

# Rapid Determination of Parvalbumin Amino Acid Sequence from *Rana catesbeiana* (pI 4.78) by Combination of ESI Mass Spectrometry, Protein Sequencing, and Amino Acid Analysis

Hikari Taka,\* Naoko Kaga,\* Tsutomu Fujimura,\* Reiko Mineki,\* Masamoto Imaizumi,† Yoshikazu Suzuki,‡ Rintaro Suzuki,‡ Masaru Tanokura,‡ Noriko Shindo,\* and Kimie Murayama\*<sup>1</sup>

\*Division of Biochemical Analysis, Central Laboratory of Medical Sciences, Juntendo University School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo 113-8421, †Department of Ophthalmology, Medical University of Oita, Hasama, Oita 879-5593, and ‡Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657

Received November 24, 1999; accepted January 31, 2000

The complete amino acid sequence of  $\beta$ -type parvalbumin (PA) from bullfrog *Rana catesbeiana* (pI 4.78) was determined by tandem mass spectrometry in combination with amino acid analysis and peptide sequencing following Arg-C and V<sub>8</sub> protease digestion. The primary structure of the protein was compared with that of  $\beta$ -type PA from *R. esculenta* (pI 4.50), with which it is highly homologous. Compared with *R. esculenta*  $\beta$ -type PA4.50, *R. catesbeiana*  $\beta$ -type parvalbumin (PA 4.78) differed in 15 out of 108 amino acid residues (14% displacement), PA4.78 had Cys at residue 64 and was acetylated at the amino terminus, but 25 residues of the carboxyl terminus were completely conserved. Several amino acid displacements were found between residues 51 and 80 (30% displacement), although the functionally important sequence of PA was completely conserved. The amino acids residues of putative calcium-binding sites were Asp-51, Asp-53, Ser-55, Phe-57, Glu-59, Glu-62, Asp-90, Asp-92, Asp-94, Lys-96, and Glu-101, which were conserved in all  $\alpha$  and  $\beta$ -types of *R. catesbeiana* as well as other parvalbumins. In addition, Arg-75 and Glu-81, which are thought to form a salt bridge located in the interior of the molecule [Coffee, C.J. *et al.* (1976) *Biochim. Biophys. Acta* 453, 67–80], were also conserved in PA4.78.

**Key words:** amino acid analysis, electrospray mass spectrometry, parvalbumin, protein sequencer, sequence.

Parvalbumin (PA) is a calcium-binding protein found in muscles, especially in lower vertebrates such as fishes, amphibians, and reptiles (1). It is an acidic protein with a molecular weight of ~12,000, which varies slightly among species, and two high-affinity calcium-binding sites. Its function remains unknown, but it is thought to be involved in the relaxation process (2, 3) and in labile heat production during muscle contraction (4).

Frog muscles, which are frequently used in physiological studies, contain parvalbumins of two types,  $\alpha$  and  $\beta$ , which are differentiated by their isoelectric points and amino acid sequences (5). Both  $\alpha$  and  $\beta$  types of PA have been identified in *Rana esculenta* and their amino acid sequences have been determined (6, 7). Tanokura *et al.* (8) have examined the biochemical properties of two forms of PA, pI 4.78 and pI 4.97, isolated from skeletal muscles of the bullfrog, *Rana catesbeiana*; and the amino acid sequence of  $\alpha$ -type PA (pI 4.97) from *R. catesbeiana* (PA4.97) has been reported (9). Comparison of the amino acid sequences of  $\alpha$ -type PA from *R. esculenta* and that from *R. catesbeiana* showed a difference of only 23 amino acids (21%), and therefore these se-

quences were considered to be highly homologous (9).

In the present study, we determined the amino acid sequence of  $\beta$ -type PA from *R. catesbeiana* (PA4.78). There are two methods for determining the primary structure of proteins. One is based on the DNA sequence of the gene precursor, the other on the amino acid sequence of the protein. The former is comparatively simple but does not allow the determination of any post-translational modifications that may occur. The latter method, on the other hand, was until recently beset by various methodological difficulties such as the amount of sample required. Recent technological advances, however, have allowed the use of very small samples for microanalysis. Here, we determined the structure of PA4.78 by tandem mass spectrometry, amino acid analysis, and peptide sequencing following enzymatic digestion, using 1 mg of  $\beta$ -type PA from *R. catesbeiana*.

## MATERIALS AND METHODS

**Reagents**—2-Mercaptoethanol (2-ME), 4-vinylpyridine, trifluoroacetic acid (high-performance liquid chromatography [HPLC] grade), and acetonitrile (HPLC grade) were from Wako Pure Chemical Industries (Osaka). Guanidine hydrochloride was from Nacalai Tesque (Kyoto). Trizma base was from Sigma Chemical (St. Louis, MO). Arginylendopeptidase (Arg-C; mouse submandibular protease) was from Takara Shuzo (Shiga), and V<sub>8</sub> protease (Glu-C; *Staph-*

<sup>1</sup>To whom correspondence should be addressed Tel: +81-3-5802-1113, Fax +81-3-3818-6330, E-mail. murayama@med.juntendo.ac.jp

*Staphylococcus aureus* protease) was from Pierce Chemical (Rockford, IL).

**Preparation of Parvalbumins**—Parvalbumins were extracted from bullfrog skeletal muscles by the TCA method (10) and purified by column chromatography as described previously (11) with a slight modification (8). All procedures were performed at 0 or 4°C. Solutions used always contained 3 or 15 mM 2-ME. The purified proteins were concentrated by TCA treatment and dialyzed against 1 mM NaHCO<sub>3</sub> containing 2 mM dithiothreitol without the addition of CaCl<sub>2</sub>.

