APPLICATION NOTE

Discerning Matrix-Cluster Peaks in Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectra of Dilute Peptide Mixtures

Bernd O. Keller and Liang Li
Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada

Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry is widely used for the analysis of peptide mixtures such as those resulting from protein digestion. Among several useful peptide matrices, α-cyano-4-hydroxycinnamic acid (4-HCCA) appears to be the most popular. This matrix does not generally give matrix-cluster peaks at the mass region covered by enzyme-digested peptides (i.e., *m/z* above ~500). However, when an analyte mixture is very dilute and/or the sample contains a large amount of salts, ion peaks from matrix clusters can be quite intense, compared to peptide peaks. This matrix-cluster interference becomes more pronounced as the amount of analyte decreases. In this paper, a simple scheme for matrix-cluster identification is reported. It is shown that matrix-cluster formation follows a systematic pattern, although the relative intensities of these cluster ions cannot be predicted. Discerning the matrix-cluster peaks from the peptide peaks is found to be critical in analyzing dilute peptide mixtures with both conventional and microspot MALDI-TOF techniques. (J Am Soc Mass Spectrom 2000, 11, 88–93) © 2000 American Society for Mass Spectrometry

Peptide mass fingerprinting based on accurate mass measurement of enzyme-digested peptides has become an important method for identification of proteins in a database [1]. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry has been widely used for this purpose. The quality of MALDI mass spectra has a direct influence on database searchability and the outcome of the data search. It is desirable to obtain mass spectra that display a number of peptide peaks reflecting the specific digestion of a given enzyme. Aside from optimization of the enzyme digestion process to generate enzyme-specific peptide fragments, the sample/matrix preparation conditions for MALDI analysis of the resulting digest is also very critical in obtaining useful mass spectra. This is particularly true in the case of dealing with a small amount of starting protein samples. Because the digestion process itself involves the use of a relatively large amount of salts and buffers, the protein digest even from a purified protein often contaminates with salts. The introduction of a purification step prior to MALDI analysis is very useful in obtaining high quality mass spectra of protein digests. However, when the amount of the starting protein sample is low, such as in the cases of low volume protein fractions collected from capillary electrophoresis and low abundant protein spots in 2D gel, a rigorous purification step may not always be applicable. This is especially true for highly sensitive techniques where the analyte solution is directly deposited onto the matrix layer and thus most of the analyte resides in a very thin layer closer to the matrix surface. Several techniques using direct deposition have been reported in recent years [2–8]. The issue of sample loss versus purification must be carefully weighed. Even with compatible purification, the salt content is higher for protein digests resulting from a small amount of proteins, compared to those obtained from a large amount of proteins. Thus the analysis of protein digests from a small amount of proteins is prone to chemical background interference.

One of the major sources of background interference is from the matrix-cluster formation. Many experimental factors can affect the extent of matrix-cluster formation. These include the type of matrix used, the salt content, pH, the relative amount of analyte and the laser power. For the analysis of protein digests, α-cyano-4-hydroxycinnamic acid (4-HCCA) is widely used as the matrix for its ability to generate mass spectra with high sensitivity and resolution [9]. For the analysis of a protein digest resulting from a picomole or higher amount of proteins, mass spectra of the digest usually do not display strong matrix peaks with *m/z* above 500. Herein we show that matrix clusters can cause interference in analyzing *dilute* peptide mixtures, i.e., typically

Address reprint requests to Liang Li, Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada. E-mail: Liang.Li@ualberta.ca

1044-0305/00/$20.00
PII S1044-0305(99)00126-9

Received August 23, 1999
Revised October 4, 1999
Accepted October 4, 1999

in the medium to low nanomolar range depending on the extent of contamination. The spectral pattern of cluster ions and their relative intensities cannot be readily predicted, causing difficulty in discerning these peaks from the actual peptide peaks. However, it is shown that these clusters are mainly from the combination of matrix, potassium, and sodium ions. Thus their masses can be calculated by using a simple scheme. Knowing the masses of the possible cluster ions, the peaks from these ions can be discerned from the peptide peaks.

