Technical Note

Optimizing sequence coverage for a moderate mass protein in nano-electrospray ionization quadrupole time-of-flight mass spectrometry

Ryan Matsuda, Venkata Kolli, Megan Woods, Eric D. Dodds, David S. Hage*

Department of Chemistry, University of Nebraska—Lincoln, Lincoln, NE 68588, USA

A R T I C L E   I N F O

Article history:
Received 21 April 2016
Received in revised form 7 June 2016
Accepted 8 June 2016
Available online 16 June 2016

Keywords:
Human serum albumin
Sequence coverage
Sample pretreatment
Nano-electrospray ionization quadrupole time-of-flight mass spectrometry
Collision-induced dissociation
Internal calibration

A B S T R A C T

Sample pretreatment was optimized to obtain high sequence coverage for human serum albumin (HSA, 66.5 kDa) when using nano-electrospray ionization quadrupole time-of-flight mass spectrometry (nESI-Q-TOF–MS). Use of the final method with trypsin, Lys-C, and Glu-C digests gave a combined coverage of 98.8%. The addition of peptide fractionation resulted in 99.7% coverage. These results were comparable to those obtained previously with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI–TOF–MS). The sample pretreatment/nESI–Q-TOF–MS method was also used with collision-induced dissociation to analyze HSA digests and to identify peptides that could be employed as internal mass calibrants in future studies of modifications to HSA.

© 2016 Elsevier Inc. All rights reserved.

Mass spectrometry (MS) is an important tool for identifying proteins and providing information on the primary sequences of proteins and their post-translational modifications (PTMs) [1,2]. Several recent studies have used MS to examine human serum albumin (HSA, 66.5 kDa) and PTMs such as non-enzymatic glycation that occur on this protein [3–5]. The effects of such modifications have been of concern given the important role played by HSA in transporting many low mass substances in blood (e.g., hormones, fatty acids, drugs) [6].

Several past studies have used matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI–TOF–MS) to investigate the structure of HSA and glycated HSA [2,3,7–12]. Sequence coverages of up to 97.4% have been reached by optimizing the MALDI matrix and conditions that were used for sample pretreatment; however, obtaining this level of coverage required multiple enzyme digestions and a series of steps for peptide fractionation [2]. The purpose of this current study was to determine whether more convenient pretreatment conditions could still obtain high sequence coverage for a moderately high mass protein such as HSA and when using nano-electrospray ionization quadrupole time-of-flight mass spectrometry (nESI–Q-TOF–MS) for analysis of the digests.

Two pretreatment procedures were initially considered (see online supplementary material for more details). Pretreatment method 1 followed a scheme similar to the previous method used with MALDI–TOF–MS [2] but without any peptide fractionation. In this method, a sample of HSA was first treated by means of alkylation and reduction, followed by dialysis to remove any excess reagents. The protein sample was then digested separately with three proteolytic enzymes—trypsin, Lys-C, and Glu-C—and analyzed by MS. In pretreatment method 2, all of the pretreatment and digestion processes were performed sequentially in a single vial, resulting in the ability to use lower concentrations of both the protein and pretreatment reagents. The pretreatment conditions for alkylation and reduction were similar to those in pretreatment method 1 with some slight variations in the reaction times and temperatures. Also in pretreatment method 2, dialysis was no longer used to remove the excess pretreatment

Abbreviations used: MS, mass spectrometry; PTM, post-translational modification; HSA, human serum albumin; MALDI–TOF–MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; nESI–Q-TOF–MS, nano-electrospray ionization quadrupole time-of-flight mass spectrometry; MS/MS, tandem mass spectrometry; CID, collision-induced dissociation.

* Corresponding author.
E-mail address: dthage1@unl.edu (D.S. Hage).

http://dx.doi.org/10.1016/j.ab.2016.06.014
0003-2697/© 2016 Elsevier Inc. All rights reserved.
residues 535 to 541. This coverage included 56 of 59 lysines in HSA.

In the previous work with HSA and MALDI–TOF–MS [2], 0.5-μl portions of the fractionated protein digests (protein concentrations: ~0.822 μg/μl in the tryptic digest, ~0.870 μg/μl in the Lys-C digest, and ~1.25 μg/μl in the Glu-C digest) were individually mixed with the MALDI matrix and spotted onto a plate for MS analysis. These multiple fractions increased the overall amount of reagents, and each digest was fractionated by using ZipTip-C₁₈ pipette tips.

Fig. 1 and the supplementary material summarize the results that were obtained by nESI–Q–TOF–MS for digests of HSA when using these pretreatment methods. The use of pretreatment method 1 with a trypsinic digest gave 42 identified peptides and a sequence coverage of 36.9% for HSA, whereas the Glu-C and Lys-C digests gave sequence coverages of 39.1 and 32.1% and 43 and 21 identified peptides, respectively. For pretreatment method 2, the sequence coverages were 86.0% for the trypsinic digest, 84.1% for the Lys-C digest, and 78.2% for the Glu-C digest, which corresponded to 104, 52, and 54 identified peptides in these digests. Because there was both improved sequence coverage and an increase in the number of identified peptides with pretreatment method 2, this method was selected for use in the remainder of this study.