**High-Performance Liquid Chromatography**—HPLC was performed on a Gilson dual pump high-pressure mixing system with the detector (Gilson Medical Electronics, Middleton, WI) operated at 230 nm and 270 or 254 nm (pyridylethylated protein). HPLC was run on a Cosmosil C<sub>18</sub> column (4.6Φ × 250 mm) (Nacalai Tesque) at a flow rate of 1 ml/min. Solvents A and B were used for all runs. Solvent A was 0.1% aqueous trifluoroacetic acid (TFA) and Solvent B was 70% acetonitrile containing 0.1% TFA.

**Pyridylethylation**—PA (100 nmol) was dissolved in 5 ml of 25 mM Tris-HCl containing 6 M guanidine and 1 mM EDTA (pH 8.5), and 250 µl of 10% 2-ME was added. Nitrogen was then introduced into the reaction mixture, the mixture was stirred at room temperature for 30 min, and 200 µl of 4-vinylpyridine was added. Nitrogen was again introduced, the mixture was incubated for a further 30 min, then allowed to stand in ice for 2 h. Pyridylethylated PA was purified by HPLC; the column was developed isocratically with solvent A for 20 min, then with a linear gradient to 100% solvent B at a rate of 2%/min. Eluted protein was detected at 230 nm for peptide bonds and at 254 nm for pyridylethyl cysteinyl residues. The fraction showing the main peak was collected and subjected to arginylendopeptidase digestion.

**Enzymatic Digestion**—(i) *Arginylendopeptidase digestion*: Pyridylethylated PA (20 nmol) was dissolved in 50 mM phosphate buffer (pH 8.0) and digested with arginylendopeptidase (Arg-C) (ratio of enzyme to substrate, 1:50) at 37°C for 16 h. The resulting peptides were separated by reverse-phase HPLC using a linear gradient of 10–100% solvent B at 2%/min.

(ii) *V<sub>8</sub> protease digestion*: Native PA (50 nmol) was dissolved in 50 mM ammonium acetate (pH 4.0) and digested with V<sub>8</sub> protease (Glu-C) (ratio of enzyme to substrate, 1:50) at 37°C for 18 h. The digested peptides were separated by reverse-phase HPLC using 10–100% solvent B in a linear gradient at 1%/min.

**Mass Spectrometry**—Electrospray ionization mass spectrometry was performed on a Finnigan Mat TSQ-700 triple quadrupole instrument (Thermoquest Finnigan Mat Instruments, San Jose, CA). Spray voltage was 4.5 kV and capillary temperature 150°C. The sample was dissolved in 0.25% acetic acid/50% methanol at a concentration of approximately 10 pmol/µl. Scans were acquired in profile mode in the *m/z* range of 200 to 2,500 over the course of 5 s. ESI-MS/MS spectra were acquired in the centroid mode. Argon was used as the target gas at a pressure of 2.7–3.0 mTorr. The collision offset voltages used in Q2 were –15 to –30 V. The sample was introduced into the Electrospray ionization ion source at 1 µl/min with an infusion pump.

**Peptide Sequencing and Amino Acid Analysis**—Automated peptide sequence analysis was performed with an

HP model H1500G pulsed liquid phase microsequencer (Hewlett-Packard Co., Palo Alto, CA). Amino acid analyses of PA and peptides were carried out using a Hitachi 835 amino acid analyzer (Hitachi, Ibaragi) after gas-phase hydrolysis by 6 M HCl including 0.2% phenol at 165°C for 25 min.

**Deacetylation of N-Terminal Ser of PA4.78**—Five hundred pmol of PA4.78 spotted onto a polyvinylidene difluoride membrane was deacetylated by gas-phase hydrolysis with 100% TFA at 45°C for 15 min.

## RESULTS AND DISCUSSION

**Purity of PA4.78 and Pyridylethylation**—The purity of PA4.78 was examined by HPLC and ESI-MS before determining its amino acid sequence. PA4.78 showed a single peak in the HPLC profile (Fig. 1). The ESI mass spectrum of PA4.78 indicated that the charge distribution was between +6 and +15. The molecular weight was measured at 11,757 from the deconvoluted spectrum. Although a minor peak appeared at 11,789 (11,757 + 32), it was considered to

