Electrospray-ionization time-of-flight mass spectrometry in protein chemistry

ARTICLE *in* ANALYTICAL CHEMISTRY · JANUARY 1994

Impact Factor: 5.64 · DOI: 10.1021/ac00073a018

5 AUTHORS, INCLUDING:

Igor V Chernushevich
AB SCIEX
45 PUBLICATIONS **2,792** CITATIONS

Anal. Chem. 1994, *66*, 99–107

Electrospray Ionization Time-of-Flight Mass Spectrometry in Protein Chemistry

Olga A. Mirgorodskaya* and Andrew A. Shevchenko
Institute for Cytology of the Russian Academy of Sciences, 4 Tikhoretsky Avenue, 194064 St. Petersburg, Russia

Igor V. Chernushevich and Alexander F. Dodonov†
Institute of Energy Problems of Chemical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

Anatoliy I. Miroshnikov
Shemyakin Institute for Bioorganic Chemistry of the Russian Academy of Sciences, Miklukho-Maklay Street 16/10, 117871 Moscow, Russia

A novel electrospray ionization time-of-flight mass spectrometer with an ion mirror has been developed. The "orthogonal acceleration" of ions combined with ion storage modulation enables large sensitivity to be obtained with moderate ($R >$ 2000) mass resolution. Application of the transient recording system allows simultaneous registration of ions of all masses, with sensitivity equal to that in the single ion monitoring regime. The potential of this technique is demonstrated by the primary structure determination of cationic peptides from porcine neutrophils and by the study of the pathways of recombinant proinsulin proteolytic processing catalyzed by various trypsin forms.

Several soft ionization mass spectrometry methods such as matrix-assisted laser desorption, plasma desorption, secondary ion, fast atom bombardment, electrospray (ES), and thermospray made a revolution in the integration of mass spectrometry with biological research during the last decade.[1-3] The ability to generate gaseous ions from large, nonvolatile, and labile molecules makes these methods the best choice for analyzing biopolymers. Electrospray seems to be one of the most popular ionization techniques today[1-8] due to its sensitivity and ability to analyze mixtures. It is used mostly with quadrupole instruments[1-3] and more rarely with magnetic sector mass spectrometers.[4-8]

It seemed attractive to design a time-of-flight (TOF) mass spectrometer with an electrospray ion source and thus enjoy the well-known advantages of TOF MS: mass range limited by ion detector only, high transmission, relatively low cost, and ability for simultaneous registration of ions of all masses,[9,10] with sensitivity equal to that in the single ion monitoring regime. Unfortunately, electrospray is an inherently continuous ionization method and cannot be directly coupled with a TOF mass analyzer without a tremendous loss of sensitivity. From 1985 to 1987 we designed and created a novel reflecting TOF mass spectrometer capable of mass analyzing any continuous ion beam without a loss of sensitivity.[11,12] It was based on "orthogonal acceleration" of ions into a TOF mass spectrometer using an ion reflector. Several ionization sources that work at atmospheric pressure have been coupled to this instrument: corona discharge, tritium radioactive foil, UV lamp, and electrospray.[13-15]

It is necessary to mention three other research groups where TOF mass spectrometers with orthogonal acceleration of ions from different ionization sources have been developed independently.[16-20] The construction of Guilhaus and

* Author to whom correspondence concerning protein chemistry applications and reprint requests should be addressed.
† Author to whom correspondence concerning the device construction and its operational principles should be addressed.

(1) Carr, S. A.; Hemling, M. F.; Roberts, G. D. *Anal. Chem.* **1991**, *63*, 2802–2824.
(2) Smith, R. D.; Loo, J. A.; Edmonds, C. G.; Barinaga, C. J.; Udseth, H. R. *Anal. Chem.* **1990**, *62*, 882–899.
(3) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Mass Spectrom. Rev.* **1990**, *9*, 37–70.
(4) Cody, R. B.; Tamura, J.; Musselman, B. D. *Anal. Chem.* **1992**, *64*, 1561–1570.
(5) Alexandrov, M. L.; Gall, L. N.; Krasnov, N. V.; Nikolayev, V. I.; Pavlenko, B. I.; Shkurov, V. A. *Dokl. Akad. Nauk SSSR* **1984**, *277*, 379–383 (in Russian).
(6) Alexandrov, M. L.; Baram, G. I.; Gall, L. N.; Grachev, M. A.; Knorre, V. D.; Krasnov, N. V.; Kusner, Yu. S.; Mirgorodskaya, O. A.; Nikolaev, V. I.; Shkurov, V. A. *Bioorg. Khim.* **1985**, *11*, 705–707 (in Russian).
(7) Aleksandrov, M. L.; Kondrat'ev, V. M.; Kusner, Yu. S.; Mirgorodskaya, O. A.; Podtelezhnikov, A. V.; Fridlyanskii, G. V. *Sov. J. Bioorg. Chem.* **1988**, *No. 4*, 473–477.
(8) Molchanova, T. P.; Mirgorodskaya, O. A.; Abaturov, L. V.; Podtelezhnikov, A. V.; Tokarev, Yu. N.; Grachev, S. A. *Mol. Biol.* **1989**, *23*, 169–180 (in Russian).
(9) Emary, W. B.; Lys, I.; Cotter, R. J.; Simpson, R.; Hoffman, A. *Anal. Chem.* **1990**, *62*, 1319–1324.
(10) Holland, J. F.; Newcomb, B.; Tecklenburg, R. E., Jr.; Davenport, M.; Allison, J.; Watson, J. T.; Enke, C. G. *Rev. Sci. Instrum.* **1991**, *62*, 69–76.
(11) Dodonov, A. F.; Chernushevich, I. V.; Dodonova, T. F.; Raznikov, V. V.; Tal'rose, V. L. Inventor's Certificate No. 168134OA1, USSR, February 25, 1987.
(12) Dodonov, A. F.; Chernushevich, I. V.; Dodonova, T. F.; Raznikov, V. V.; Tal'rose, V. L. International Patent Appl., PCT/SU89/00228}, April 1989.
(13) Tal'rose, V. L.; Dodonov, A. F.; Chernushevich, I. V., 41st Pittsburg Conference on Analytical Chemistry and Applied Spectrometry, Chicago, IL, March 1991; Paper 299P.
(14) Dodonov, A. F.; Chernushevich, I. V.; Laiko, V. V. *Book of Abstracts*, 12th International Mass Spectrometry Conference, Amsterdam, August 1991; Abstract 153.
(15) Chernushevich, I. V. Ph.D. Thesis, Institute of Energy Problems of Chemical Physics of Russian Academy of Sciences, Chernogolovka, 1991 (in Russian).
(16) Dawson, J. H. J.; Guilhaus, M. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 155–159.
(17) Dawson, J. H. J.; Guilhaus, M. Australian Provisional Patent, P16079 Unisearch Limited, December 1987.
(18) Coles, J.; Stutsel, M.; Guilhaus, M. *Abstracts of Papers*, 40th ASMS Conference on Mass Spectrometry and Allied Topics, 1992; pp 10–11.