Experimental

Mass spectra of protein digests and matrix clusters are collected on a home-built linear time-lag focusing MALDI-TOF mass spectrometer. This instrument has been described in detail elsewhere [10]. A 337-nm and 3-ns pulse width laser beam from a nitrogen laser (model VSL 337ND, Laser Sciences, Newton, MA) was used for desorption. In general, 150–200 laser shots (3–5 μJ pulse energy) are averaged to produce a mass spectrum. Spectra are acquired and processed with Hewlett-Packard supporting software and reprocessed with the Igor Pro software package (WaveMetrics, Inc Lake Oswego, OR). Each spectrum is normalized using the most intense signal.

Bovine lactoferrin (MW ~80 kDa), dithiothreitol (DTT), iodoacetamide, 2,5 dihydroxybenzoic acid (DHB), α-cyano-4-hydroxycinnamic acid (4-HCCA), and trypsin (98%, L-1-tosylamide-2-phenylethyl chloromethyl ketone treated for reduction of chymotrypsin activity) were from Sigma-Aldrich Canada (Oakville, Ontario). 4-HCCA was recrystallized from ethanol (95%) at 50 °C.

For the macro digest of lactoferrin, 13 μL of 1.4 μg/μL lactoferrin is pipetted into a 0.5 mL siliconized polypropylene sample vial (Rose Scientific, Edmonton, Alberta, Canada). 0.6 μL of 1 M NH_4HCO_3 and 0.6 μL of 250 mM DTT are added. After incubation for 25 min at 37 °C, the vial is cooled down to room temperature and 0.6 μL of 400 mM iodoacetamide is added. After incubation in the dark for another 25 min at room temperature, 0.2 μL of a 3.5 μg/μL freshly prepared trypsin stock solution are added and the resulting 15 μM lactoferrin solution in 0.04 M digestion buffer incubated overnight at 37 °C. Dilution of the digest was done for analysis as described in Figures 1 and 2. For on-target washing, distilled water at room temperature was used. Generally 0.5 μL of distilled H_2O were deposited onto the sample spot and left there for ~30 s before the drop was blown off using pressurized air.

For in-capillary digestion, a nanoliter chemistry station is used. The sample handling procedures and sample deposition onto the matrix-covered probe tip have been described in more detail elsewhere [11]. Briefly, a 20-μm-diameter capillary tube connected to a syringe is used to draw a subnanoliter volume of a protein sample or reagents. Initially a 200–300 pL volume of 45 mM DTT in 0.04 M NH_4HCO_3 is drawn into the capillary and allowed to dry close to the capillary entrance. Then a 100–200 pL volume of the protein sample is loaded and the capillary closed with a Parafilm. After a reaction time of ~5 min the mixture is allowed to dry again close to the capillary entrance. A 200–300 pL volume of 100 mM iodoacetamide in 0.04 M NH_4HCO_3 is drawn into the capillary and the capillary is closed again. After ~10 min reaction time the mixture is allowed to dry and a 100–200 pL volume of freshly prepared 2 μM trypsin in 0.04 M NH_4HCO_3 is loaded. Sufficient digestion is usually achieved after 10–20 min. Before deposition of the final mixture onto the matrix-covered target a 100–200 pL volume of saturated 4-HCCA in 40% methanol/water (v/v) is loaded into the capillary. The sample mixture and matrix solution are hereby separated by a small air gap. Both solutions are then simultaneously deposited onto a matrix covered probe tip. Matrix layer formation is done according to the two-layer matrix/sample preparation method [4, 12]. To produce a very thin first layer, about 1 μL of a 5 mg/mL solution of 4-HCCA in 80% acetone/methanol (v/v) are deposited on the probe. After drying, a second layer of 0.4 μL of 4-HCCA saturated in 35% methanol/water (v/v) is deposited and allowed to dry. The sample/matrix mixture from the nanoliter sample preparation procedure is deposited on the probe as the third layer.

Results and Discussion

In routine MALDI-TOF analysis of tryptic digests of proteins using picomole amounts, interferences of matrix-cluster peaks from 4-HCCA can be eliminated by rigorous washing procedures in the two-layer matrix/sample preparation. This is demonstrated in Figure 1. An untreated digest, which was diluted in 150 mM saline solution, produces only matrix-cluster peaks; specific lactoferrin peptide fragments are not identified. After the on-probe washing procedure, many specific lactoferrin peptide fragments are detectable and interference from matrix clusters is almost eliminated. In this case the sample loss due to washing is largely compensated by the larger gain of the peptide signals. However, when dilute and/or heavily contaminated peptide mixtures are analyzed, matrix-cluster peaks can create a severe problem for accurate peptide identification. Figure 2 shows the MALDI spectra of the same tryptic digest, but with 1200-fold dilution using distilled water. For the unwashed preparation, matrix clusters interfere with peptide signals and the assignment of peaks to specific protein fragments becomes very difficult. As is shown in the other two spectra, washing can decrease this interference, but signals from specific protein fragments can also be lost. Other than interference from trypsin autolysis, the matrix-cluster peaks are not reproducible and their appearance and relative intensities in the spectra cannot be predicted. Note that in general the trypsin autolysis signals can be readily discerned