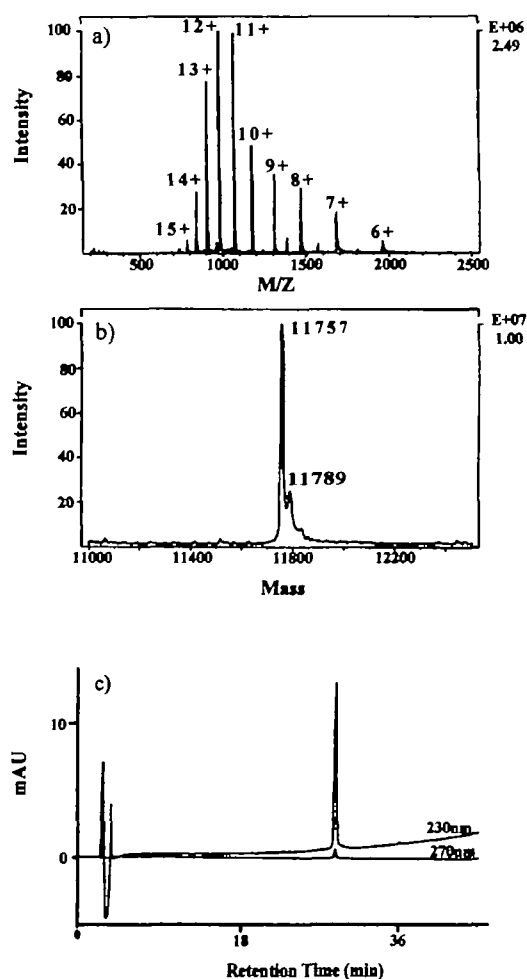


Fig. 1 Electrospray ionization mass spectra and HPLC profile of native PA4.78. (a) Convolved mass spectrum, (b) deconvoluted spectrum, and (c) HPLC profile. HPLC was performed using a linear gradient of solvent B (10–100%, 45 min), and protein was detected at 230 and 270 nm.

be a single peak under these conditions.

Since the previously reported amino acid composition suggested the presence of Cys, PA4.78 was pyridylethylated before enzymatic digestion (8, 12).

**Digestion of Pyridylethylate-PA4.78 with Arginylendopeptidase (Arg-C)**—Since the highly homologous  $\beta$ -type *R. esculenta* PA4.50 contains three Arg, 11 Lys, and eight Glu residues (7), and PA4.78 had similar residue proportions (8, 12), we presume that Arg-C digestion results in one large and two small fragments of PA4.78. However, when specific digestion of the C-terminal side peptide bond of Arg residues was carried out for only 2–3 h, native PA4.78 was insufficiently digested. When the reaction was performed for 16 h for complete digestion of the native PA4.78, some C-terminal side peptide bond of Lys residues were cleaved in addition to the C-terminal side peptide bond of Arg residues (data not shown). After these preliminary experiments, pyridylethylated PA4.78 was digested for 16 h. This resulted in the appearance of more than 12 peaks, referenced as peaks A1 to A12 in Fig. 2. Although this prolonged digestion reduced enzymatic specificity, it allowed us to get appropriately long fragments. If PA4.78 had been digested with Lys-C, it might have been difficult to put many small fragments together.

The molecular weight, amino acid composition, and peptide sequence of the obtained peptides are summarized in Table I. The amino acid sequence of A1 conserved in residues 33–38 was confirmed by ESI-MS. The sequences of some peptides were identified by a combination of ESI-MS/MS and the amino acid composition. The amino acid sequences of A2 and A3 were also identified by ESI-MS/MS (data not shown). The amino acid corresponding to the  $Y_1^+$  ion (13, 14) on A2 was considered to be either Gln or Lys, but could be identified as Lys in view of the enzyme specificity. As the same reason the difference of 113 in amino acid residue weight on A3 indicated the presence of either Ile or Leu. This was later confirmed to be Leu by amino acid analysis. Thus, the amino acid sequence could not be determined by MS analysis alone since the molecular weight of Ile is equal to that of Leu, and that of Gln is equal to that of Lys.

Fractions A4 and A5 included either Gln or Asn because

the difference between the molecular weights calculated from the amino acid composition and ESI-MS was only 1 mass unit. This finding indicated that Gln or Asn was converted to Glu or Asp.

Previous studies have shown the possible modification of the N-terminal residue of PA4.78 (9). In peak A6, the difference between the molecular weight calculated from the amino acid composition and that derived from ESI-MS was 39.8. The difference number ( $m/z$  39.8) indicated that A6 was acetylated at the N-terminal ( $+m/z$  42) and contained two amide residues ( $-m/z$  1)  $\times$  2 (Gln and Asn).

In long peptides, such as A4–A6, ESI-MS/MS data were assigned with the Finnigan database Pep Match after predicting the amino acid sequence of the peptides by amino acid composition and ESI-MS. Since it was difficult to determine a difference of 1 mass unit, like the difference between Asp and Asn, with ESI-Qpole-MS due to the limitation of the low resolution mass spectrometry, localization of Asp and Asn in A4 was determined by peptide sequencing. Since A9 appeared on HPLC at 254 nm (Fig. 2), a pyri-

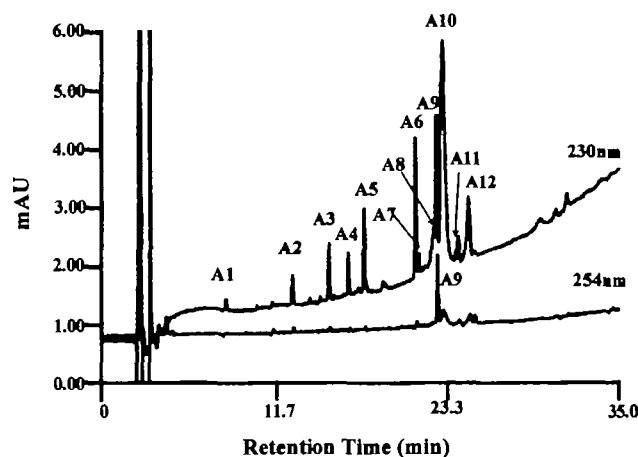


Fig 2 HPLC profile of pyridylethylated PA4.78 digested with arginylendopeptidase. HPLC was performed using a linear gradient of solvent B (10–100%, 45 min), and peptides were detected at 230 and 254 nm

TABLE I. Arginylendopeptidase digestion of pyridylethylated PA4.78.

