Vol. 173, No. 3, 1990
December 31, 1990

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 927-931

ANALYSIS OF MIXTURES OF CLOSELY RELATED FORMS OF BOVINE TRYPSIN BY ELECTROSPRAY IONIZATION MASS SPECTROMETRY: USE OF CHARGE STATE DISTRIBUTIONS TO RESOLVE IONS OF THE DIFFERENT FORMS

S.K. Chowdhury and B.T. Chait

The Rockefeller University, New York, N.Y. 10021

Received October 12, 1990

SUMMARY: Electrospray ionization mass spectrometry has been utilized for the analysis of bovine trypsin. Commercial active bovine trypsin is comprised of a mixture of three closely related different forms: β-, α-, and ψ-trypsin with similar molecular masses (23293, 23311, and 23329 mass units respectively). Peak broadening caused by the natural isotopic abundance of atoms in the protein molecules results in overlap of the ion peaks from the different trypsin forms. Therefore, accurate determination of the molecular masses of these trypsins in a mixture is not straightforward. However, because the trypsin isoforms have different number of basic groups and different higher order structures, the different forms acquire different numbers of charges. Therefore, ions from the different trypsins appear in different parts of the mass spectrum allowing accurate determination of their molecular masses.

The recent development of two new desorption ionization techniques, namely, electrospray ionization, ESI (1,2), and matrix-assisted laser desorption (3), have revolutionized the mass spectrometry of proteins. Utilizing the two new techniques, molecular masses of proteins up to ~40,000 u can now be determined with an accuracy of less than 200 ppm and up to ~100,000 u with lesser accuracy (2-7). In electrospray ionization, multiply charged gas phase ions of large biopolymers are produced as a result of the attachment of protons and/or cations (e.g. Na^+) to the available basic sites on the molecule. Because of the attachment of a large number of charges to protein molecules, the mass-to-charge ratio (m/z) of protein ions become small and typically range between 500-2500. Therefore, conventional mass analyzers such as quadrupoles with limited m/z range can be utilized to analyze large proteins.

Recently, we have reported the investigation of a number of proteins using a simple, inexpensive electrospray ionization quadrupole mass spectrometer that was constructed at our laboratory (6). In the present Communication, we describe the accurate molecular mass (MM) determination of three different forms of a commonly

0006-291X/90 $1.50

used enzyme mixture, bovine trypsin, by the ESI technique. Commercial active bovine trypsin is comprised of a mixture of three forms: β-trypsin, α-trypsin and ψ-trypsin (8,9). The difference between the molecular masses of β- and α-trypsins is 18 u and the same is true for α- and ψ-trypsins. Because of the natural isotopic abundance of atoms in protein molecules, the high mass tail of ions from one form overlaps with the low mass end of the distribution of ions from the other. Therefore, it is difficult to accurately determine the molecular masses of the three closely related trypsins, particularly when one component is present in small amount relative to the others. Even when various trypsin forms are present in nearly equal amounts, a resolution of approximately 1300 would be required to resolve ions from the different trypsins; a resolution higher than is frequently used with quadrupole mass analyzers coupled to ESI sources. Furthermore, because of the presence of small amounts of sodium and other low MM adducts that are often observed in the electrospray ionization mass spectra of proteins (5,6), the ions of a second closely related material become more difficult to resolve and result in a decrease in the accuracy in MM measurements from the unresolved ion peaks. We describe here a method whereby the closely related forms of trypsin in a mixture can be resolved in the mass spectrum according to the different number of charges acquired by the different forms, allowing the accurate determination of their molecular masses.

MATERIALS AND METHODS. Bovine trypsin (catalog # T8642) and bovine trypsinogen (catalog # T1143) were obtained from Sigma Chemical Co., St. Louis, MO and were used without any further purification. High purity water and methanol were obtained from Burdick and Jackson, Muskegon, MI. Ultrapure acetic acid was obtained from J.T. Baker Inc., Phillipsburg, NJ. The protein was dissolved in a solution containing water, methanol and acetic acid with a ratio 47.5:47.5:5 % v/v, to obtain a concentration of 20 μM.

