PROTEOMICS METABOLOMICS GENOMICS INFORMATICS
GLYILEVALCYSGLUGLNALASERLEUASPARG
CYSVALLYSPROLYSPHETYRTHRLEUHISLYS

Accurate-Mass LC/TOF-MS for Molecular Weight Confirmation of Intact Proteins

Author

Ning Tang
Senior Applications Scientist
Agilent Technologies, Inc.
Santa Clara, California, USA

Abstract

The use of liquid chromatography/mass spectrometry (LC/MS) methods for protein analysis is well documented and has had a significant impact on understanding the complexities of biological systems. Less well publicized is the role that mass spectrometry has had in the characterization of expressed proteins in molecular biology laboratories. LC/MS is one of the most effective technologies for protein identification in both simple and complex samples. This application note will demonstrate the use of liquid chromatography and high-mass-accuracy time-of-flight mass spectrometry (LC/TOF-MS) for the characterization of expressed proteins. Accurate-mass determinations can be completed in less than an hour and provide the molecular biologist with vital information about the success of protein expression via recombinant DNA methods.

Introduction

Since the introduction of recombinant DNA (r-DNA) techniques in 1973 by Stanley Cohen and Herbert Boyer,[1] they have become some of the most common molecular biology methodologies. r-DNA techniques have been used to produce protein-based pharmaceuticals such as insulin, erythropoietin, and human growth hormone; and produce genetically modified proteins to study biological processes. One of the many challenges in implementing r-DNA techniques is rapid characterization of the expressed proteins. Since experimental parameters such as growth conditions, construct design, and protein purification methods can affect the final protein product, an efficient method is needed to ensure the final protein product is correct.

One method that has been used is one-dimensional gel electrophoresis. However, one-dimensional gels do not have the mass accuracy or specificity necessary to detect subtle protein changes, such as a point mutation due to a PCR error or proteolytic clipping at the amino or carboxyl terminus. It is desirable to have a method with which to quickly, and with high accuracy, assess the integrity of the expressed protein.

Time-of-flight mass spectrometers, with resolving power greater than 10,000 and mass accuracy better than 2 ppm for monoisotopic masses or 10–20 ppm for the average mass of proteins, provide researchers with a rapid, sensitive, and accurate technique to determine the molecular weight of proteins. Combining liquid chromatography with TOF-MS greatly reduces the need for time-consuming sample preparation, and can provide molecular weight confirmation in less than one hour. This note describes several examples of using LC/TOF-MS to assess the quality of expressed proteins.

Experimental

Chromatographic separation of the proteins was performed by an Agilent 1100 LC configured with a capillary pump, micro well plate sampler with cooler, and thermostatted column compartment. A C18 capillary column (Vydac 218TP5105) was used. Depending on the protein concentration in the sample, between 0.1 and 8 µL of sample was injected for a protein load of approximately 0.5–1 µg. Column flow was 20 µL/min and column temperature was 30 °C. Solvent A was 0.05% trifluoroacetic acid (TFA) in water; solvent B was 0.05% TFA in acetonitrile.

The 0.05% TFA was added to the LC solvents to improve the chromatographic separation and, especially, to increase the solubility of eluted proteins in acetonitrile.[2] Recovery of larger proteins is often very poor if the TFA is omitted. However, TFA can severely suppress ionization of proteins, so 20 µL/min neat acetic acid was introduced into the LC eluent using a tee placed between the HPLC UV detector and the mass spectrometer ion source. The resulting 50% acetic acid displaced the TFA from the protein.[3]

Mass spectrometry was performed using an Agilent LC/MSD TOF (G1969A) equipped with a dual-nebulizer electrospray (ESI) ion source. The second nebulizer and the LC/MSD TOF's built-in calibrant delivery system were used to continuously introduce a reference compound at very low levels relative to the samples. Reference mass correction was enabled with ions at 173.9616, 186.8781, 245.8914, 304.9047, 360.8397, 419.8530, 478.8663, 533.8852, 537.8796, 922.0098.[4]

Complete system control of both the LC and MS was accomplished using Agilent LC/MSD TOF software (version A.02.01). Deconvolution of the protein spectra was performed using Agilent MassHunter protein confirmation software (version A.02.00).