0003-2700/94/0366-0099$04.50/0

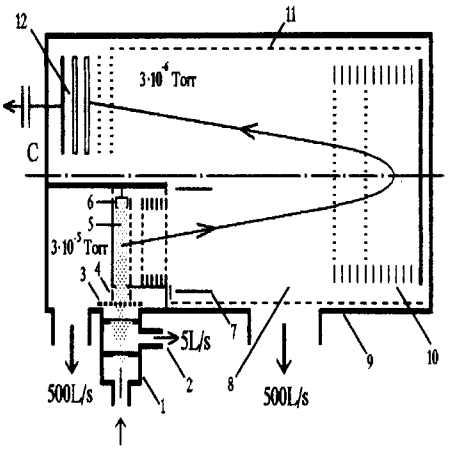


Figure 1. Schematic diagram of ES TOF mass spectrometer: (1) electrospray source, (2) first stage of differential pumping, (3) focusing grid, (4) quadrupole lens, (5) time-of-flight ion storage modulator, (6) ion collector, (7) deflection plates, (8) ion drift region, (9) vacuum chamber, (10) two-stage ion mirror, (11) guard grid at high potential, and (12) ion detector.

Dawson[16–18] resembles our TOF device, although differing in at least two important features: there is no ion mirror, and accumulation of ions is not utilized to maximize sensitivity. The first difference results from a wide-spread misunderstanding that the application of the reflectron is restricted to sources where ions are accelerated simultaneously from a well-defined plane (e.g., laser or plasma desorption). Cotter and coauthors used a similar method in their continuous SIMS.[19] The apparatus described in ref 20 was designed to probe ion clusters. Both orthogonal acceleration and accumulation of ions were used in this device. During 1991–1993, three more groups constructed TOF mass spectrometers with orthogonal acceleration of ions using atmospheric pressure ionization of gases[21] and ES ionization of solutions.[22,23]

The present paper describes our instrument and its further development with the emphasis on the results obtained with the electrospray ion source. The potential of this technique is demonstrated by the primary structure determination of cationic peptides from porcine neutrophils and by the study of the pathways of recombinant proinsulin proteolytic processing catalyzed by various trypsin forms. These examples show that a number of problems in protein chemistry can be solved by a direct ES MS identification of peptides in mixtures obtained by chemical or enzyme protein treatment. The reactions conditions or sample preparation procedures are chosen to ensure the possibility of direct ES MS analysis of unfractionated digests.

INSTRUMENTATION

ES TOF Mass Spectrometer Design. The electrospray ionization reflecting time-of-flight mass spectrometer is schematically shown in Figure 1. The electrospray ion source (1) (described below) is arranged outside the vacuum chamber (9), which is differentially pumped by a 5 L/s rotary pump and two 500 L/s turbomolecular pumps. Low-energy ions (5–25 eV) enter vacuum in the direction normal to the axis of the drift region (8), are focused into a parallel beam by a grid (3) and a quadrupole lens (4), and are introduced into a time-of-flight ion storage modulator (5) where they continue moving in field-free conditions. Approximately half of the ions entering vacuum through the skimmer orifice enter the ion storage modulator through a slit (2×10^{-2} mm^2) which separates the second and the third stages of the differential pump. A detailed description of the time-of-flight ion storage modulator has been given in refs. 11 and 24. In brief, it works in the following way. Initially, a repeller plate and the first grid are at the ground potential, and ions move toward the ion collector (6), which helps to adjust and monitor ion current entering the modulator. When the space between the plate and the first grid is filled with the input ions, both the push-out and draw-out voltage pulses are applied simultaneously to the plate and to the second grid, and the ion packet is drawn to the acceleration region. Ions entering the modulator at the pulse time are deflected by the push-out pulse and routed to the frame of the grid. To prevent field distortions in the acceleration region, guard frames are placed between the grids. The frames are connected by a built-in voltage divider. Ions are pushed out from the ion storage modulator (5) with a set frequency (approximately several kilohertz, depending on the largest mass to be analyzed) and accelerated to the drift region (8). A pair of deflection plates (7) is placed in the drift region to correct the ion trajectory. In order to keep ion source and ion storage modulator potentials close to ground, the drift region is floated at the accelerating voltage 2 kV. A cylindrical guard grid (11) serves to prevent the penetration of ground potential of the vacuum tube (9) to the drift region. Once having passed the drift region, the ions enter a two-stage electrostatic mirror (10). The ion mirror design is similar to that of the ion-storage modulator. It consists of two grids, 18 guard frames, and a plate forming two regions with uniform fields of different strength. The first grid is connected to the cylindrical guard grid, whereas the second grid and the plate potentials are adjusted to ensure time focusing of ions with different initial coordinates and energies.[24,25] Reflected ions pass the drift region once again (total drift length (L) of 1508 mm) and, finally, reach the detector (12). The latter is a secondary electron multiplier made of two rectangular microchannel plates (63×43 mm^2). The multiplier is isolated from the ground potential; therefore, postacceleration of heavy ions and negative ion detection are possible. A decoupling capacitor ($C = 4700$ pF) is used to transmit ion-current pulses to the ground base. The pulses are amplified and recorded by a homemade stroboscopic registration system with 10-ns resolution. Recently, the stroboscopic registration system was replaced by an integrating transient recording system (Strob Ltd., Novosibirsk, Russia). It is capable of storing transients up to 81.92 μs in length, with time resolution of 5 ns and with minimum dead time. Typically, from 1000 to 100 000 transients were accumulated. In the negative-ion registration mode, all potentials except that of the multiplier are reversed.

(19) Olthoff, J. K.; Lys, I. A.; Cotter, R. J. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 171–175.
(20) Kennedy, R. A.; Kung, C.-Y.; Miller, J. P. In *Ion and Cluster Ion Spectroscopy and Structure*; Maier, J. P., Ed.; Elsevier, New York, 1989; pp 213–239.
(21) Sin, C. H.; Lee, E. D.; Lee, M. L. *Anal. Chem.* **1991**, *63*, 2897–2900.
(22) Boyle, J. G.; Whitehouse, C. M. *Anal. Chem.* **1992**, *64*, 2084–2089.
(23) Verentchikov, A. N.; Standing, K. J.; Ens, W. *Anal. Chem.*, in press.
(24) Dodonov, A. F.; Chernushevich, I. V.; Laiko, V. V. In *Time-of-Flight Mass Spectrometry*; Cotter, R. J., Ed.; American Chemical Society: Washington, DC, in press.
(25) Mamyrin, B. A.; Karataev, D. V.; Schmikk, D. V.; Zagulin, V. A. *Sov. Phys.—JETP (Engl. Transl.)* **1973**, *37*, 45–53.