Peak	Mass <sup>a</sup>	MS-MS	MW <sup>b</sup>	Amino acid composition	MW-mass	Peptide sequence	Residue
A1	545	VGLAGK					33–38
A2	671	TFFQK					28–32
A3	1,208	ALSSAETAAFLK	1,208.4	T(1)S(2)E(1)A(4)L(2)F(1)K(1)	0.4		76–87
A4	1,877	SANDAKKVFEILDRDR	1,878.3	D(4)S(1)E(1)A(2)V(1)I(1)L(1)F(1)-K(2)R(2)	1.3	SANDAKKVEE-ILDRDR	39–54
A5	2,108	AGDSDGDGKIGVEEFQALVKA	2,107.2	D(3)S(1)E(3)G(4)A(3)V(2)I(1)L(1)-F(1)K(2)	1.2		88–108
A6	2,985	AcSITDIVSVSQKDIEAALESV-KAEGSFNYK	2,941.2	D(3)T(1)S(4)E(4)G(1)A(3)V(2)I(3)-L(1)Y(1)F(1)K(3)	-39.8		Ac1–27
A7	2,694						
A8	3,637, 4,010						Ac1–32, 39–71
A9	2,151					SGFIEEEEL(PE)-CLFLQN(FK)	55–71
A10	6,022						Ac1–54
A11	8,583						Ac1–75
A12	11,863						PE-PA4.78

Amino acids in parentheses (peptide sequence) indicate non-detected amino acids. PE (peptide sequence): pyridylethylated amino acid. <sup>a</sup>Molecular mass determined by deconvolution of ESI mass spectrum. <sup>b</sup>M.W. was calculated from amino acid composition.

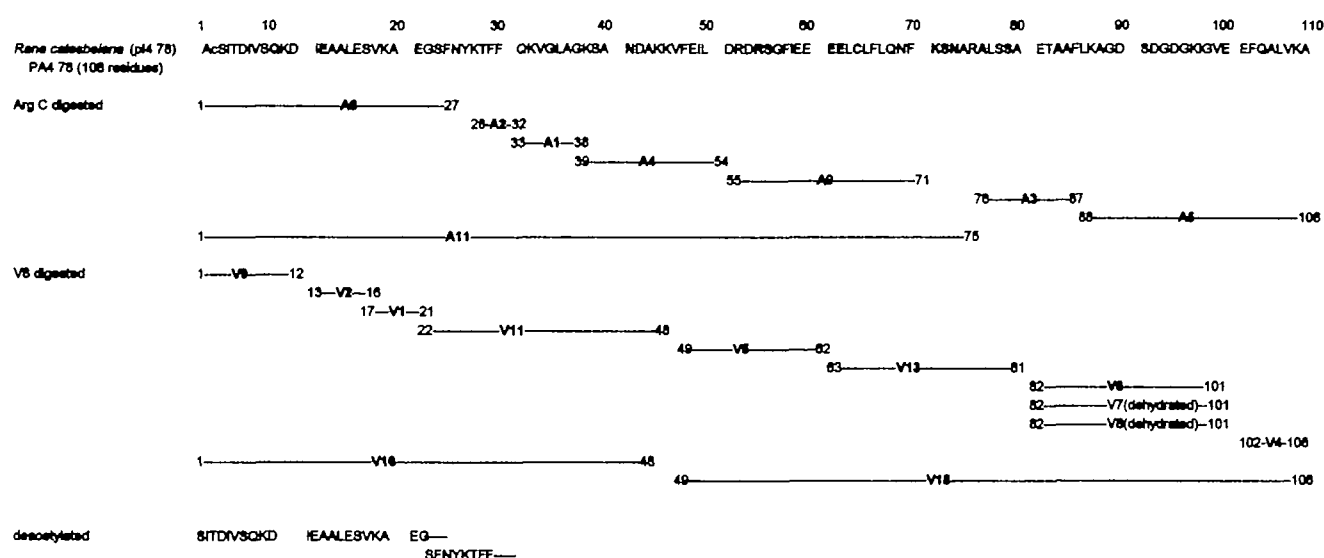


Fig. 3 Amino acid sequence of  $\beta$ -type parvalbumin in *Rana catesbeiana* PA4.78. Bold letters indicate amino acid residues that are different from those of *R. esculenta* PA4.50.

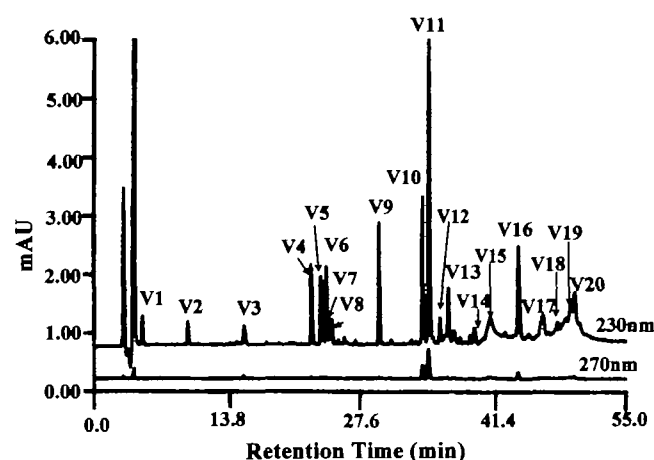


Fig. 4 HPLC profile of PA-1 digested by  $V_8$  protease. HPLC was performed using a linear gradient of solvent B (10–100%, 90 min), and peptides were detected at 230 and 270 nm.

dylethylated Cys residue of this peptide was confirmed by peptide sequencing.