LC Conditions	
Column:	Vydac capillary C18, #218TP5105
Flow rate:	20 µL/min
Column temperature:	30 °C
Injection volume:	0.1– 8.0 µL (achieve protein load of 0.5–1.0 µg protein)
Mobile phase:	A: 0.05% trifluoroacetic acid (TFA) in water B: 0.05% TFA in acetonitrile
Gradient:	00% B at 00.0 min 70% B at 35.0 min 90% B at 40.0 min 00% B at 40.1 min

MS Conditions	
Ionization mode:	Positive ESI
Drying gas flow:	10 L/min
Nebulizer:	10 psig
Ion source temperature:	350 °C
Vcap:	5000 V
Fragmentor:	235 V
Skimmer:	60 V
Octopole RF:	250 V
Scan range:	*m/z* 145–3200

Results and Discussion

Two sets of intact protein samples were analyzed to demonstrate the utility of LC/TOF-MS for characterization of expressed proteins. Another sample was analyzed to demonstrate the ability of LC/TOF-MS to determine labeling efficiency for isotopically labeled proteins.

Analysis of glutamine synthetase

Glutamine synthetase lies at the crossroads of nitrogen metabolic pathways in almost all organisms, from bacteria to mammals. In many Gram-negative bacteria, as well as in Mycobacteria, the activity of glutamine synthetase is regulated by multiple mechanisms including reversible modification of a tyrosine residue by adenylylation.[5] When over-expressing in *E. coli*, the protein must be checked both for possible sequence errors in the plasmid construct and for the extent of adenylylation.

The expected mass of unmodified glutamine synthetase is 51,772.7 u. Adenylylation increases this by 329 u. Figure 1 shows the electrospray mass spectrum (top) and the deconvoluted spectrum (bottom) of glutamine synthetase expressed under growth conditions expected to give the unmodified form. Each ion peak in the top spectrum represents a different charge state. For example, the ion at *m/z* 1205 carries a charge of +43. The series of multiply charged ions generated by ESI is deconvoluted via an algorithm to determine the molecular weight of the protein. When ESI is combined with a high-resolution and high-mass-accuracy TOF-MS, the mass accuracy observed is generally better than 20 ppm or 0.002%. The measured mass was 51,772.9, which is 0.2 mass units higher than calculated, giving an error of 6 ppm. The small molecular weight difference between the calculated and observed mass provides high confidence that the desired sequence was actually expressed and that it was not post-translationally modified. The complete LC/TOF-MS analysis took 45 minutes.

Analysis of an NPr construct

No sequence errors were introduced during the construction of the glutamine synthetase vector, but such errors are not rare during manipulation by PCR. An example of an undesired sequence change came from the analysis of NPr, a member of a phosphotransfer cascade of proteins implicated in potassium transport in bacteria.[6] The measured mass was 18 u too low (Table 1). The protein was then digested with LysC, and the resulting peptides were analyzed by LC/TOF-MS. All but one of the peptides had the expected molecular weight.

The peptide containing residues 12–21 had a mass that was 17.95 u too low. The only DNA mutation that could produce this result was a change of a Met to Ile or Leu.

Table 1. Calculated and observed molecular weights, and mass differences, of the NPr constructs.