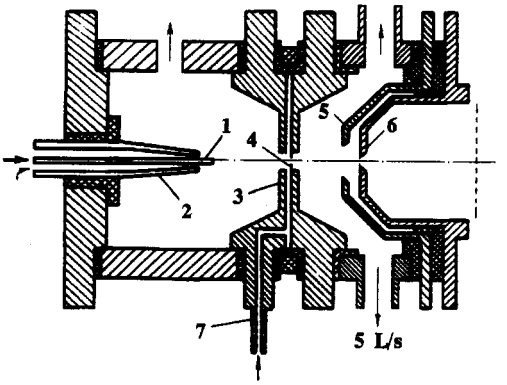


Figure 2. Detailed schematic diagram of the ES ion source: (1) stainless steel capillary (0.1-mm i.d., 0.5-mm o.d.), (2) glass capillary (1-mm i.d. at the tip), (3) gas curtain electrode, (4) nozzle (orifice diameter 0.4 mm), (5) additional focusing electrode, (6) skimmer (orifice diameter 0.1 mm), and (7) gas (N_2) inlet.

Table 1. Typical Values of Ion Current at Various Points of the Ion Trajectory

ion current from capillary, A	10^{-6}–10^{-7}
ion current through nozzle orifice, A	10^{-8}–10^{-10}
ion current of collector,[6a] A	10^{-11}–10^{-13}
ion current of detector,[12a] A	10^{-12}–10^{-15}

[a] The numbers correspond to positions in Figure 2.

Ion Source Design. The ES ion source is schematically shown in Figure 2. It is similar in part to those conventionally used.[2-4] A voltage of between +2.5 and +5 kV (for positive ions) is applied to the capillary (1). The distance between the tip of the capillary (1) and the counterelectrode (3) is 15–20 mm. A flow of dry air or oxygen (~100 cm^3/s, ambient temperature), introduced through a glass capillary (2), assists nebulization and suppresses corona discharge. A syringe pump controls the flow of analyte liquid at, typically, 2 μL/min. A countercurrent flow of dry nitrogen (100 cm^3/s, ambient temperature) through tube 7 is used as a gas curtain for large droplets. A voltage of 1 kV is applied between the electrode (3) and nozzle (4) to pull ions through the gas curtain toward the nozzle. Both the nozzle and skimmer electrodes are flat, and the orifices are made of 0.1-mm-thick stainless steel foil. A strong nonuniform field near the skimmer orifice, which usually helps to focus and decluster ions in the case of a conical skimmer, is obtained by an additional focusing electrode (5), placed 3.5 mm from the skimmer. Another more complicated ion source construction (not shown here) was used to obtain the optimum conditions for ion transmission through the skimmer orifice.[15] It allows the experimenter to move the nozzle along the axis with respect to the skimmer. It turned out that for any nozzle diameter in the range from 0.2 to 0.5 mm, and any pumping speed from 1 to 5 L/s, the nozzle–skimmer separation (Δz) should be $\Delta z \approx 2.5\ x_m$, where x_m is the distance between the nozzle orifice and the Mach disk. The empirical relationship provides the maximum ratio of ion current per neutral gas flow through the skimmer orifice. In our standard conditions ($\Delta z \approx 8$ mm), an additional requirement must be fulfilled for electrode potentials to provide focusing of ions on the skimmer orifice: nozzle potential (U_n) must be $\sim 0.75\Delta U$, where ΔU is the focusing electrode potential, with both potentials measured relative to skimmer. It is noteworthy that, under these conditions, ions move against the electric field part of their way from nozzle to skimmer. Typical voltages are ΔU = 200 V and U_n = 150 V, corresponding to total declustering and a low degree of fragmentation of ions. Typical values of ion current measured for different electrodes are listed in Table 1.

EXPERIMENTAL SECTION

Sample Preparation and Spectra Recording. For MS analysis, samples were dissolved in 15 μL of a 2% acetic acid/acetonitrile mixture (1:1, v/v). Routine sample concentration were ~10^{-4} M. The sample was injected at 2 μL/min flow rate using a microsyringe pump. Time of spectrum registration was 30 s (when the integrating transient recording system was used) so that 30 000 transients were accumulated for each spectrum.

Determination of Primary Structure of Porcine Neutrophil Peptides. The total preparation of porcine neutrophil peptides (PNPs) was obtained by Dr. V. N. Kokryakov and co-workers (Pavlov Institute for Experimental Medicine of the Russian Academy of Medicine Sciences, St. Petersburg) as described in ref 26.

The total preparation of PNPs was acetylated by an acetic anhydride/methanol mixture[1]) at 10 °C for 5 min ("mild acetylation") and 2) at 40 °C for 2 h ("hard acetylation"). The reaction mixtures were vacuum-dried and analyzed by ES MS.

The individual peptides from the total peptide preparation were purified by HPLC on a 2 × 62 mm column packed with Nucleosil 5 C_{18}. Elution was carried out in a 15–35% (v/v) linear gradient of acetonitrile (Merck) in 0.1% aqueous trifluoroacetic acid/acetonitrile gradient. Individual PNPs were reduced and alkylated by 4-vinylpyridine as described in ref 27. The reduced and pyridylethylated individual PNPs were digested with pepsin (Serva) in 1% acetic acid at 37 °C for 5 h. The enzyme/substrate ratio was 1:50. Digestion products were separated by HPLC and subjected to hydrolysis by heptafluorobutyric acid (HFBA) vapors as described in ref 28, with some modifications.

Amino acid composition was determined using a Model D-500 amino acid analyzer (Durrum). The PNPs' primary structure was determined by the Edman degradation technique, using a Model 816 automatic peptide sequencer (Knauer) equipped with a Model 120A PTH analyzer (Applied Biosystems), by Prof. Ts. A. Egorov and co-workers (Engelhardt Institute for Molecular Biology of the Russian Academy of Sciences, Moscow).

Recombinant Human Proinsulin Trypsinolysis Study. Recombinant human proinsulin (ProIns) was manufactured in the Shemyakin Institute for Bioorganic Chemistry of the Russian Academy of Sciences, Moscow. Trypsin from bovine pancreas was purchased from Serva. Trypsin immobilized on a silica matrix was produced as described in ref 29 and was packed into stainless steel column (2 × 62 mm).

(26) Mirgorodskaya, O. A.; Shevchenko, A. A.; Kamal Omer, M. A. A.; Chernushevich, I. V.; Egorov, Ts. A.; Musoliamov, A. X.; Kokryakov, V. N.; Shamova, O. V. *FEBS Lett.*, in press.
(27) Inglis, A. S.; McMahon, D. T. W.; Roxburgh, C. M.; Takayanagi, H. *Anal. Biochem.* **1976**, *72*, 86–94.
(28) Tsugita, A.; Takamoto, K.; Kamo, M.; Iwade, H. *Eur. J. Biochem.* **1992**, *206*, 691–696.
(29) Ivanova, G. P.; Mirgorodskaya, O. A.; Moskvichev, B. V. *Bioorg. Khim.* **1981**, *7*, 92–97 (in Russian).

