

RESEARCH ARTICLE

Identification of commercial hake and grenadier species by proteomic analysis of the parvalbumin fraction

Mónica Carrera¹, Benito Cañas², Carmen Piñeiro¹, Jesús Vázquez³ and José Manuel Gallardo¹

¹ Instituto de Investigaciones Marinas, Vigo, Pontevedra, Spain

² Dpto. Química Analítica, Facultad de Química, UCM, Madrid, Spain

³ Centro de Biología Molecular Severo Ochoa, Laboratorio de Química de Proteínas, Madrid, Spain

Analysis of parvalbumin fractions through proteomic methodologies allowed the differential classification of ten commercial, closely related species of the family Merlucciidae. Muscle extracts from nine hake species of the genus *Merluccius* including two subspecies of *Merluccius australis* (*australis* and *polylepsis*) and one grenadier species *Macruronus novaezelandiae* with two populations (*novaezelandiae* and *magellanicus*) were evaluated by 2-DE and MALDI-TOF MS. 2-DE demonstrated that the species tested displayed a low intra-specific degree of polymorphism and the isoform patterns were noticeably species-specific. MALDI-TOF mass fingerprints showed clear differences in the pattern of peptides produced by tryptic digestion between the *Merluccius* and the *Macruronus*, making the genus differentiation possible. In addition, a selective peptide mass present in the spectra from certain hakes allowed its classification in two groups: Euro-African and American hakes. Besides, some specific masses allowed a clearly individual identification for *M. bilinearis*, *M. australis polylepsis*, *M. australis australis*, *M. productus*, *M. paradoxus* and *M. polli*, while the rest of the hake species can be grouped in two clusters, comprising *M. hubbsi* and *M. gayi* in one and *M. merluccius* and *M. capensis* in the other.

Received: December 14, 2005

Revised: May 16, 2006

Accepted: June 4, 2006



Keywords:

Fish identification / MALDI-TOF MS / Merlucciidae / Parvalbumin

1 Introduction

In some countries, the family Merlucciidae constitutes an important percentage of the overall fish consumption. This family comprises four different genera: *Lyconodes*, *Lyconus*, *Merluccius* and *Macruronus* [1] but only the last two are of interest to the fishing industry. From a qualitative point of view, some members of this family are more appreciated than others for consumption, due to their different sensory features, giving rise to different market values. Among them, the specimens belonging to the genus *Merluccius*, commonly known as hakes, have higher commercial value, being pre-

ferred over the species belonging to the genus *Macruronus*, known as grenadiers. Within this genus, only *Ma. novaezelandiae* deserves special fishing interest.

Conventional identification of unprocessed fish is done by the examination of their morphological characteristics. However, due to the development of the fishing industry, marine products can be processed, either on board or after landing, making often impossible the appreciation of their external anatomical or morphological features. To avoid cases of substitution of certain fish species by others with less commercial value, a number of global regulations have been implemented [2] to assure a complete and correct information, guaranteeing market transparency. As a consequence, these regulations make highly recommendable the development of the necessary analytical tools to make possible distinguishing between closely related species. DNA-based [3–5] and protein identification methodologies [6–8] have been used with authentication purposes. Recently, the proteomics tools have been applied in the fishing indus-

Correspondence: Mónica Carrera, Instituto de Investigaciones Marinas, Eduardo Cabello 6 E-36208 Vigo, Pontevedra, Spain

E-mail: mcarrera@iim.csic.es

Fax: +34-986-292-762

Abbreviation: EF-hand, two helix-loop-helix

try with several purposes, among others, the characterization of post-mortem changes [9, 10], the study of the effect of some additives during the processing of fish muscle [11] and the identification of species in seafood products [12–14]. This asseveration can be exemplified with the use of 2-DE for the authentication of several species belonging to the family Merlucciidae, which has allowed the designation of several specific proteins for five closely related species from the genus *Merluccius*, [15, 16]. By contrast, little effort has been made to elucidate differences among closely related fish species using MS [17].

Parvalbumins are proteins relatively abundant in the sarcoplasmic fraction of the white fish muscle. They are thermo stable proteins with a molecular weight around 10–12 kDa and an *pI* in the range of 4.0–5.2. Two helix-loop-helix (EF-hand) motifs [18–20], which are important for calcium-binding and are involved in Ca^{2+} and Mg^{2+} transport [21], are included in the parvalbumin structure. It is assumed that these proteins are important for the relaxation/contraction cycle of the muscle fibers by soaking up free calcium [21–24]. Besides biochemical connotations, parvalbumins also play an important role in food safety, as they have been defined as major cross reactive allergens (induce Type-I allergy) [25–28].