Thus, we determined the major part of the amino acid sequence of PA4.78 with Arg-C fragments. However, fragments of residues 72–75 (molecular weight 403) could not be found on the chromatogram (Fig. 3).

**Digestion of PA4.78 with  $V_8$  Protease**—To identify the amino acid sequence between residues 72 and 75 and to confirm the other sequences of PA4.78 examined by Arg-C, native PA4.78 was digested with  $V_8$  protease. The HPLC profile after digestion by  $V_8$  protease is shown in Fig. 4. Twenty peptides were identified, corresponding to peaks V1 to V20. The molecular weight, amino acid composition, and peptide sequence of V8-digested peptides are summarized in Table II.

V13 peptide consisted of residues 63–81 and therefore

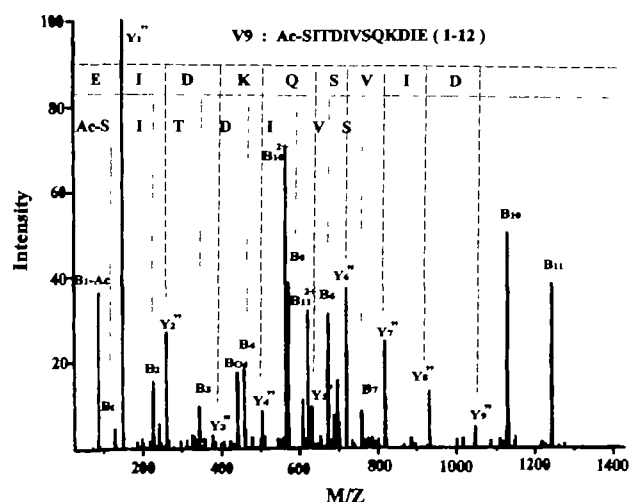


Fig. 5 ESI-MS/MS spectrum of peptide V9 (N-terminal peptide) digested by  $V_8$  protease.

included residues 72–75 (–SNAR–). Its sequence was determined by peptide sequencing. V7 and V8 contained a sequence comprising an aspartylglycine bond and a dehydrated Asp-94 or 92 (succinylation), respectively. The molecular weight of V12 was equal to that of fragment V13+32, which corresponded to the minor component, M+32 (11,789), seen in the ESI-MS spectrum (Fig. 1), confirming the purity before enzymatic digestion. This fragment is under investigation. Acetylation of the N-terminus of V9 was also confirmed by the ESI-MS/MS spectrum (Fig. 5).  $B_1$  ( $m/z$ =129) was found to be acetylated Ser, and a sequence of 12 residues was subsequently deduced using both the B ion and  $Y''$  ion (13, 14). The N-terminal amino acid sequence of the deacetylated PA4.78 was also confirmed by peptide sequencing.



Thus, the complete amino acid sequence of  $\beta$ -type parvalbumin of *R. catesbeiana* PA4.78 could be determined from the results of Arg-C and  $V_8$  protease digestion (Fig. 3).

**Comparison of Amino Acid Sequences of  $\beta$ -Type Parvalbumin of *R. catesbeiana* PA4.78 and Other Parvalbumins**—The sequence homology between *R. catesbeiana* PA4.78 and *R. catesbeiana*  $\alpha$ -type PA4.97 (9), *R. esculenta*  $\alpha$ -type PA4.88 (6), and *R. esculenta*  $\beta$ -type PA4.50 (7) was 53, 51, and 86%, respectively (Fig. 6). These results suggest that PA4.78 belongs to the  $\beta$ -type. This was confirmed by the phylogenetic tree analysis shown in Fig. 7. The phylogenetic tree was constructed based on multiple alignments of parvalbumins (GENETIX Version 3, Software Development

USA). The above results matched well the taxonomy, which clearly showed that *R. catesbeiana* PA4.78 is a member of the  $\beta$  family (Fig. 7). The molecular weight of *R. catesbeiana* PA4.78 calculated from electrophoretic mobility was 10,100 (8), but this value did not match the actual sequence of the protein in this study, suggesting that the electrophoretic behavior of *R. catesbeiana* PA4.78 is peculiar. This is in contrast to the molecular weight of *R. catesbeiana* PA4.97 estimated by electrophoresis, which was in agreement with that estimated from the amino acid sequences (9). *R. catesbeiana* PA4.78 is a common PA in amino acid sequence, and the difference between the two frog  $\beta$  parvalbumins is small compared to the difference between the  $\alpha$

TABLE II.  $V_8$  protease digestion of PA4.78.

Peak	Mass	MS-MS	Amino acid composition	Peptide sequence	Residue
V1	532	SVKAE			17–21
V2	402	AALE			13–16
V3	—	—			—
V4	776	FQALVKA	F(1)A(2)V(1)L(1)K(1)	FQA(LVKA)	102–108
V5	1,449	ILDRDRSGFIEE			49–60
	1,708	ILDRDRSGFIEEE			49–62
V6	1,980	TAAFLKAGDSGDGKIGVEE			82–101
V7	1,962	—		TAAFLKAGDSGDG(-H <sub>2</sub> O)GKI-GVEE	82–101 dehydrated
V8	1,962	—		TAAFLKAGDSGDG(-H <sub>2</sub> O)GKI-GVEE	82–101 dehydrated
V9	1,389	AcSITDIVSQKDIE			Ac1–12
V10	3,497	—			17–48 (V1+V11)
V11	2,982	—		GSFNYKTFFQKVGLAGKSAN-DAKK(VFE)	22–48
V12	2,144	—		LCLFLQNFKSNARALSSAE+32	63–81 added 32
V13	2,112	—		LCLFLQNFKSNARALSSAE	63–81
V14	4,223, 4,481	—		—	(63–81)×2, —
V15	6,185, 5,082	—		—	—
V16	5,253	—		—	Ac1–48 (V9+V2+V1+V11)
V17	5,746, 5,765	—		—	49–101 dehydrated, 49–101
V18	6,522	—		—	49–108 (V5+V13+V6+V4)
V19	9,037, 11,034	—		—	Ac1–81 (V16+V5+V3)
V20	11,000	—		—	Ac1–101

Numbers in parentheses (amino acid composition) indicate the number of amino acids. Amino acids in parentheses (peptide sequence) indicate non-detected amino acids.