Protein	Calculated MW (u)	Observed MW (u)	Difference (u)
NPr Construct	9115.4	9097.00	18.40
NPr Construct, peptide P12-21	1110.58	1092.63	17.95
NPr New Construct	9115.4	9115.00	0.40

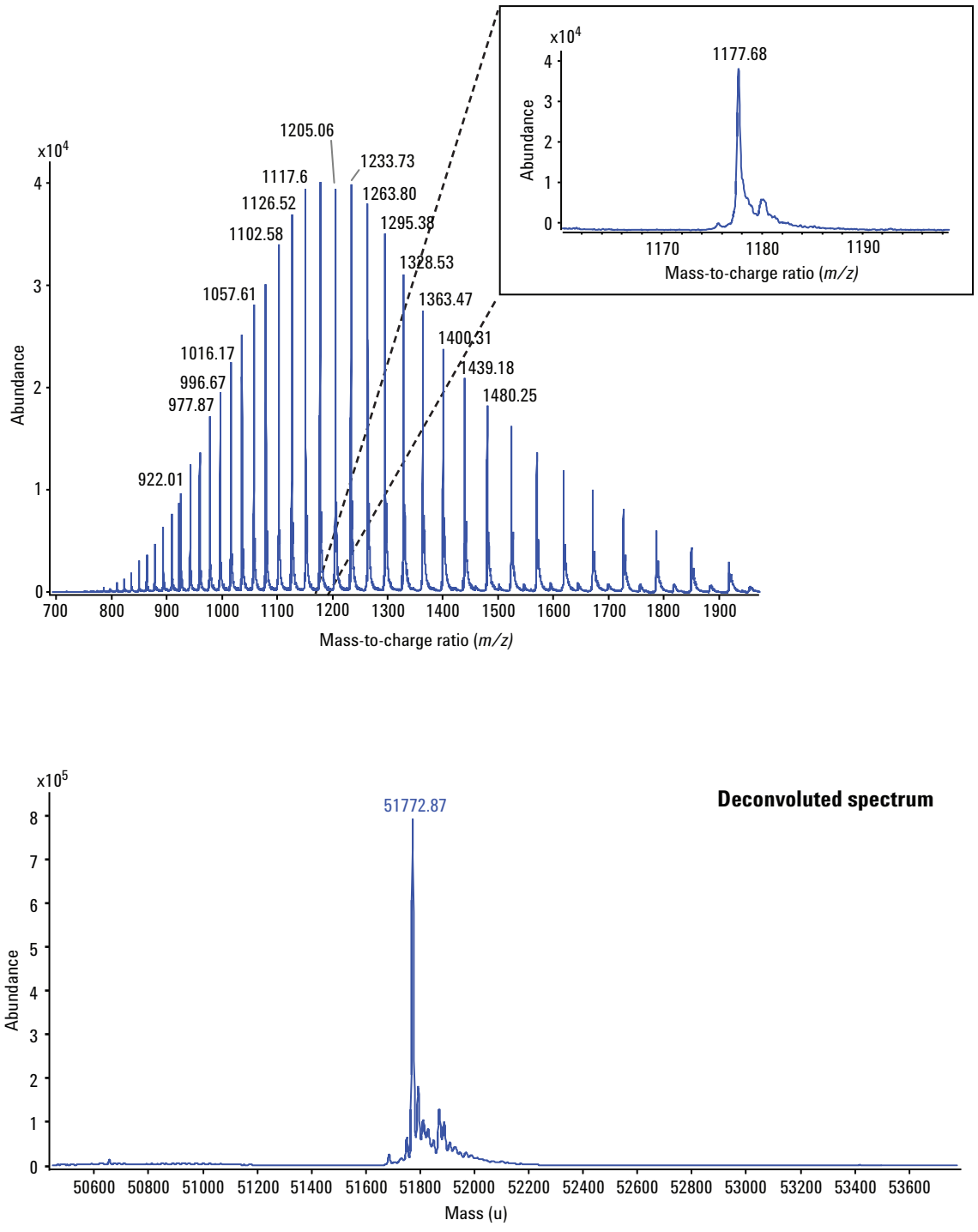

Figure 1. Mass spectrum (top) and deconvoluted mass spectrum (bottom) of expressed glutamine synthetase.

Further LC-MS analyses showed that Met 15 was the mutation site. The original DNA construct was then sequenced, demonstrating that the ATG codon for residue 15 had been mutated to TTG, which codes for Leu (data not shown here). Thus, the LC/TOF-MS analysis quickly detected a presumptive PCR error in the construct.