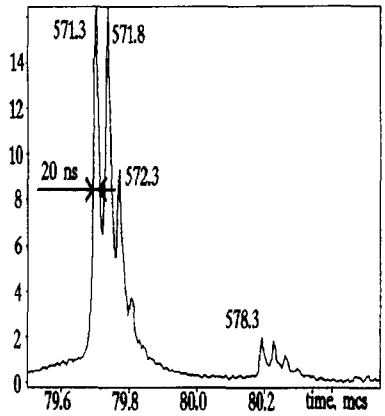


Figure 3. Portion of ES TOF mass spectrum of gramicidin S near the doubly charged ion $[M + 2H]^{2+}$ (monoisotopic m/z value is 571.3). The figure demonstrates the resolving power of the instrument.

ProIns was digested by trypsin in 0.1 M Tris-HCl buffer pH 8.0 at 25 °C. The initial ProIns concentration was 0.1 mM; the enzyme/substrate ratio was 1:1000 (w/w). At certain intervals, hydrolysis was stopped by the addition of 10% acetic acid. The aliquots of hydrolyzates were desalted on the 2 × 62 mm column packed with Nucleosil 5 C_{18}. Sample elution was performed with 50% acetonitrile in 2% acetic acid.

ProIns hydrolysis by immobilized trypsin was performed by passing 50 μL of proinsulin solution (0.1 mM concentration in 0.1 M Tris-HCl buffer, pH 8.0) through the column preequilibrated with the same buffer. The fraction was collected under spectrophotometric control followed by desalting on a HPLC column.

RESULTS AND DISCUSSION

ES TOF Mass Spectrometer Operational Features. The most significant limitation on the resolution of the ES TOF mass spectrometer may be imposed by the following factors: coordinate spread of ions in the ion storage modulator and consequent energy spread, initial energy spread of ions, deflection of ions in the ion drift region, rise time of the push-out and draw-out pulses, and ion scattering in a nonuniform field near the grids. The influence of the aforementioned factors and the means to increase resolution were discussed elsewhere.[24] Provided that all grids and plates are planar and parallel to each other, ions are focused into a parallel beam, and the deflection of ions is not used, the overall resolution of the ES TOF instrument is limited by the ion scattering from the grids. Taking into account ion scattering from the grids, the maximum resolution value is estimated to be 2500.[24]

To evaluate the resolving power experimentally, a monoisotopic peak of a doubly charged molecule of gramicidin S was used (Figure 3). (The scale of the ions' flight time is presented as the x-axis of all the spectra. The actual m/z values of the ions are presented near the corresponding peaks.) The total flight time of the molecular ion (m/z = 571.3) is t = 79.7 μs, time spread at half-height is 20 ns, and resolution is therefore 2000. Taking into account the time resolution of the registration system (5 ns), it is clear that the resolution of the mass spectrometer itself is much better. Mass resolution could be determined more precisely for the peaks of ions with a greater m/z ratio, if they were not broadened by an isotopic distribution. Furthermore, the time spread of heavy ions

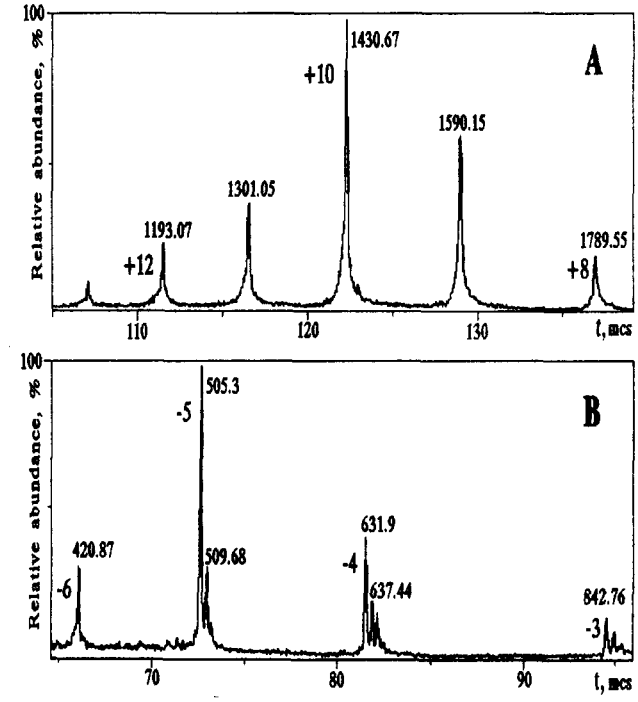


Figure 4. Mass spectra of hen egg lysozyme (A) and oxidized A-chain of bovine insulin (B) recorded in positive- and negative-ion modes, respectively. Molecular masses calculated from mass spectra data are shown in Table 2.

(starting with $m \approx 1000$ Da) is not proportional to flight time, as is generally assumed, and resolution decreases for higher masses.[24] The reasons for this fact are not yet clear. Insufficiently high vacuum can account for only a part of this additional time spread.

The accurate mass determination known to be the merit of ES MS is easily obtainable on the ES TOF instrument. One or two ion peaks in a mass spectrum are necessary to calibrate the mass scale. Usually, singly and doubly protonated ions of gramicidin S, with m/z = 1142.5 and 571.75, were used for calibration. Both the reference m/z values, and those obtained from the calibrated spectrum, were isotopically averaged. A typical accuracy of mass determination for ion masses is 0.02%; accuracy is a little worse for fragment ions and noisy spectra. The spectra of hen egg lysozyme (positive-ion mode) and bovine insulin A-chain (negative-ion mode) are presented in Figure 4. The results of the molecular mass determination for these and some other proteins are shown in the Table 2. A standard procedure described by Fenn[3] was used for the molecular mass calculation.

The average conventional sample concentrations were ~0.1 mM. The individual concentration limits are strongly dependent on the protein's molecular mass and amino acid composition (primarily from the content of basic amino acid residues). Hence, doubly charged ion of gramicidin S can be detected in lower than 10^{-6} M sample concentration.

The average sample amount necessary for spectra recording is estimated as 2 μL/min × 0.5 min × 10^{-4} M = 10^{-10} M, which is comparable with that conventionally necessary for quadrupole instruments.[1,2] It should be mentioned that sample amount is mostly limited by preceding sample preparation procedures (e.g., by sample dissolving and uptaking) than by the sensitivity of mass spectrometer.

Table 2. Analysis of Proteins on the ES TOF Mass Spectrometer

protein	molecular mass, Da: calcd from protein sequence	measd
melittin	2846.5	2846.3 ± 0.2
human insulin	5807.6	5807.6 ± 0.4
human proinsulin	9388.6	9390 ± 1.3
ribonuclease A[a]	13682.2	13703 ± 3
ribonuclease A[b]	14531.3	14533 ± 3
lysozyme	14306.2	14304 ± 3
recombinant protein of proinsulin	16671.6	16678 ● 3
oxidized A-chain of bovine insulin	2531.6	2531.4 ± 0.2

[a] Sodium or water adduct's peak is not resolved (peak maximum used). [b] Disulfide-reduced and pyridylethylated. [c] The spectrum was recorded in negative-ion mode.