The electrospray ionization mass spectrometer and the procedure adopted to acquire mass spectra have been described previously (6). Briefly, the sample solution was electrosprayed from a syringe needle in ambient air at a rate of 0.6 μl/min. The resulting highly charged droplets and solvated ions were transported into the vacuum of the quadrupole mass spectrometer through a 20 cm long, 0.5 mm i.d. heated metal capillary tube for mass-to-charge ratio (m/z) analysis. The spectra were acquired using a commercial data system, vector-1 (Teknivent, St. Louis, MO), on an IBM AT compatible computer. Molecular masses of proteins were determined from the observed m/z values using the procedure described in the literature (1,2,5).

RESULTS AND DISCUSSION. The electrospray ionization mass spectrum of bovine trypsin is shown in Figure 1. Two intensity distributions of charge states can be discerned. One distribution centers around the 18+ ion and the other around the 14+ ion. 18+ and 14+ represent ions with 18 and 14 protons respectively attached to the

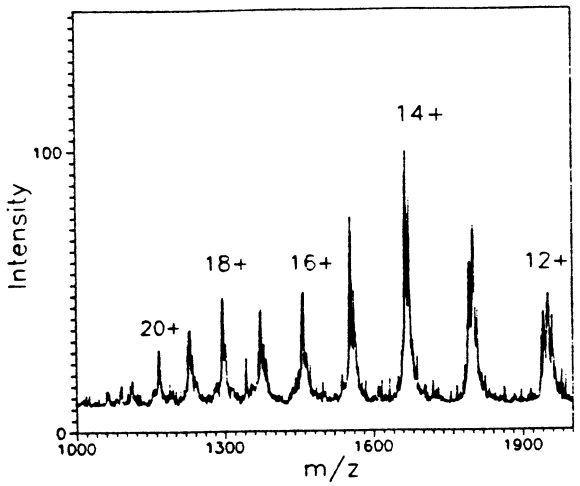

Figure 1. Electrospray ionization mass spectrum of bovine trypsin. The spectrum is an average of 5 scans, each acquired in 100 s from a 20 μM trypsin solution of the enzymes with solvent composition: water/methanol/acetic acid = 47.5/47.5/5 % v/v. The flow rate was 0.6 μl/min. The peaks labelled n+ arise from ions with n protons attached to the trypsin molecule.

trypsin molecule. The calculation of molecular mass (MM) from the observed m/z values of ions designated as 20+ to 12+ indicates the presence of three different proteins (Table 1). The MM obtained from the 12+ to 14+ ions is 23295.7 ± 1.8, that from the 17+ to 20+ ions is 23324.9 ± 2.3, and that from the 16+ ion is 23309.6. Ions with intermediate charge states (15+, 16+, and 17+) may have unresolved contributions from the proteins producing the two flanking ion distributions. The multiple resolved ion components observed with the 12+ to 15+ ions arise from the attachment(s) of sulfuric acid and/or phosphoric acid molecule(s) to the protein ions (10). These acids are

Table 1. Observed and calculated molecular masses of trypsinogen and different trypsin forms

Protein	Observed MM(u)	Charge States	Calculated[a] MM(u)	Δ[b] (ppm)
bovine trypsinogen	23981.6 ± 2.0	13 - 16	23981.1	+ 20
bovine trypsin	23295.7 ± 1.8	12 - 14	23293.3(β)	+ 103
	23309.6[c]	16	23311.3(α)	- 73
	23324.9 ± 2.3	17 - 20	23329.4(ψ)	+ 193

a: Molecular masses are calculated using the sequences compiled in the Dayhoff Protein Sequence Database (11).

b: Difference between the observed (column 2) and the calculated molecular mass (column 3).

c: Data from a single determination.

formed in the acidic spray solution from the sulfate/phosphate impurities present in the protein sample. Many other proteins are also observed to produce such adduct ions in ESI mass spectra (6).