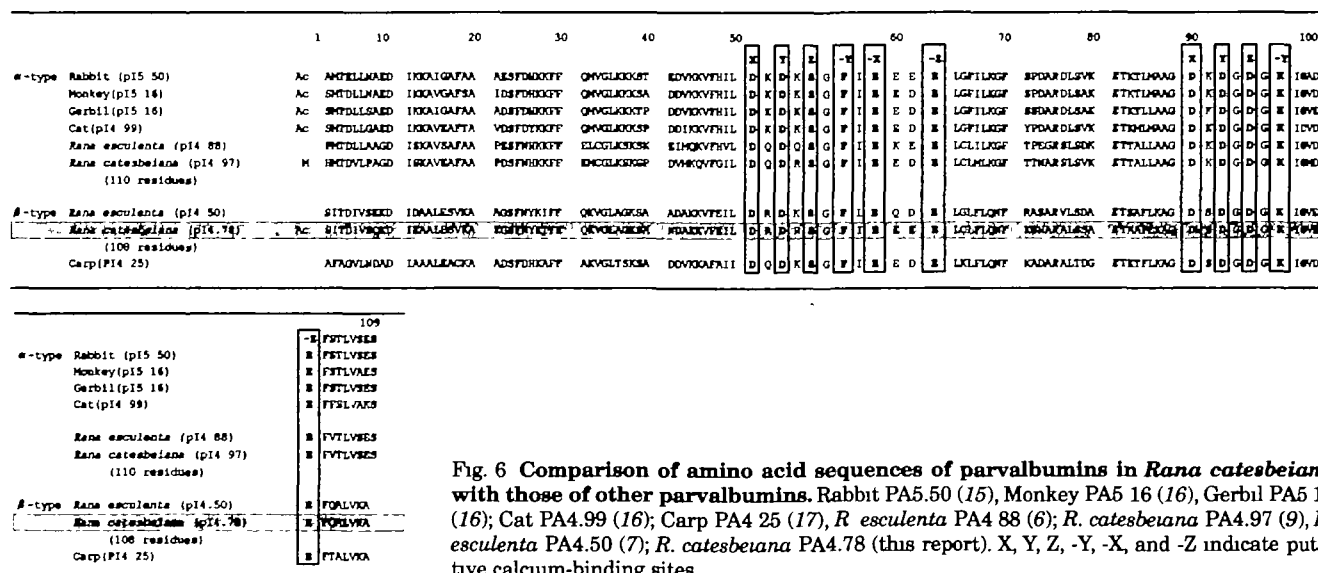


Fig. 6 Comparison of amino acid sequences of parvalbumins in *Rana catesbeiana* with those of other parvalbumins. Rabbit PA5.50 (15), Monkey PA5.16 (16), Gerbil PA5.16 (16); Cat PA4.99 (16); Carp PA4.25 (17), *R. esculenta* PA4.88 (6); *R. catesbeiana* PA4.97 (9), *R. esculenta* PA4.50 (7); *R. catesbeiana* PA4.78 (this report). X, Y, Z, -Y, -X, and -Z indicate putative calcium-binding sites.

parvalbumins (Fig. 7).

The amino acid displacement in *R. catesbeiana*  $\beta$ -type parvalbumin PA4.78 compared with *R. esculenta* PA4.50 was 15 out of 108 residues (14% displacement) (Fig. 6). *R. catesbeiana*  $\beta$ -type PA4.78 contained Cys at residue 64 and was acetylated at the amino terminus. The 25 residues of the carboxyl terminus were completely conserved, but many amino acid displacements were found between residues 51 and 80 (30% displacement), although the functionally important sequence of PA was completely conserved in this region. The amino acid residues of putative calcium-binding sites were X:(Asp-51), Y:(Asp-53), Z:(Ser-55), -Y:(Phe-57), -X:(Glu-59), -Z:(Glu-62), X:(Asp-90), Y:(Asp-92),