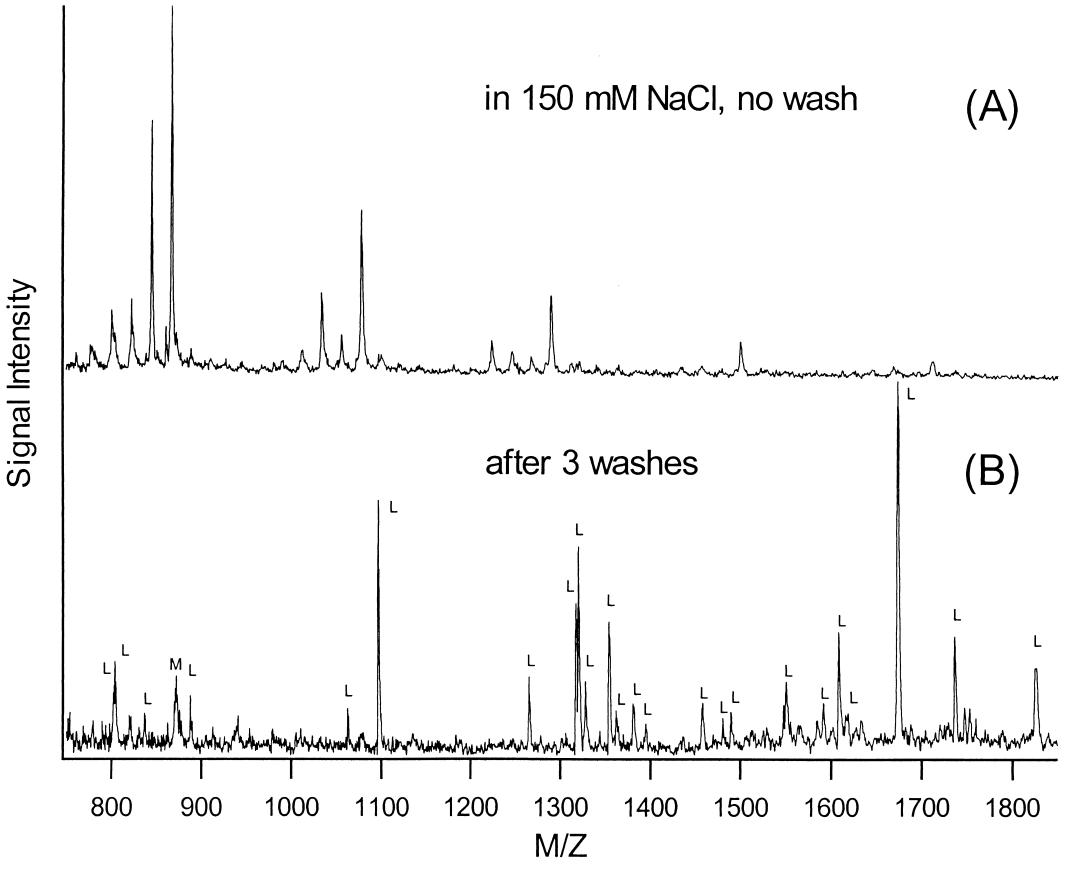

Figure 1. MALDI mass spectra of a tryptic digest of lactoferrin. An aqueous 15-μM lactoferrin solution was used for digestion and the resulting digest was then diluted by 100-fold in 150 mM NaCl (physiological saline). Assuming 100% digestion, the concentration of the diluted digest is ~150 nM. For each spectrum, 0.2 μL of this solution were directly deposited onto the 4-HCCA layer (i.e., the amount of each peptide loaded is about 30 fmol). (**A**) was not washed, whereas (**B**) was washed three times with distilled water as described in the text.

from the protein digest peptide peaks by doing a control experiment where all conditions are the same as in the actual protein digestion except no protein is added. In the cases of Figures 1 and 2, trypsin autolysis peaks are not detected, because the molar ratio of trypsin to protein is ~ 1:8 and thus autolysis peptides are not very abundant in the diluted digests. The commonly observed trypsin autolysis peaks include *m/z* 805.9, 907.1, 1112.3, 1154.3, 1434.7, 2164.3, and 2274.6.

