when searching within ± 5 ppm MS window and a 0.02 Da MS/MS tolerance.

Acknowledgments

The authors would like to thank Dr. Hunter Walker for his helpful discussions. The authors thank Dr. Troy Ghashghaei and Dr. Nagendran Muthusamy from NCSU College of Veterinary Medicine for providing the mouse tissue samples. This work has been supported by National Institutes of Health grant R01GM08764 and North Carolina State University.

References

1. Cravatt BF, Simon GM, Yates JR. 3rd, The biological impact of mass-spectrometry-based proteomics. *Nature*. 2007; 450(7172):991–1000. [PubMed: 18075578]
2. Eng JK, McCormack AL, Yates JR. An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database. *Journal of the American Society for Mass Spectrometry*. 1994; 5(11):976–989. [PubMed: 24226387]
3. Yates JR, Eng JK, McCormack AL, Schieltz D. Method to Correlate Tandem Mass-Spectra of Modified Peptides to Amino-Acid-Sequences in the Protein Database. *Analytical chemistry*. 1995; 67(8):1426–1436. [PubMed: 7741214]
4. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 1999; 20(18):3551–3567. [PubMed: 10612281]
5. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 2004; 20(9):1466–1467. [PubMed: 14976030]
6. Michalski A, Damoc E, Hauschild JP, Lange O, Wiegand A, Makarov A, Nagaraj N, Cox J, Mann M, Horning S. Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Molecular & cellular proteomics : MCP*. 2011; 10(9):M111 011015. [PubMed: 21642640]
7. Kelstrup CD, Young C, Lavalley R, Nielsen ML, Olsen JV. Optimized Fast and Sensitive Acquisition Methods for Shotgun Proteomics on a Quadrupole Orbitrap Mass Spectrometer. *Journal of Proteome Research*. 2012
8. Jung HJ, Purvine SO, Kim H, Petyuk VA, Hyung SW, Monroe ME, Mun DG, Kim KC, Park JM, Kim SJ, Tolic N, Slys GW, Moore RJ, Zhao R, Adkins JN, Anderson GA, Lee H, Camp DG 2nd, Yu MH, Smith RD, Lee SW. Integrated post-experiment monoisotopic mass refinement: an integrated approach to accurately assign monoisotopic precursor masses to tandem mass spectrometric data. *Analytical chemistry*. 2010; 82(20):8510–8. [PubMed: 20863060]
9. Hsieh EJ, Hoopmann MR, MacLean B, MacCoss MJ. Comparison of database search strategies for high precursor mass accuracy MS/MS data. *Journal of Proteome Research*. 2010; 9(2):1138–43. [PubMed: 19938873]
10. Hao P, Ren Y, Alpert AJ, Sze SK. Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. *Molecular & cellular proteomics : MCP*. 2011; 10(10):O111 009381. [PubMed: 21784994]
11. Shuford CM, Li QZ, Sun YH, Chen HC, Wang J, Shi R, Sederoff RR, Chiang VL, Muddiman DC. Comprehensive Quantification of Monolignol-Pathway Enzymes in *Populus trichocarpa* by Protein Cleavage Isotope Dilution Mass Spectrometry. *Journal of Proteome Research*. 2012; 11(6):3390–3404. [PubMed: 22524869]
12. Ren D, Pipes GD, Liu DJ, Shih LY, Nichols AC, Treuheit MJ, Brems DN, Bondarenko PV. An improved trypsin digestion method minimizes digestion-induced modifications on proteins. *Analytical Biochemistry*. 2009; 392(1):12–21. [PubMed: 19457431]
13. Hong J, Schoneich C. The metal-catalyzed oxidation of methionine in peptides by Fenton systems involves two consecutive one-electron oxidation processes. *Free radical biology & medicine*. 2001; 31(11):1432–41. [PubMed: 11728815]
14. Morand K, Talbo G, Mann M. Oxidation of peptides during electrospray ionization. *Rapid communications in mass spectrometry : RCM*. 1993; 7(8):738–43. [PubMed: 8374164]
15. Robinson NE, Robinson AB. Molecular clocks. *Proceedings of the National Academy of Sciences of the United States of America*. 2001; 98(3):944–9. [PubMed: 11158575]