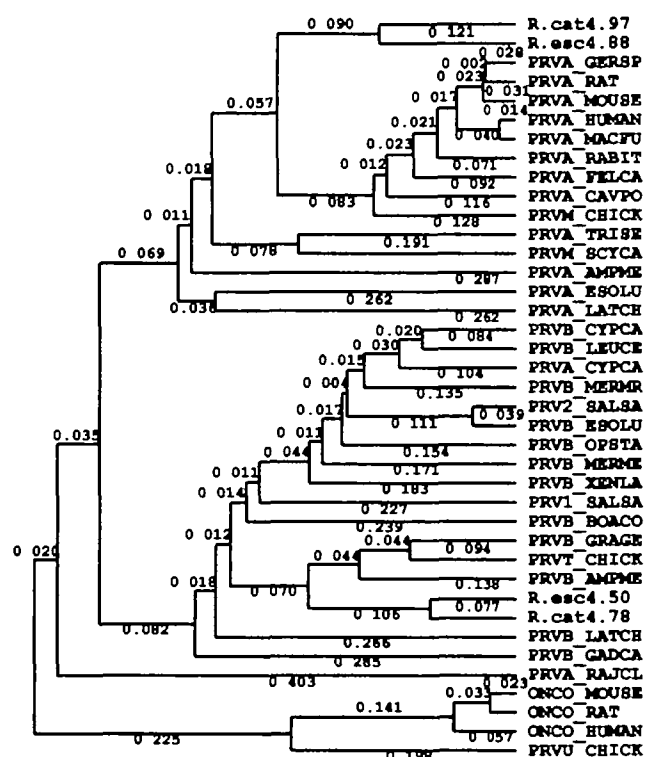


Fig. 7. Phylogenetic tree of parvalbumins. Branch lengths and values indicated show the relative numbers of mutations per site. *R.cat*PA1, bullfrog PA-1 4.78, *R.cat*PA2, bullfrog PA-2 4.97, *R.esc*4.50, edible frog PA4.50, and *R.esc*4.88, edible frog PA4.88. Other parvalbumins are designated as ID numbers of SWISS-PROT: PRVA\_GERSP, gerbil PA $\alpha$ ; PRVA\_RAT, rat PA $\alpha$ ; PRVA\_MOUSE, mouse PA $\alpha$ ; PRVA\_HUMAN, human PA $\alpha$ ; PRVA\_MACFU, Japanese macaque PA $\alpha$ ; PRVA\_RABIT, rabbit PA $\alpha$ ; PRVA\_FELCA, cat PA $\alpha$ ; PRVA\_CAVPO, guinea pig PA $\alpha$ ; PRVM\_CHICK, chicken PA from muscle; PRVA\_TRISE, leopard shark PA $\alpha$ ; PRVM\_SCYCA, spotted dogfish PA $\alpha$ ; PRVA\_AMPME, salamander PA $\alpha$ ; PRVA\_ESOLU, northern pike PA $\alpha$ ; PRVA\_LATCE, coelacanth PA $\alpha$ ; PRVB\_CYPCA, carp PA $\beta$ ; PRVB\_LEUCE, chub PA $\beta$ ; PRVA\_CYPCA, carp PA $\alpha$ ; PRVB\_MERMR, whiting PA $\beta$ ; PRV2\_SALSA, Atlantic salmon PA $\beta$ ; PRVB\_ESOLU, northern pike PA $\beta$ ; PRVB\_OPSTA, oyster toadfish PA $\beta$ ; PRVB\_MERME, European hake PA $\beta$ ; PRVB\_XENLA, African clawed frog PA $\beta$ ; PRV1\_SALSA, Atlantic salmon PA $\beta$ ; PRVB\_BOACO, boa PA $\beta$ ; PRVB\_GRAGE, map turtle PA $\beta$ ; PRVT\_CHICK, chicken PA (thymic hormone), PRVB\_AMPME, salamander PA $\beta$ ; PRVB\_LATCE, coelacanth PA $\beta$ ; PRVB\_GADCA, Baltic cod PA $\beta$ ; PRVA\_RAJCL, thornback ray PA $\alpha$ ; ONCO\_MOUSE, mouse oncomodulin (PA $\beta$ ); ONCO\_RAT, rat oncomodulin (PA $\beta$ ); ONCO\_HUMAN, human oncomodulin (PA $\beta$ ); PRVU\_CHICK, chicken PA from thymus.

Z:(Asp-94), -Y:(Lys-96), and -Z:(Glu-101). These were conserved in all  $\alpha$  and  $\beta$  types of *R. catesbeiana* PA and other parvalbumins. In addition, Arg-75 and Glu-81, which are thought to form a salt bridge located in the interior of the molecule (18), were also conserved in PA4.78 (Fig. 6).

The results of peptide sequencing after deacetylation showed that Ser was present at the N-terminus, and more than 20 residues could be identified from the N-terminus (Fig. 3).

**Conclusion**—We determined the primary structure of *R. catesbeiana*  $\beta$ -type parvalbumin PA4.78 by ESI-MS, amino acid analysis, and peptide sequencing. However, it was impossible to determine the primary structure by using only one method, such as mass spectrometry. Similarly, although peptide sequencing allows the determination of the exact amino acid sequence of small peptides, in many cases the last several residues (for example, in A9, V4, and V11) cannot be determined by Edman degradation. The post-translational modification of the protein cannot be established. Most important is the inability to sequence the modified N-terminal residues of the protein.

However, the combination of ESI-MS with amino acid analysis allowed the establishment of the structure of the N-terminus and/or any post-translational modification by comparing the difference in molecular weight. When only MS is used, it is often difficult to determine the sequence because some amino acids have the same molecular weight, such as Ile and Leu, or Glu and Lys, or differ by only 1 mass unit, for example, Asp and Asn, or Glu and Gln. Thus, ESI-MS, amino acid analysis, and peptide sequencing each offers certain advantages and disadvantages, and the use of only one of these methods does not allow proper determination of the sequence. For determination of the primary structure, therefore, these techniques should be combined to compensate for their disadvantages. Recent technological advances allow effective microanalysis using matrix-assisted laser desorption ionization mass spectrometry (19, 20) and aminoquinolyl-carbamyl amino acid analysis (21, 22).