Applications of ES TOF Mass Spectrometer in Protein Chemistry. (1) Porcine Neutrophil Peptide Mapping and Sequencing. As mentioned above, the principal advantages of ES MS are the determination of protein molecular masses with high accuracy and direct qualitative analysis of peptide or protein mixtures without preliminary fractionation.[1-8] These features enable one to form a MS map of proteins,[1,8] to reveal posttranslational modifications,[1,2,26] to determine partial N- or C-terminal sequences by acid or exopeptidase degradation,[6,26] and to verify the determined sequence by the exact molecular mass value. The combined usage of the ES TOF mass spectrometer and the traditional Edman degradation technique greatly facilitates primary structure determination, especially when low molecular weight, homologous polypeptides are studied. In the present work, the primary structure of three main antimicrobial cationic peptides[30] from porcine neutrophils (PNPs) was determined by the aforementioned approach.

Direct ES MS analysis of the total peptide preparation (Figure 5A) reveals three peptides with the masses shown in Table 3.

Mass spectra of the reaction mixtures obtained by treatment of the total PNP preparation with an acetic anhydride/methanol mixture are shown in Figure 5B,C. After "mild" acetylation, when only the N-terminal amino group is known to be acetylated,[31] the masses of all PNPs were increased by 42 Da (Figure 5B, Table 3) corresponding to only one acetyl group attached to each peptide. The "hard" acetylation of the total preparation yields only mono- and diacetyl derivatives of each of the PNPs with the complete absence of methyl ester formation (Figure 5C, Table 3). As was previously found in model experiments, all free COOH groups in peptides undergo a complete esterification under these conditions (data not shown). It can be concluded that C-terminal COOH groups of all the PNPs appear to be blocked. At the same time, each PNP contains a free N-terminal amino group and one amino acid residue with an OH or NH_2 group in its side chain.

Individual PNPs were purified by HPLC, and their amino acid compositions were determined (Table 4). The Table 4 data show that the second site of acetylation in each PNP is at the OH group of a Tyr residue.

(30) Lehrer, R. I.; Ganz, T.; Selsted, M. E. *Cell* **1991**, *64*, 229–230.
(31) Knapp, D. R. *Methods Enzymol.* **1990**, *193*, 314–329.

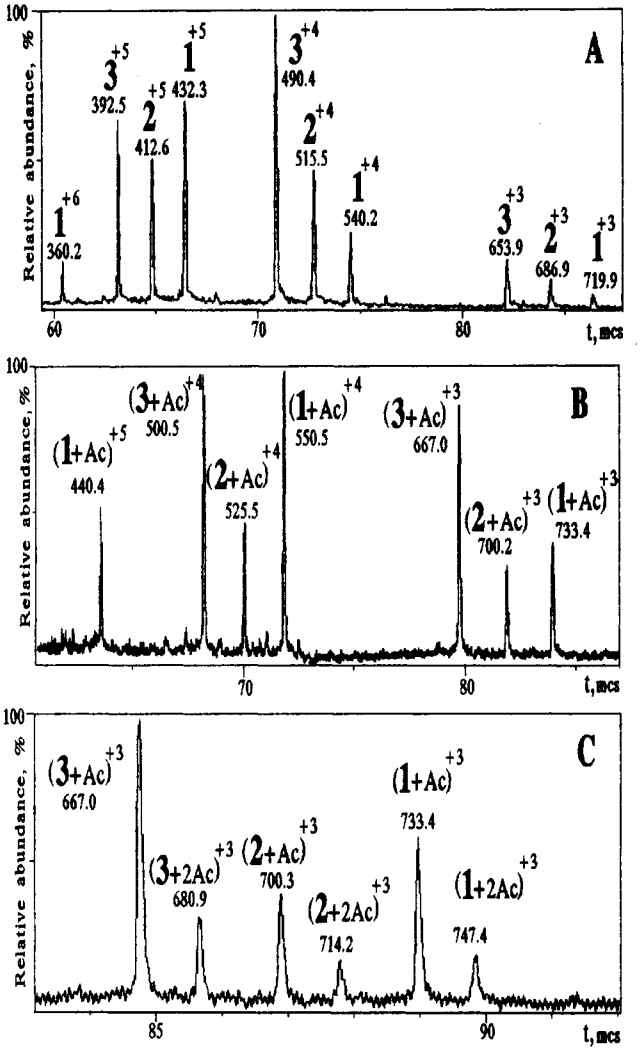


Figure 5. Mass spectra of the total preparation of PNPs: (A) native preparation; (B, C) after its acetylation under "mild" and "hard" conditions, respectively. Only the part of the spectrum corresponding to triply charged ions is presented in (C). The ions of PNP 1, PNP 2, and PNP 3 are designated as 1, 2, and 3, respectively. Acetyl residue is marked as Ac. Molecular masses calculated from mass spectra data are shown in Table 3.

Table 3. Results of Mass Spectrometric Analysis of Native and Chemically Derivatized Total PNP Preparation[a,b]

peptide	mol mass[c]	native prep	after treatment with acetic anhydride/methanol	after reduction and pyridylethylation
PNP 1	2155.7	2156.6 ± 0.5	2197.2 ± 0.2 2238.9 ± 0.1	2580.7 ± 0.4
PNP 2	2056.5	2057.6 ± 0.5	2098.1 ± 0.1 2140.1 ± 0.4	2481.8 ± 0.1
PNP 3	1956.5	1957.5 ± 0.5	1998.1 ± 0.1 2039.7 ± 0.1	2381.4 ± 0.2

[a] All masses are obtained by mass spectra deconvolution in accordance with the procedure described by Fenn[3] and expressed in daltons. [b] The peptides are numbered in accordance with the order of elution from a reversed-phase HPLC column. [c] Mass calculated from sequence data.

Treatment of the PNP preparation with Ellman's reagent showed the absence of free SH groups. After reduction of the PNP by β-mercaptoethanol, followed by pyridylethylation of the formed cysteine residues, the molecular masses of each of the PNPs were increased by 424.6 Da (Table 3). This shows that each of the PNPs contains four cysteine residues involved

Table 4. Main Characteristics for the Peptides from Porcine Neutrophils

peptide	measd amino acid comp[a]			mass spectrometric mapping data: mol mass, Da	mass spectrometric mapping data: variants of amino acid comp[b]
PNP-1	Gly	3.1 (3)	P_1	557.7	G_2R_2L (VG_3RL)
	Val	2.1 (2)			
	Leu	1.1 (1)	P_2	389.3	CY
	Tyr	0.8 (1)			
	Phe	0.9 (1)	P_3	841.9	CR_3F ($CGVR_2F$, G_2V_2RCF)
	Arg	5.8 (6)			
	Cys_2	2 (2)	P_4	844.8	V_2C_2GR (VR_2C_2)
PNP-2	Gly	4.1 (4)	P_1	458.5	G_3RL
	Val	2.1 (2)			
	Leu	1.1 (1)	P_2	389.3	CY
	Tyr	0.8 (1)			
	Phe	1.0 (1)	P_3	841.9	CR_3F ($CGVR_2F$, G_2V_2RCF)
	Arg	4.9 (5)			
	Cys_2	2 (2)	P_4	844.8	V_2C_2GR (VR_2C_2)
PNP-3	Gly	2.1 (2)	P_1	557.7	G_2R_2L (VG_3RL)
	Val	1.0 (1)			
	Leu	1.6 (1)	P_2	389.3	CY
	Tyr	0.8 (1)			
	Phe	1.0 (1)	P_3	841.9	CR_3F ($CGVR_2F$, G_2V_2RCF)
	Arg	4.9 (5)			
	Ile	0.9 (1)	P_4	645.6	C_2VI
	Cys_2	2 (2)			

[a] Found from sequence data. [b] C, pyridylethylcysteine residues. Alternative variants of amino acid composition are shown in parentheses. The symbols of amino acid residues are given in one-letter code. The amount of cystine groups was determined mass spectrometrically (see text).