We found that matrix-cluster formation can be influenced by changing the sodium and potassium salt content in the matrix solution before the matrix is deposited for crystallization on the probe tip. Also, basic conditions lead to enhanced matrix-cluster formation. Similar observations were reported for calcium adducts with dihydroxybenzyoic acid (DHB) as matrix [13]. Sodium and potassium salts are common impurities found in 4-HCCA and many protein samples, it is therefore not surprising that these ions play an important role in matrix-cluster formation. Different matrix preparations with deliberately altered concentrations of these salts lead to different cluster peaks, but there are some salient features. Figure 3 shows three mass spectra of typical matrix clusters from three different matrix preparations. Cluster peaks always appear in assemblies of several members, with one or two members in the center having the highest intensity. Assemblies of cluster peaks are usually separated by 190–228 mass units, which represent the molecular weight range of the matrix 4-HCCA plus a proton, sodium, or potassium ion. Generally the intensity of cluster peaks decreases with increasing mass. Interferences can appear in the mass region up to ~2000 Da, which is unfortunately the region of interest for peptide mass mapping.

There are slight differences in the spectral pattern of Figures 1A and 3C, although both preparations were done with physiological saline solution. These differences are most likely due to the different way of preparation. For Figure 1A the saline solution was deposited directly onto the second 4-HCCA layer, whereas for Figure 3C the saline solution was mixed in a basic environment with the second matrix solution, then acidified prior to deposition onto the first 4-HCCA layer. This further underlines the unpredictability of

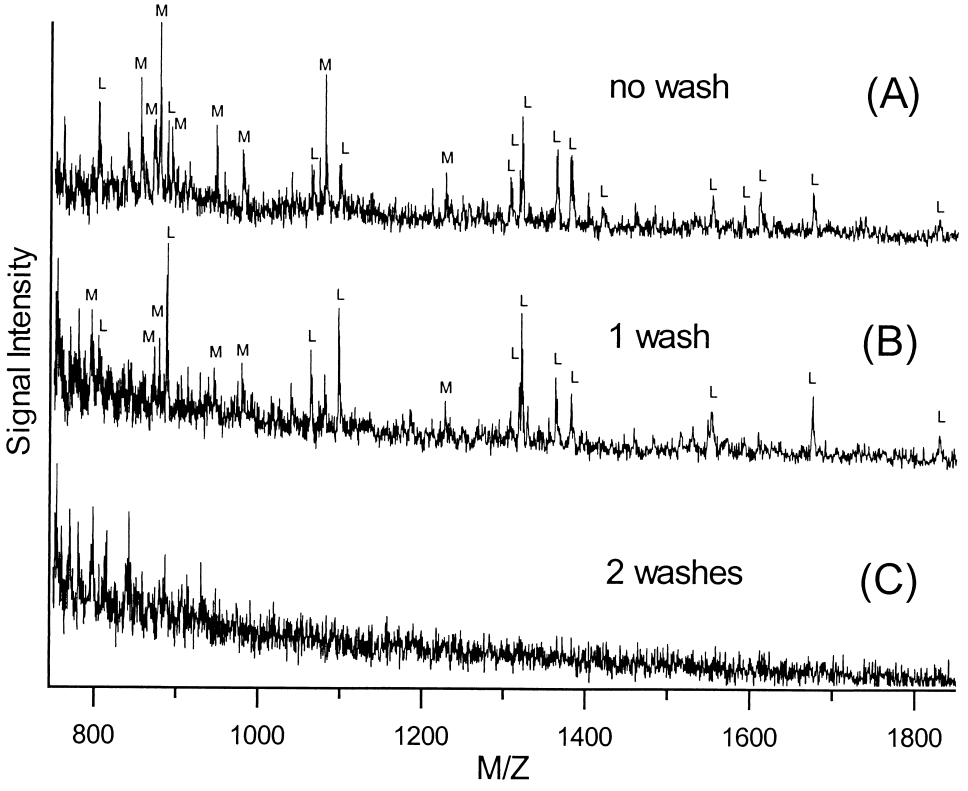

Figure 2. MALDI mass spectra of the same initial digest as in Figure 1, but with 1200-fold dilution in distilled water. For each spectrum, 0.2 μL of the diluted digest were directly deposited onto the 4-HCCA layer (i.e., the amount of each peptide loaded is about 2.5 fmol). Specific peaks from matrix clusters and lactoferrin peptide fragments are labeled M and L, respectively. (**A**) was not washed with distilled water, whereas (**B**) and (**C**) were washed once and twice, respectively.