## REFERENCES

- Wnuk, W., Cox, J.A., and Stein, E.A. (1982) Parvalbumins and other soluble high-affinity calcium-binding proteins from muscle in *Calcium and Cell Function* (Cheung, W.Y., ed.) Vol. II, pp 243–278, Academic Press, New York
- Gillis, J.M., Thomason, D., Lefevre, J., and Krestinger, R. H. (1982) Parvalbumins and muscle relaxation: a computer simulation study. *J. Muscle Res. Cell Motil.* **3**, 377–398
- Heizmann, C.W., Berchtold, M.W., and Rowleson, A.M. (1982) Correlation of parvalbumin concentration with relaxation speed in mammalian muscles. *Proc. Natl. Acad. Sci. USA* **79**, 7243–7247
- Tanokura, M. and Yamada, K. (1985) A calorimetric study of  $Ca^{2+}$  binding to two major isoforms of bullfrog parvalbumin. *FEBS Lett.* **185**, 165–169
- Goodman, M. and Pechère, J.-F. (1977) The evolution of muscular parvalbumins investigated by the maximum parsimony method. *J. Mol. Evol.* **9**, 131–158
- Jauregui-Adell, J., Pechère, J.-F., Briand, G., Richet, C., and Demaille, J.G. (1982) Amino-acid sequence of an  $\alpha$ -parvalbumin, pI = 4.88, from frog skeletal muscle. *Eur. J. Biochem.* **123**, 337–345
- Capony, J.-P., Demaille, J., Pina, C., and Pechère, J.-F. (1975) The amino-acid sequence of the most acidic major parvalbumin from frog muscle. *Eur. J. Biochem.* **58**, 215–227
- Tanokura, M., Aramaki, H., Goto, K., Hashimoto, U., Toyomori,

- Y., and Yamada, K. (1986) Preparation and characterization of two major isotypes of parvalbumins from skeletal muscle of bullfrog (*Rana catesbeiana*). *J. Biochem.* **99**, 1211–1218
9. Sasaki, T., Tanokura, M., and Asaoka, K. (1990) The complete amino acid sequence of bullfrog (*Rana catesbeiana*) parvalbumin pI4.97. *FEBS Lett.* **268**, 249–251
10. Yazawa, M., Sakuma, M., and Yagi, K. (1980) Calmodulins from muscles of marine invertebrates, scallop and sea anemone. *J. Biochem.* **87**, 1313–1320
11. Pechère, J.-F., Demaille, J., and Capony, J.-P. (1971) Muscular parvalbumins: preparative and analytical methods of general applicability. *Biochim. Biophys. Acta* **236**, 391–408
12. Tanokura, M., Goto, K., Toyomori, Y., and Yamada, K. (1987) Preparation and characterization of the major isotype of parvalbumin from skeletal muscle of the toad (*Bufo bufo japonicus*). *J. Biochem.* **102**, 1133–1139
13. Roepstorff, P. and Fohlman, J. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **11**, 601
14. Biemann, K. and Martin, S. (1987) Mass spectrometric determination of amino acid sequence of peptides and proteins. *Mass Spectrom. Rev.* **6**, 1–76
15. Capony, J.-P., Pina, C., and Pechère, J.-F. (1976) Parvalbumin from rabbit muscle: Isolation and primary structure. *Eur. J. Biochem.* **70**, 123–135
16. Hauer, C.R., Staudenmann, W., Kuster, T., Neuheiser, F., Hughes, G.J., Seto-Ohshima, A., Tanokura, M., and Heizmann, C.W. (1992) Protein sequence determination by ESI-MS and LSI-MS tandem mass spectrometry: parvalbumin primary structures from cat, gerbil and monkey skeletal muscle. *Biochim. Biophys. Acta* **1160**, 1–7
17. Coffee, C.J. and Bradshaw, R.A. (1973) Carp muscle calcium-binding protein I. Characterization of the tryptic peptides and the complete amino acid sequence of component B. *J. Biol. Chem.* **248**, 3302–3312
18. Coffee, C.J. and Solano, C. (1976) Preparation and properties of carp muscle parvalbumin fragments A (residues 1 leads to 75) and B (residues 76 leads to 108). *Biochim. Biophys. Acta* **453**, 67–80
19. Scheler, C., Lamer, S., Pan, Z., Li, X.P., Salmikow, J., and Jungblut, P. (1998) Peptide mass fingerprint sequence coverage from differently stained proteins on two-dimensional electrophoresis patterns by matrix assisted laser desorption/ionization-mass spectrometry (MALDI-MS). *Electrophoresis* **19**, 918–927
20. Scheler, C., Müller, E.C., Stahl, J., Müller-Nerdam, U., Salmikow, J., and Jungblut, P. (1997) Identification and characterization of heat shock protein 27 protein species in human myocardial two-dimensional electrophoresis patterns. *Electrophoresis* **18**, 2823–2831
21. Shindo, N., Nojima, S., Fujimura, T., Taka, H., Mineki, R., and Murayama, K. (1997) Separation of 18 6-aminoquinolyl-carbamyl-amino acids by ion-pair chromatography. *Anal. Biochem.* **249**, 79–82
22. Shindo, N., Fujimura, T., Nojima-Kazuno, S., Mineki, R., Furukawa, S., Sasaki, K., and Murayama, K. (1998) Identification of multidrug resistant protein 1 of mouse leukemia P388 cells on a PVDF membrane using 6-aminoquinolyl-carbamyl (AQC)-amino acid analysis and World Wide Web (WWW)-accessible tools. *Anal. Biochem.* **264**, 251–258