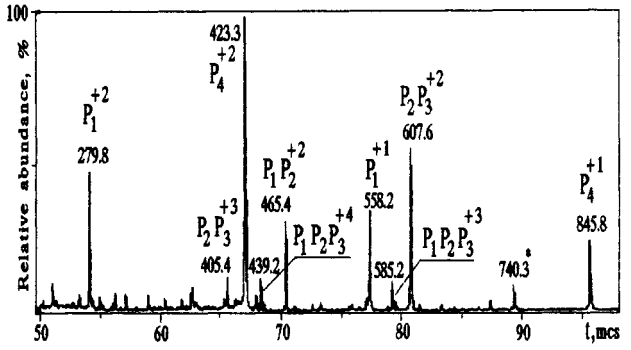
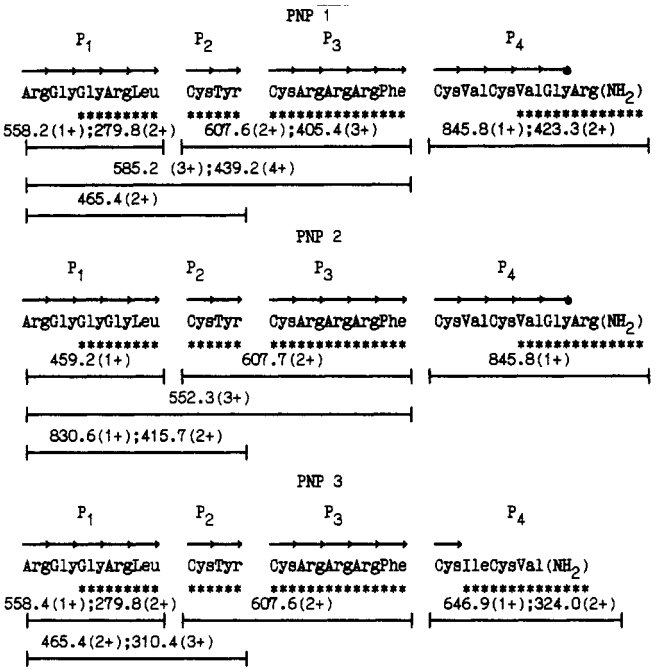


Figure 6. Mass spectrum of the pepsin digest of reduced and pyridylethylated PNP 1. Ion identification is given in the Table 4 and Scheme 1. The fragment ion of P_4 ion formed by pyridylethyl group loss (M − 105.2) is marked by an asterisk.

in two intermolecular disulfide bonds. Thus, from such an intrinsic property as molecular weight, the number of disulfide bridges and the absence of free C-terminal carboxylic groups were revealed simultaneously for each of the PNPs by direct MS analysis of the initial and chemically derived total preparations.

For ES MS mapping and sequencing, individual PNPs were reduced and alkylated by 4-vinylpyridine. The treated PNPs were digested with pepsin. The obtained pool of products was injected directly into the mass spectrometer. As an example, the mass spectrum of a pepsin digest of PNP 1 is presented in Figure 6. MS analysis clearly demonstrates the presence of four structural fragments, revealed by comparison of the masses of the products obtained by complete (P_1, P_4) and partial (P_1P_2, P_2P_3, $P_1P_2P_3$) cleavage of PNPs (Table 4, Scheme 1). All PNPs have two identical structural fragments

Scheme 1. Amino Acid Sequences of the Peptides from Porcine Neutrophils[a]



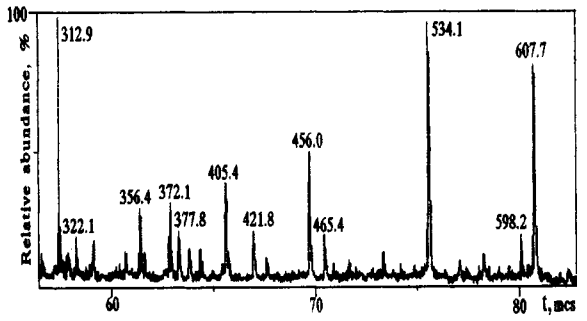
[a] Sequences obtained by Edman degradation are shown by arrows. Sequences determined by ESI MS are shown by asterisks. Fragments identified by ESI MS mapping are shown by lines with the corresponding *m/z* and *z* values.

(P_2, P_3). The sequence of the P_3 fragment can only be Cys-Tyr, as follows from the PNP amino acid composition and the known substrate specificity for pepsin.

The amino acid composition of fragments P_1, P_3, and P_4 cannot be calculated unambiguously from only their masses. However, as follows from the calculation of the amino acid composition, P_1 and P_3 fragments contain Leu and Phe residues, respectively. Since pepsin hydrolyzes polypeptides at Leu, Phe, and Tyr residues, this means that only the P_4 fragment is a C-terminal one. All masses of P_4 calculated from their amino acid composition exceed their measured values by 1 Da (Table 4, Scheme 1). Consequently, the C-terminals of all PNPs appear to have an amide group.

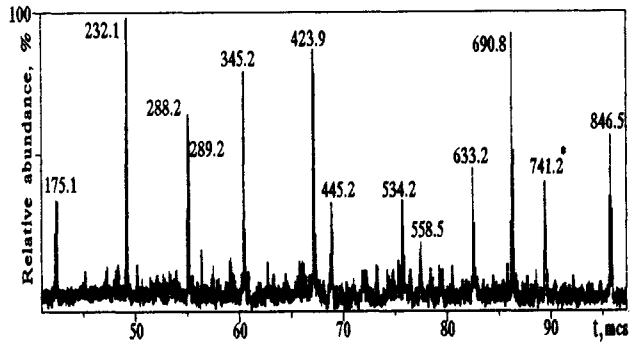
The attempts to sequence the fragments P_1–P_4 were done using a PNP pepsin digest separated into two fractions: one was the P_2P_3 fragment and the other was a mixture of P_1 and P_4 fragments. The fractions were treated with HFBA vapors to perform a partial C-terminal peptide hydrolysis. As an example, mass spectra of the HFBA hydrolyzates of PNP 1 fractions and their interpretation are presented in Figure 7 and Figure 8. PNPs sequences determined from ES MS data are shown in Scheme 1.

The obtained data show that the initially proposed peptide hydrolysis by HFBA vapors[28] actually leads to C-terminal degradation. Nonspecific cleavage at both sides of Gly residues appeared to be negligible in spite of a high Gly content. It should also be mentioned that the consecutive C-terminus degradation was not halted by the presence of a $CO(NH_2)$ group at the C-terminals of the PNPs. This method appears to be a good alternative to the often used digestion by carboxypeptidase Y when the C-terminus sequence is to be determined.