matrix-cluster appearance. However, when looking at the observed masses for cluster peaks in different matrix preparations, we found that all theoretically possible cluster masses from combinations of intact 4-HCCA and potassium and/or sodium can be calculated with the following simple equations:

$$M_{\text{Cluster}} = n\text{M} - x\text{H} + y\text{K} + z\text{Na}$$

$$x = y + z - 1$$

$$y + z \leq n + 1 \qquad (\text{thus } x \leq n)$$

where M_{Cluster} is the observed mass of the ionized cluster, $n = 1, 2, 3, \ldots, y$ or $z = 0, 1, 2, 3, \ldots$, and x an integer meeting the above condition. M, K, and Na stand for the masses of 4-HCCA, potassium, and sodium, respectively.

Equipped with this knowledge, possible interferences from matrix-cluster peaks can now be easily detected and eliminated. This greatly facilitates the analysis of the digests of low amount of proteins in conventional sample handling. As in Figure 2B, the matrix peaks can be sorted out and these peaks will not be entered into the search program for database protein identification. Note that, if a peptide peak has the same mass as one of the clusters, some judgment based on peak intensity is needed to ascertain the origin of the peak.

Knowing the matrix-cluster peaks is also crucial for microspot MALDI analysis of very low abundant proteins. This is illustrated in Figure 4 where MALDI spectra of tryptic digests of 0.5 fmol lactoferrin is shown. Note that 0.5 fmol is the actual amount of the starting protein used for the digestion, not the result of dilution of tryptic digests resulting from the use of much higher amount of proteins required in conventional protein sample handling with microliter volumes. In this case, when a single shot spectrum is collected from inside the microspot and if it is unclear whether the peaks originate from peptides or matrix clusters, the spectrum is carefully checked for typical matrix-cluster characteristics. If several cluster assemblies separated by ~190–230 Da are detected, their intensities decrease with increasing mass, and no strong signal of peptides in the mass region between the cluster assemblies are detectable, then the single shot spectrum can be discarded.

Figure 4 also illustrates the importance of selecting the desorption sites on the MALDI probe. Sampling outside the microspot (see Figure 4A) only produces