A new construct was made and the expressed protein had the correct molecular weight as shown in Table 1 and Figure 2.

Determining isotope labeling efficiency

Isotopically labeled proteins are often used in NMR studies. Determining the extent of isotopic labeling prior to NMR analysis is another task that can be accomplished by LC/TOF-MS analysis. For this experiment, the NPr protein was expressed in a growth medium containing ^{15}N labeled ammonia and ^{13}C labeled glucose. The efficiency of labeling was determined by comparing the observed average molecular weight of the $^{13}C/^{15}N$ labeled NPr (9610.8 u) with the calculated average molecular weights of fully $^{13}C/^{15}N$ labeled NPr (9615.5 u) and unlabeled NPr (9115.4 u) as shown in Table 2. The mass difference between the observed labeled NPr and calculated unlabeled NPr is 9610.8 u – 9115.4 u = 495.4 u. If the NPr was fully labeled, the calculated mass difference between labeled and unlabeled should be 9615.5 u – 9115.4 u = 500.1 u. Thus, the efficiency of labeling was 495.4 u/500.1 u = 99.1%.

Table 2. ^{13}C and ^{15}N labeling efficiency of NPr construct.

Observed average MW – labeled (u)	Calculated average MW – labeled (u)	Calculated average MW – unlabeled (u)	Difference (u)
9610.8		9115.4	495.4
	9615.5	9115.4	500.1
			495.4/500.1 = 99.1%