MH^+	MH_2^{+2}	MH_3^{+3}	P_2P_3	
	607.7	405.4	Cys*-Tyr-Cys*-Arg-Arg-Arg-Phe	
	534.1	356.4	Cys*-Tyr-Cys*-Arg-Arg-Arg	ΔM = 147
	456.0		Cys*-Tyr-Cys*-Arg-Arg	ΔM = 156
	377.8		Cys*-Tyr-Cys*-Arg	ΔM = 156
598.2			Cys*-Tyr-Cys*	ΔM = 156

Figure 7. Mass spectrum of the HFBA (vapor) hydrolyzate of P_2P_3 fragment of PNP 1.



MH^+	MH_2^{+2}	P_4	
846.5	423.9	(Cys_2,Val)-Val-Gly-Arg	
690.8	345.2	((Cys_2,Val)-Val-Gly	ΔM = 158
633.2		(Cys_2,Val)-Val	ΔM = 57
534.2		(Cys_2,Val)	ΔM = 99
MH^+	MH_2^{+2}	P_1	
558.5	279.8	(Arg,Gly)-Gly-Arg-Leu	
445.2		(Arg,Gly)-Gly-Arg	ΔM = 113
289.2		(Arg,Gly)-Gly	ΔM = 147
232.1		(Arg,Gly)	ΔM = 57

Figure 8. Mass spectrum of the HFBA (vapor) hydrolyzate of P_1 and P_4 fragments of a PNP 1 mixture. These fragments were unresolved by HPLC. The fragment ion of P_4 ion formed by pyridylethyl group loss (M – 105.2) is marked by an asterisk. Amino acid compositions of the N-terminal sections of P_1 and P_4, calculated from their molecular masses, are shown in parentheses.

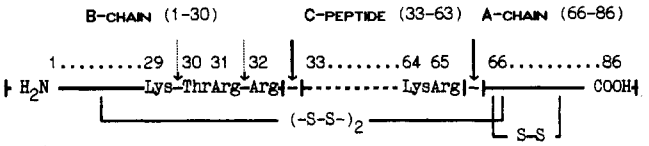
The overlapping sequences of products obtained by partial peptic digestion of PNPs (Scheme 1) allow for determination of the location of P_1–P_2 structural fragments in the total sequence. The data obtained show that the location of P_1–P_4 is in agreement with their numbering.

It can be seen that only N-terminal sequences of P_1 and P_4 fragments remain ambiguous. To complete and confirm PNP sequences, the Edman degradation technique was applied (Scheme 1).

In conclusion, it should be mentioned that study of the primary structure of homologous polypeptides reveals some advantages and disadvantages of the ES TOF mass spectrometric technique.

Mass spectrometric analysis of the partial enzyme digest of a polypeptide chain enables discovery of the structural

Scheme 2. Trypsinolysis Pathways of Human Proinsulin[a]



[a] The sites of primary proinsulin cleavage are marked by solid arrows. The sites of secondary cleavages of Ins-RR are marked by dotted arrows.

Table 5. Mass Spectrometric Identification of Proinsulin Hydrolysis Products

	mol mass, Da	
product[a]	calcd[b]	measd
ProIns	9389.4	9390.8 ± 1.3
Intermed	9407.4	9408.7 ± 1.3
Ins-RR	6120.6	6120.0 ± 1.0
Ins-R	5963.9	5964.1 ± 1.3
desT-Ins	5707.1	5707.5 ± 0.5
C-KR	3304.8	3305.5 ± 0.4

[a] The abbreviations used are explained in the text. [b] From sequence data.

fragments and determination of their location. Comparison of masses of structural fragments reveals the polypeptides which are homologous and thus allows the structural fragments with amino acid substitutions to be located.

The relatively low resolution of the TOF mass spectrometer makes it impossible to measure the values of monoisotopic masses for polypeptides so that the mass difference of 1 Da cannot hardly be reliably calculated from the series of three to five charged ions of the native PNPs (Figure 5A, Table 3). The C-terminal amidations discerned through measurement of the masses of the P_4 fragments are therefore more reliable (the m/z values of singly and doubly charged ions were measured). Moreover, the TOF mass spectrum can be overcalibrated by using the two exact m/z values of ions corresponding to the interior fragments of PNP's sequences (P_1P_2 or P_2P_3 in our case).

MS analysis of peptide hydrolyzates by HFBA vapors enables determination of partial C-terminal sequences.

At the same time, the ESI MS technique does not allow determination of Leu and Ile positions in the polypeptide sequence, so they must be determined by other methods.

(2) ES Mass Spectrometric Monitoring of the Proteolytic Processing of Recombinant Human Proinsulin. The limited proteolysis of pro-proteins is an important and commonly used procedure in the manufacturing of several recombinant biologically active proteins.[32] The monitoring of pro-proteins proteolysis in order to ensure their selective and complete cleavage is a common problem in analytical biotechnology.[1,32] HPLC separation of the reaction mixture is commonly used to solve this problem. Chromatographic peaks are usually identified by their retention times (if the standards of the proteolysis products are available) or by amino acid analysis data. This approach is labor- and time-consuming. The results obtained are crucially dependent on the peaks' homogenity and on the accuracy of amino acid analysis.

The present work shows that the monitoring procedure can be sufficiently simplified and accelerated by ES TOF

(32) *Biotechnology. Principles and Applications*; Higgins, I. J., Best, D. J., Jones, J., Eds.; Blackwell Science Publ.: London, 1985.