16. Robinson NE, Robinson AB. Amide molecular clocks in drosophila proteins: potential regulators of aging and other processes. *Mechanisms of Ageing and Development*. 2004; 125(4):259–67. [PubMed: 15063101]
17. Dasari S, Wilmarth PA, Rustvold DL, Riviere MA, Nagalla SR, David LL. Reliable detection of deamidated peptides from lens crystallin proteins using changes in reversed-phase elution times and parent ion masses. *Journal of Proteome Research*. 2007; 6(9):3819–26. [PubMed: 17696381]
18. Robinson NE, Robinson AB. Deamidation of human proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 2001; 98(22):12409–13. [PubMed: 11606750]
19. Li XJ, Cournoyer JJ, Lin C, O'Connor PB. Use of O-18 labels to monitor deamidation during protein and peptide sample processing. *Journal of the American Society for Mass Spectrometry*. 2008; 19(6):855–864. [PubMed: 18394920]
20. Krokhn OV, Antonovici M, Ens W, Wilkins JA, Standing KG. Deamidation of -Asn-Gly-sequences during sample preparation for proteomics: Consequences for MALDI and HPLC-MALDI analysis. *Analytical chemistry*. 2006; 78(18):6645–50. [PubMed: 16970346]
21. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nature Methods*. 2009; 6(5):359–U60. [PubMed: 19377485]
22. Gokce E, Franck WL, Oh Y, Dean RA, Muddiman DC. In-Depth Analysis of the Magnaporthe oryzae Conidial Proteome. *Journal of Proteome Research*. 2012; 11(12):5827–5835. [PubMed: 23039028]
23. Andrews GL, Shuford CM, Burnett JC Jr, Hawkrigde AM, Muddiman DC. Coupling of a vented column with splitless nanoRPLC-ESI-MS for the improved separation and detection of brain natriuretic peptide-32 and its proteolytic peptides. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences*. 2009; 877(10):948–54.
24. Randall S, Cardasis H, Muddiman D. Factorial Experimental Designs Elucidate Significant Variables Affecting Data Acquisition on a Quadrupole Orbitrap Mass Spectrometer. *Journal of the American Society for Mass Spectrometry*. 2013:1–12.
25. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical chemistry*. 2002; 74(20):5383–5392. [PubMed: 12403597]
26. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. *Analytical chemistry*. 2003; 75(17):4646–4658. [PubMed: 14632076]
27. Weatherly DB, Atwood JA, Minning TA, Cavola C, Tarleton RL, Orlando R. A heuristic method for assigning a false-discovery rate for protein identifications from mascot database search results. *Molecular & Cellular Proteomics*. 2005; 4(6):762–772. [PubMed: 15703444]
28. Venable JD, Xu T, Cociorva D, Yates JR. 3rd, Cross-correlation algorithm for calculation of peptide molecular weight from tandem mass spectra. *Analytical chemistry*. 2006; 78(6):1921–9. [PubMed: 16536429]
29. Krokhn OV, Ying S, Cortens JP, Ghosh D, Spicer V, Ens W, Standing KG, Beavis RC, Wilkins JA. Use of peptide retention time prediction for protein identification by off-line reversedphase HPLC-MALDI MS/MS. *Analytical chemistry*. 2006; 78(17):6265–6269. [PubMed: 16944911]
30. Wiener MC, Sachs JR, Deyanova EG, Yates NA. Differential Mass Spectrometry: A Label-Free LC-MS Method for Finding Significant Differences in Complex Peptide and Protein Mixtures. *Analytical chemistry*. 2004; 76(20):6085–6096. [PubMed: 15481957]
31. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, Resing KA, Ahn NG. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Molecular & cellular proteomics : MCP*. 2005; 4(10):1487–502. [PubMed: 15979981]

Figure 1.

(a) A plot of the number of identified proteins with increasing precursor peptide mass tolerance. Data was searched at three separate values for the MS/MS mass tolerances, 0.06 Da, 0.02 Da, and 0.005 Da. 0.02 Da provided the greatest number of proteins identified at all precursor peptide mass tolerances. (b) Histogram of the mass accuracy of all peptides that were identified when searching data with Deamidation as a variable modification, $\pm$ 100 ppm precursor mass tolerance, and $\pm$ 0.02 Da MS/MS mass tolerance. There are two observable distributions, the second distribution which ranges from 5–20 ppm illustrates a large number of deamidated peptides. (c) Deamidation was not included in the search parameters, clearly proving that the second distribution is a result from using deamidation as a variable modification.

Figure 2.

(a) A deamidated and amidated form of the AIGKDNFTAPEGTNGVEER peptide were identified from the MASCOT results. On the left side, the monoisotopic peak of 1060.02783 was established to be the deamidated form due to the mass accuracy of 8.75 ppm. From the MS/MS spectrum below, it is alluded that the deamidated peptide is in fact misassigned. The right illustrates the identification of the amidated peptide. And below is the MS/MS spectra that is comparable to that of the deamidated peptide. (b) Are the selected ion chromatograms for each peptide.

Figure 3.

(a) Another example of a deamidated peptide that has a mass measurement accuracy of 10.91 ppm which results in a misidentification. The software chose the wrong peak (♦), in doing so it results in the identification of a deamidated peptide. (b) The correct amidated peptide was identified since the correct monoisotopic peak was used. (Bottom) Shows the selected ion chromatogram for both the deamidated and amidated peptide.

Figure 4.

The correct identification of a deamidated peptide. The correct monoisotopic peak was chosen leading to a -2.04 mass error. The peptide was then confirmed from the MS/MS spectrum (right). The y₇ fragment ion confirms that it is a deamidated peptide.

SIC – 709.8441 Deamidated Peptide

SIC – 709.3497 Amidated Peptide

Figure 5. Selection ion chromatograms for both the deamidated (top) and amidated (bottom) peptide. Different elution profiles indicate that both are identified correctly.

Table 1

Searching With Deamidation (N and Q) Variable Modification		Searching Without Deamidation (N and Q) Variable Modification	
Search Parameters	Number of Proteins Identified	Search Parameters	Number of Proteins Identified
±5 ppm Precursor MS Tolerance	4822	±5 ppm Precursor MS Tolerance	4933
±20 ppm Precursor MS Tolerance	5049	±20 ppm Precursor MS Tolerance	4907
±100 ppm Precursor MS Tolerance	4937	±100 ppm Precursor MS Tolerance	4832

Comparison of the number of proteins identified when searching with (left) and without (right) deamidation (NQ) as a variable modification. The total number of proteins identified is shown when using either ± 5, 20 and 100 ppm as a precursor mass tolerance. All data was searched with ± 0.02 Da MS/MS tolerance.