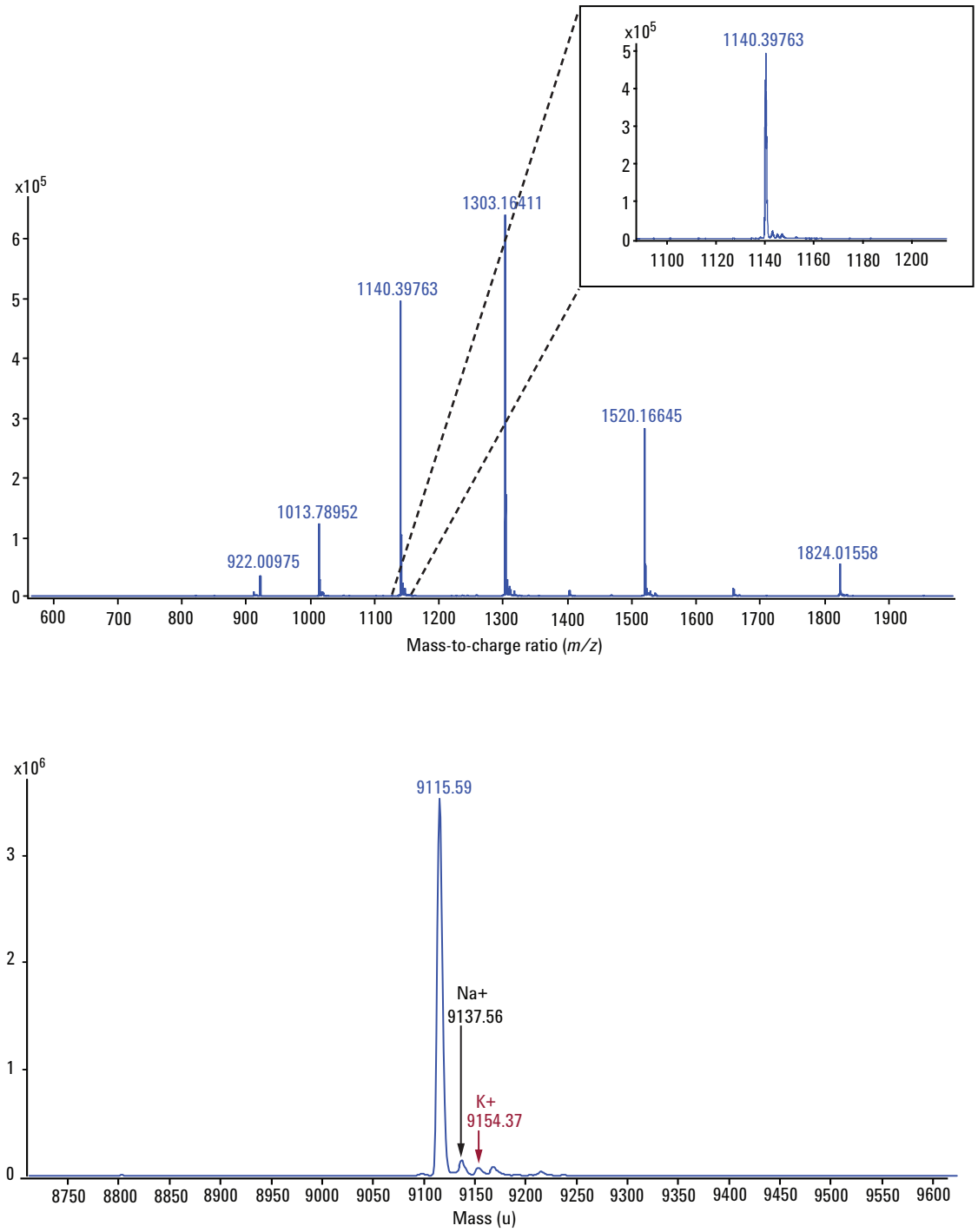

Figure 2. Mass spectrum and deconvoluted mass spectrum of new construct NPr.

Conclusion

This note has demonstrated the capability of LC/TOF-MS to quickly assess the accuracy of an expressed protein by determining its molecular weight with high mass accuracy. The procedure requires minimal sample preparation and can be automated to run unattended. It is sensitive and usually requires $\leq$1 µg of protein. It has also been shown that this approach can aid in the troubleshooting of expressed proteins.

References

1. Cohen, S.N., Chang, A. C., Boyer, H. W., Helling, R. B. Construction of Biologically Functional Bacterial Plasmids in vitro. *Proc. Nat. Acad. Sci.* 70: 3240–3244; 1973.
2. Mahoney, W. C., Hermodson, M. A. Separation of Large Denatured Peptides by Reverse Phase High Performance Liquid Chromatography; Trifluoroacetic Acid as a Peptide Solvent. *J. Biol. Chem.* 255: 11199–11203; 1980.
3. Apffel, A., Fischer, S., Goldberg, G., Goodley, P. C., Kuhlmann, F. E. Enhanced Sensitivity for Peptide Mapping with Electrospray Liquid Chromatography-Mass Spectrometry in the Presence of Signal Suppression due to Trifluoroacetic Acid-Containing Mobile Phases. *J. Chromatogr. A.* 712: 177–190; 1995.
4. Levine, Rodney L. Fixation of nitrogen in an electrospray mass spectrometer. *Rapid Commun. Mass Spectrom.* 20: 1828–1830; 2006.
5. Stadtman, E. R. Discovery of Glutamine Synthetase Cascade. *Methods Enzymol.* 182: 793–809; 1990.
6. Lee, C. R., Cho, S. H., Yoon, M. J., Peterkofsky, A., Seok, Y. J. *Escherichia coli* Enzyme II A^{Ntr} Regulates the K^+ Transporter TrkA. *Proc. Natl. Acad. Sci* U.S.A. 104: 4124–4129; 2007.

About Agilent Technologies

Agilent Technologies is a leading supplier of life science research systems that enable scientists to understand complex biological processes, determine disease mechanisms, and speed drug discovery. Engineered for sensitivity, reproducibility, and workflow productivity, Agilent's life science solutions include instrumentation, microfluidics, software, microarrays, consumables, and services for genomics, proteomics, and metabolomics applications.

Buy online:
www.agilent.com/chem/store

Find an Agilent customer center in your country:
www.agilent.com/chem/contactus

U.S. and Canada
1-800-227-9770
agilent_inquiries@agilent.com

Europe
info_agilent@agilent.com

Asia Pacific
adinquiry_aplsca@agilent.com

Printed in the U.S.A. October 9, 2007

5989-7406EN

Agilent Technologies