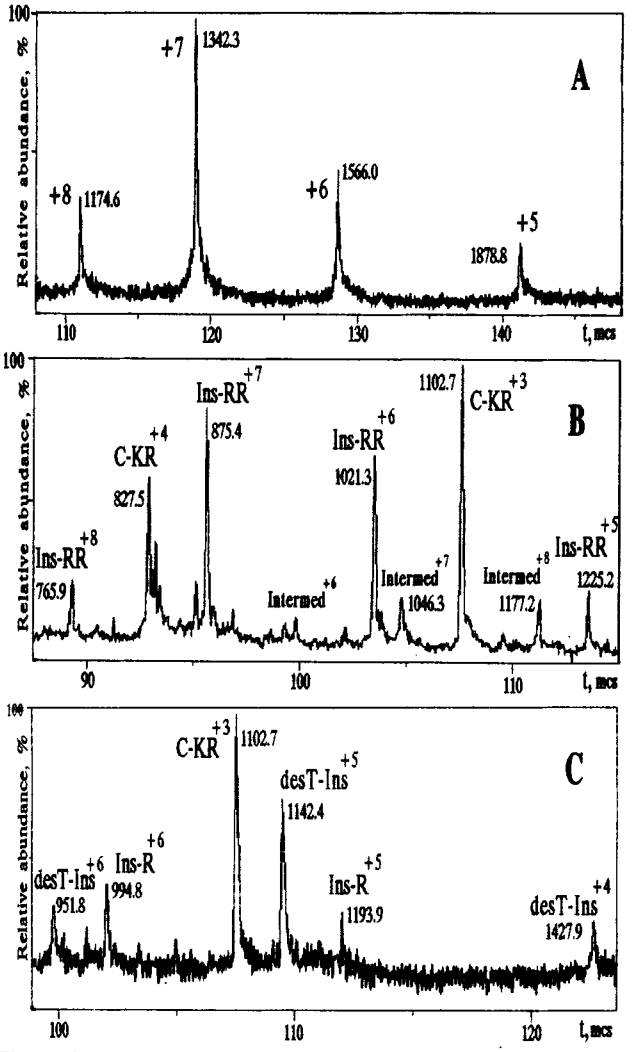


Figure 9. Mass spectra of the desalted hydrolyzates of human recombinant proinsulin by native trypsin: (A) hydrolysis time $t = 0$; (B) $t = 10$ min; (C) $t = 3.5$ h. The abbreviations used are explained in the text. The molecular masses of the proteolysis products, calculated from mass spectra data, are presented in Table 5.

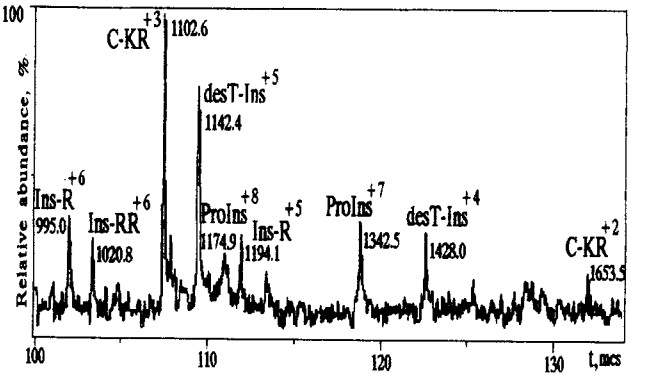


Figure 10. Mass spectrum of the desalted hydrolyzate of human recombinant proinsulin by silicate matrix-immobilized trypsin. The abbreviations used are explained in the text. The molecular masses of the proteolysis products, calculated from mass spectra data, are presented in Table 5.

mass spectrometer application. This work considers a comparative study of recombinant human proinsulin (ProIns) proteolysis by native and silicate matrix-immobilized trypsin. The trypsinolysis of ProIns is used to produce recombinant human insulin by the process technology developed in the Shemyakin Institute for Bioorganic Chemistry of the Russian Academy of Sciences, Moscow.

It is known that ProIns trypsinolysis is accompanied by the formation of two intermediate products associated with the hydrolysis of C-terminal peptide bonds of the pairs of basic amino acids—Arg-Arg and Lys-Arg.[33] Further cleavage of both intermediate products (having equal molecular weights and therefore named as Intermed) gives the desired [Arg^{B31},-Arg^{B32}]insulin (Ins-RR) and Arg,Lys-C-peptide (C-KR) (Scheme 2). The masses of possible proteolysis products are listed in the Table 5.

Tryptic digests of ProIns at different stages of hydrolysis were desalted on the HPLC column and injected into the mass spectrometer (Figure 9). Initially ($t = 0$), only the series of multiply protonated proinsulin ions were registered (Figure 9A). The measured molecular mass of proinsulin corresponds to the calculated mass from the sequence data. With increasing hydrolysis time, a product pool, formed in fairly good agreement with Scheme 2, was discovered by ES MS (Figure 9B). Proteolysis products were identified by their molecular masses calculated from mass spectra data (Table 5). Mass spectra data show that Ins-RR is the main product of ProIns trypsinolysis during the initial period of the reaction. Continuation of the trypsinolysis results in Ins-RR digestion yielding Ins-R and desT-Ins products (Figure 9C).

For a number of technological reasons, the use of immobilized enzymes is preferable to use of native ones.[29] However, the mass spectrum of ProIns digest by immobilized trypsin unambiguously shows that in this case ProIns hydrolysis does not give the desired Ins-RR, but rather the unwanted products of further Ins-RR cleavage—[Arg^{B31}]insulin (Ins-R) and des-Thr^{B30}-insulin (desT-Ins) (Figure 10). The degradation of Ins-RR takes place when the reaction mixture still contains the initial ProIns.

The present work unambiguously shows that trypsin immobilization results in sufficient changes in its specificity with regard to high molecular weight substrates. Consequently, the composition of the product pool yield following repeated proteolytic cleavage appears to be unpredictable. Therefore, the approach proposed is possible when manufacture processes are developed.

ES TOF mass spectrometer application enables one to identify the proteolysis products without preliminary fractionation of the digests. Standards are not needed for the identification. At the same time, two limitations inherent for the ES technique were also encountered: (1) The samples to be analyzed need to be first desalted. (2) The technique does not allow simultaneous quantitative measurement of the concentrations of digestion products, so combined usage with HPLC is necessary to solve this problem.

CONCLUSIONS

We have shown here that the application of the TOF mass spectrometry is not confined to pulsed ionization methods. Most of the continuous ionization techniques may be coupled to TOF mass analyzers in the way described in this paper. The main advantage of the original construction lies in the fact that a large sensitivity is compatible with a moderate

(33) Steiner, D. F.; Docherty, K. *Annu. Rev. Physiol.* **1982**, *44*, 625–638.

mass resolution. The initial energy spread of ions affects, only slightly, both resolution and sensitivity. An integrating transient recording system allows simultaneous registration of ions of all masses. The possibility of fast spectrum acquisition together with the pneumatically assisted nebulization of large amounts of analyte liquid enables TOF MS to be interfaced with high-performance liquid chromatography.

The technical features of the device enable its successful application for structural and functional studies of proteins and polypeptides. The results obtained show that ES mass spectrometry is a powerful tool for direct qualitative analysis of peptide and protein digests. Digestion products can be identified without their preceding purification and without the use of any additional techniques such as amino acid analysis. ES MS appeared to be a good supplement to the traditional methods used for primary structure determination. Peptide mapping by ES MS allows not only discovery of the definite structural fragments in a protein's sequence but also determination of its relative location by accounting for the overlapping amino acids. Peptide hydrolysis by HFBA vapors, followed by MS analysis of the hydrolyzate, enables determination of a partial C-terminal sequence. It is a good supplement to the Edman sequencing technique in which the C-terminal sequences are often left undetermined.

ACKNOWLEDGMENT

The authors are grateful to Prof. Ts. A. Egorov and A. X. Musoliamov for Edman sequencing of PNPs, Dr. V. N. Kokryakov and O. V. Shamova for the kind gift of the total PNP preparation, Dr. A. V. Tolmachev for computer simulation of the ion trajectories and S. V. Dobretsov for computer data processing. The authors are grateful to academician of the Russian Academy of Sciences D. G. Knorre for his support of the present work and useful advise.

Received for review April 7, 1993. Accepted September 24, 1993.•

• Abstract published in *Advance ACS Abstracts,* November 1, 1993.