Commercially available active bovine trypsin normally contains a mixture of three different forms of the enzyme, namely β-, α-, and ψ-trypsin (8,9). β-trypsin is a single chain enzyme, resulting from the cleavage of the Lys(6) - Ile(7) bond of trypsinogen (obs. MM=23981.6 u, cal. MM=23981.1 u; Table 1) with the release of the hexapeptide, VDDDDK. The calculated MM of bovine β-trypsin is 23293.3 u. α-trypsin originates from the additional hydrolysis of Lys(145) - Ser(146) and, therefore, has a MM 18 u higher than the β-form. ψ-trypsin refers to a form with an additional split at Lys(188) - Asp(189) and its MM is 18 u higher than the α-enzyme.

The observed MM of 23295.7 u closely agrees to that calculated for β-trypsin, the value 23324.9 corresponds to ψ-trypsin, and the intermediate value, 23309.6 u appears to be in agreement with the calculated MM of α-trypsin. No correlation of the amount of each enzyme form present in the sample with the observed relative ion abundances can be made from the electrospray ionization mass spectrum (Fig. 1) because the relative mass spectrometric responses of these different trypsin forms are not known. It is expected that ions arising from α-trypsin will have higher charge states than those arising from β-trypsin because it has an additional free NH_2-group (from the internal cleavage) and because its less constrained structure allows it to acquire and hold on to a higher number of charges (12). Similarly, it is expected that ions arising from ψ-trypsin will acquire a higher number of charges than both α and β-trypsin. The observed electrospray ionization charge state distributions of the bovine trypsin isoforms (Figure 1) confirm these expectations. The results described above are of practical importance because the different trypsin forms produce ions with different charge state distributions and therefore, yield ions in different regions of the mass spectrum. Thus, an accurate determination of the molecular masses of the three closely related different trypsin forms that differ in MM by only 18 u is obtained. We predict that this method for analyzing closely related forms of proteins will be generally applicable to other cases where the different forms adopt different higher order structures. Such measurements would be quite difficult to perform with other techniques in which the different forms do

not produce ions with the different charge state distributions and/or the mass spectrometric resolution is not sufficient to resolve ions from the different forms.

ACKNOWLEDGMENT. This work was supported in part by grants RR00862 and RR07065, Division of Research Resources, NIH, and GM38274, National Institute of General Medical Sciences, NIH.

REFERENCES

1. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989) Science 246, 64-71.
2. Smith, R.D., Loo, J.A., Edmonds, C.G., Barinaga, C.J., and Udseth, H.R. (1990) Anal. Chem. 62, 882-899.
3. Karas, M., Bahr, U., and Hillenkamp, F. (1989) Int. J. Mass Spectrom. Ion Processes 92, 231-242.
4. Beavis, R.C. and Chait, B.T. (1990) Anal. Chem. 62, 1836-1840.
5. Covey, T.R., Bonner, R.F., Shushan, B.I., and Henion, J.D. (1988) Rapid Commun. Mass Spectrom. 2, 249-256.
6. Chowdhury, S.K, Katta, V., and Chait, B.T. (1990) Rapid Commun. Mass Spectrom. 4, 81-87.
7. Henry, K.D., Williams, E.R., Wang, B.H., McLafferty, F.W., Shabanowitz, J., and Hunt, D.F. (1989) Proc. Natl. Acad. Sci. USA 86, 9075-9078.
8. Schroeder, D.D. and Shaw, E. (1968) J. Biol. Chem. 243, 2943-2949.
9. Fehlhmmer, H. and Bode, W. (1975) J. Mol. Biol. 98, 683-692.
10. Chowdhury, S.K., Katta, V., Beavis, R.C., and Chait, B.T. (1990) J. Am. Soc. Mass Spectrom. 1, 382-388.
11. Protein Identification Resource (PIR), National Biomedical Research Foundation (NBRF) Release 19.0 (3187 December 1988).
12. Chowdhury, S.K., Katta, V., and Chait, B.T. (1990) J. Am. Chem. Soc. (in press).