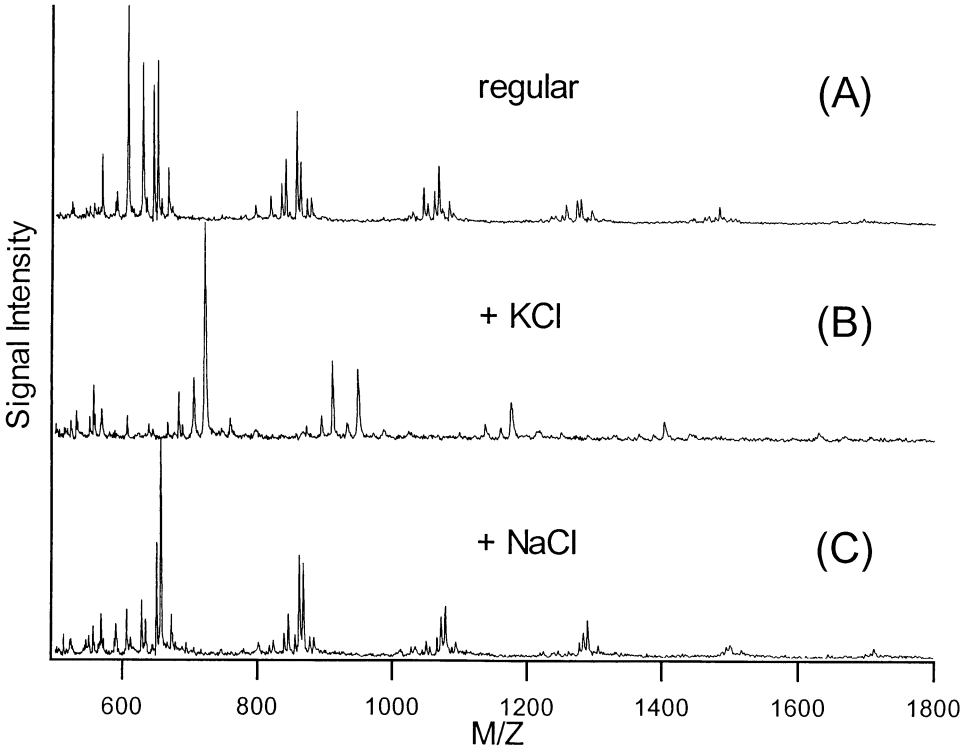

Figure 3. Mass spectra of 4-HCCA matrix clusters obtained under different matrix preparation conditions. (**A**) Untreated matrix from conventional two-layer preparation as described in the text. (**B**) 5 μL of second-layer 4-HCCA solution was first mixed with 5 μL 0.1 M NH_4HCO_3 then 5 μL of 1% KCl was added. To neutralize the solution, 2 μL of a slurry of precipitated 4-HCCA were finally added and the solution centrifuged. As in the conventional preparation, 0.4 μL of this mixture was deposited onto a first layer of 4-HCCA matrix. (**C**) Same preparation as in (**B**), except the 1% KCl solution was replaced with a 1% NaCl solution.

matrix-cluster peaks at high laser power since of course no peptides are present. In the border region of the microspot (see Figure 4B), peptide concentrations might be very low and thus only weak signals of peptides are obtained and only at medium laser power. In this case, significant interferences of matrix clusters in the MALDI spectrum of tryptic digest are observed. Only when sampling inside the microspot (see Figure 4C) strong peptide signals at low laser power are obtainable. In this case we discarded all spectra with any interfering cluster peaks (200 single shot spectra out of 600 shots were summed).

It is certainly desirable to reduce the extent of matrix-cluster formation whenever possible. Sample/matrix preparation should be carefully optimized. It is our experience that matrix purification using the simple ethanol recrystallization method can assist in reducing matrix-cluster formation. The improvement can be quite dramatic when a batch of 4-HCCA as received from the manufacturer contains a lot of salts and impurities. For the microspot experiment, we find that, by loading a small amount of matrix solution into the capillary after digestion and then depositing the digest *and* the matrix solution onto the matrix-covered probe tip, the spectral quality can be significantly improved, compared to direct deposition of the digest to the matrix layer. Specifically, it can help to neutralize the digestion buffer and thus avoid excessive matrix-cluster formation, and decrease the extent of salt adducts formation from tryptic peptides.

It should be pointed out that discerning matrix peaks is done manually at present. Nevertheless, it should be possible to perform this task using software control. Such an implementation could also become a valuable quality control tool for routine and especially for automated high-throughput techniques.

Finally we note that the above described method of discerning matrix-cluster peaks for 4-HCCA should be applicable to other monoacidic matrix substances, when their specific characteristics (e.g., H_2O loss or other group losses, different metal ions) are taken into account. For example, 2,5-dihydroxybenzoic acid (DHB) is another widely used matrix substance for peptide analysis. It forms similar clusters in the presence of salts, although not to the same extent as 4-HCCA (data not shown). We find that the observed cluster masses for DHB can be calculated according to

$$M_{\text{Cluster}} = n\text{M} + p[\text{M} - \text{H}_2\text{O}] - x\text{H} + y\text{K} + z\text{Na}$$