The results obtained by pretreatment method 2, nESI–Q–TOF–MS, and using a single elution step for peptide fractionation were compared with the results from the previous pretreatment method that had been used with MALDI–TOF–MS to examine HSA (see supplementary material) [2]. The previous MALDI–TOF–MS method gave sequence coverages of 82.9, 77.6, and 64.6% and 54, 21, and 21 identified peptides when HSA was digested with trypsin, Lys-C, and Glu-C, respectively, and these digests were fractionated on ZipTip-C₁₈ tips through a series of several elution steps [2]. The approach based on pretreatment method 2 and nESI–Q–TOF–MS, which used a simpler alkylation/reduction protocol and only a single elution step for peptide fractionation, gave 3.1–13.6% higher sequence coverages for these digests and a greater number of identifiable peaks (i.e., 11–48 more peptides per digest). The use of multiple elution steps for peptide fractionation, as described in the supplementary material, gave a small increase in the sequence coverages for pretreatment method 2 and nESI–Q–TOF–MS to levels of 83.2 and 81.5% for the tryptic and Glu-C digests and a slight decrease to 72.6% for the Lys-C digest.

When a single elution step was used for peptide fractionation in pretreatment method 2, the total sequence coverage obtained by nESI–Q–TOF–MS for HSA was 98.8%. This coverage included 578 of the 585 amino acids in HSA, with the only excluded region being residues 535 to 541. This coverage included 56 of 59 lysines in HSA and all of the arginine residues, which were of interest because lysine and arginine are both potential sites for glycation-related modifications (the 3 lysines in residues 535–541 that were not included in this coverage are not major sites for such modifications) [3]. When multiple elution steps were used for peptide fractionation, the combined digests gave an overall sequence coverage of 99.7%, in which 583 of 585 residues were now covered. The only missing residues in this second case came from residues 312 and 313; however, previous reports have shown that K313 is highly susceptible to modification by glycation [3].

Both of these situations were an improvement over the previous pretreatment method used with MALDI–TOF–MS, in which a total sequence coverage of 97.4% was obtained for HSA when data were combined for the three enzyme digests [2]. In this prior work, 570 of the 585 amino acids were included as well as 54 of the 59 lysines and all of the arginines. In the missed regions for this previous method, K4, K536, K538, and K541 are not major sites for glycation-related modifications, but K317 can be an important location for such PTMs [3]. Neither the previous MALDI–TOF–MS method nor the use of pretreatment method 2 with nESI–Q–TOF–MS and a single elution step for fractionation covered the region 535 to 541 of HSA. However, if the data were combined for nESI–Q–TOF–MS and pretreatment method 2 when using both a single step and multiple steps for peptide fractionation, then all 585 amino acids of HSA were included and effectively 100% coverage was achieved.
sample that was needed for analysis. Although a 20-μl sample of the purified protein digest (protein concentration: ~1 μg/μl in each digest) was prepared for use in pretreatment method 2 and nESI–Q-TOF–MS, the sample could be placed directly into the nano-ESI emitter and sprayed in nanoliter volumes to obtain a mass spectrum. In addition, the remaining sample from the nano-ESI emitter could be recovered and stored for future experiments.

Another advantage of the nESI–Q-TOF–MS method over MALDI–TOF–MS was its ability to perform tandem mass spectrometry (MS/MS) experiments on the digests such as by using collision-induced dissociation (CID) to create peptide fragmentation [13]. This type of experiment was performed on the various digests of HSA to confirm the identity of the peptides, as demonstrated in Fig. 2A. Some of these peptides were also tested for use as internal m/z calibrants to improve mass accuracy when searching for PTMs [14,15] such as glycation-related modifications [2,3,7–10], as illustrated in Fig. 2B. Approximately 6 or 7 unique peptides were selected through these experiments from each type of enzyme digest of HSA, with each peptide having a relatively high signal intensity (see supplementary material). One of these peptides is shown in Fig. 2A and corresponded to residues 337 to 348 (RHPDYSVVLLLR); this peptide was selected because it contained no lysine residues and would not be subject to PTMs at such sites. Other peptides were chosen because they appeared in a variety of regions in the mass spectra.

These peptides were next used as internal calibrants in the analysis of a separate set of HSA digests that were prepared by using trypsin, Lys-C, and Glu-C. A calibrated spectrum of a Glu-C digest of HSA resulted in a root mean square mass error of 19.9 ppm, with similar results being obtained with the trypsin and Lys-C digests. This low mass measurement error made it possible to lower the mass tolerance that was used for sequence analysis (i.e., a previous value of 50 ppm); this change had the advantage of reducing the possibility of false positive matches while also improving confidence in the assignment of putative peptides, including those harboring PTMs. The effect of this change on the sequence coverage was evaluated by using the internally calibrated mass spectra for the tryptic, Lys-C, and Glu-C digests along with a 25-ppm mass tolerance. Under these conditions, the individual digests had sequence coverages of 79.8, 86.5, and 78.9% for the tryptic, Lys-C, and Glu-C digests, respectively, with a 96.1% total coverage for the combined results. These results were comparable to the previous values acquired by MALDI–TOF–MS [2] and were only slightly lower than the coverages seen earlier with nESI–Q-TOF–MS when using a 50-ppm mass tolerance and no internal calibration.

In summary, a sample pretreatment method was developed for use with nESI–Q-TOF–MS to obtain high sequence coverage for HSA. The combined use of nESI–Q-TOF–MS and the final pretreatment method with single or multiple elution steps for peptide fractionation gave sequence coverages of 98.8—99.7% for HSA, with a combined coverage of essentially 100%. It was also shown how this combined approach could be used with CID to confirm the structures of the peptides and to select some of these peptides for use as internal calibrants in MS/MS experiments. Although this study focused on the analysis of HSA by nESI–Q-TOF–MS, the same methods could be applied to other proteins with moderately high masses.

Acknowledgments

This work was funded by the National Institutes of Health under grant R01 DK069629. R. Matsuda was supported under a fellowship through the Molecular Mechanisms of Disease program at the University of Nebraska.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ab.2016.06.014.

References