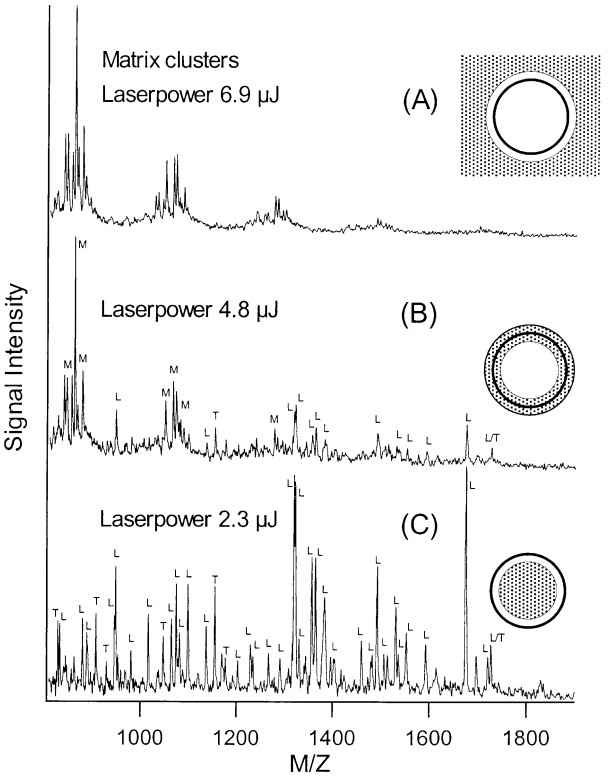

Figure 4. MALDI mass spectra of 390 pL of 0.1 $\mu g/\mu L$ lactoferrin (0.5 fmol) after incapillary reduction, carbamidomethylation and tryptic digestion. (**A**) The laser beam is sampling outside the microspot (shadow areas in the picture), (**B**) laser is sampling in the border region of the microspot, and (**C**) laser is sampling inside the microspot. Specific peaks from matrix clusters, trypsin autolysis, and lactoferrin peptide fragments are labeled M, T, and L, respectively.

$$x = y + z - 1$$

$$y + z \leq n + p + 1 \qquad (\text{thus } x \leq n + p)$$

where M_{Cluster} is the observed mass of the ionized cluster, n or $p = 0, 1, 2, 3, \ldots, n + p > 0$, y or $z = 0, 1, 2, 3, \ldots$, and x an integer meeting above condition. M, $[M - H_2O]$, K, and Na stand for the molecular weights of DHB, DHB $- H_2O$, potassium, and sodium, respectively. The difference between the above equation and the equation used for 4-HCCA is the addition of the $M - H_2O$ term for DHB. DHB loses H_2O much more readily than 4-HCCA. For 4-HCCA, cluster formations containing $[M - H_2O]$ are usually not observed in the mass region above ~ 500 Da.

Acknowledgments

This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Genomic Solutions, Inc. BOK thanks the University of Alberta for a Dissertation Fellowship.

References

1. See for example, (a) *Proteome Research: New Frontiers in Functional Genomics*; Wilkins, M. R.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F., Eds.; Springer: Berlin, 1997; (b) Quadroni, M.; James, P. *Electrophoresis* **1999,** *20*, 664–677; (c) Jungblut, P.; Thiede, B. *Mass Spectrom. Rev.* **1997,** *16*, 145–162.
2. Vorm, O.; Roepstorff, P.; Mann, M. *Anal. Chem.* **1994,** *66*, 3281–3287.
3. Zhang, H.; Caprioli, R. M. *J. Mass Spectrom.* **1996,** *31*, 1039–1046.
4. Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1996,** *68*, 2721–2725.
5. Li, L.; Golding, R. E.; Whittal, R. M. *J. Am. Chem. Soc.* **1996,** *118*, 11662–11663.
6. Allmaier, G. *Rapid Commun. Mass Spectrom.* **1997,** *11*, 1567–1569.
7. Kussmann, M.; Nordhoff, E.; Rahbek-Nielsen, H.; Haebel, S.; Rossel-Larson, M.; Jakobsen, L.; Gobom, J.; Mirgorodskays, E.; Kroll-Kristensen, A.; Palm, L.; Roepstorff, P. *J. Mass Spectrom.* **1997,** *32*, 593–601.
8. Murray, K. K. *Mass Spectrom. Rev.* **1997,** *16*, 283–299.
9. Beavis, R. C.; Chaudhary, T.; Chait, B. T. *Org. Mass Spectrom.* **1992,** *27*, 156–158.
10. Whittal, R. M.; Li, L. *Anal. Chem.* **1995,** *67*, 1950–1954.
11. Whittal, R. M.; Keller, B. O.; Li, L. *Anal. Chem.* **1998,** *70*, 5344–5347.
12. Dai, Y. Q.; Whittal, R. M.; Li, L. *Anal. Chem.* **1999,** *71*, 1087–1091.
13. Dubois, F.; Knochenmuss, R.; Steenvoorden, R. J. J. M.; Breuker, K.; Zenobi, R. *Eur. Mass Spectrom.* **1996,** *2*, 167–172.