From the evolutionary point of view, there are two distinct parvalbumin phylogenetic lineages, named α and β . In the case of fish, most of the parvalbumin sequences belong to the β type [22, 29]. In the special case of family Merlucciidae there are only two entries in the protein databases: β parvalbumin from European hake (*Merluccius merluccius*) [30] and β parvalbumin from Silver hake (*Merluccius bilinearis*) [31], with accession numbers PRVB_MERME, P02620 and PRVB_MERBI, P56503, respectively.

Due to parvalbumin high inter-specific variability, thermo stability [32–35] and high concentration in the white fast-contracting muscles from most fishes [36], we can forecast that they could be used as a good authenticity biomarker in fresh or processed seafood products [17, 37]. In this paper we have applied for the first time the proteomics methodology to achieve the differential characterization of all the parvalbumin isoforms from ten commercial, closely-related species, belonging to the family Merlucciidae. This work includes the analysis of 36 different parvalbumin isoforms, belonging to two different genera. From the genus *Merluccius*: *M. merluccius*, *M. capensis*, *M. polli*, *M. paradoxus*, *M. hubbsi*, *M. gayi*, *M. productus*, *M. bilinearis* and two different sub-species of *M. australis*, *M. australis polylepsis* and *M. australis australis* and from the genus *Macruronus*: two different populations of *Ma. novaezelandiae*, *Ma. nov. nov.* and *Ma. nov. magellanicus*. For this purpose we have applied 2-DE coupled with MALDI-TOF MS analysis. This study is the first step towards the total sequencing of these interesting proteins. Results obtained will allow the characterization of all the different parvalbumin isoforms and the identification of species-specific peptides for each of the

species included in the family Merlucciidae. This approach would be usable by fisheries as a good tool for fresh or processed seafood products authentication, providing a rapid and effective identification and guaranteeing the quality of the product to consumers.

2 Materials and methods

2.1 Fish material

The hake species evaluated in this study included: *Merluccius merluccius* (European hake), *Merluccius capensis* (Cape hake), *Merluccius polli* (Benguela hake), *Merluccius paradoxus* (Deep-water Cape hake), *Merluccius hubbsi* (Patagonian hake), *Merluccius gayi* (Peruvian hake), *Merluccius australis* (Southern hake) from two different locations (from New Zealand, *M. australis australis* and from Patagonian coasts, *M. australis polylepsis*), *Merluccius productus* (Pacific hake) and *Merluccius bilinearis* (Silver hake). Specimens were caught off the coast of Spain (*M. merluccius*), South and Northwest Africa (*M. capensis*, *M. polli* and *M. paradoxus*), North America (*M. bilinearis* and *M. productus*) or South America (*M. australis*, *M. gayi*, and *M. hubbsi*). Except for the European hake, fish were frozen on board at -30°C , and shipped by plane to our laboratory for the analyses. Special care was taken in keeping their morphological characters in good shape.

The two grenadier populations of *Macruronus novaezelandiae* included: *Macruronus novaezelandiae novaezelandiae* (Blue grenadier) and *Macruronus novaezelandiae magellanicus* (Patagonian grenadier). These were caught off the coast of the New Zealand and Argentina. These fish were frozen and shipped as per the hake species.

The weight of the specimens studied was in the range of 3 to 6 kg. At least three individuals belonging to each of the ten different species were subjected to taxonomical study and analyzed, according to their anatomical and morphological features, by an expert marine biologist and by genetic identification with the fishID Kit (Bionostra SL., Madrid, Spain). Muscle of these fish was then subjected to 2-DE and MS analysis.

2.2 Extraction of sarcoplasmic proteins

Extraction of the sarcoplasmic proteins was carried out by homogenizing 5 g of raw white muscle from each species in two volumes of 10 mM Tris-HCl buffer, pH 7.2, supplemented with 5 mM PMSF, in an Ultra-Turrax blender for 30 s. The fish extracts were then centrifuged at 40 000g for 20 min at 4°C (J221-M centrifuge; Beckman, Palo Alto, CA). The supernatants were filtered using Ultrafree CL (0.22 μm) filters (Millipore, Bedford, MA) and maintained at -80°C until the electrophoretic analysis. Protein concentration in the extracts was determined by the bicinchoninic acid method (Sigma Chemical Co., USA).

2.3 2-DE Analysis

First dimension native IEF was performed at 10°C in a Multiphor II electrophoresis unit (Amersham Biosciences, Sweden) as described elsewhere [38]. Protein extracts (40 mg/mL) were loaded on gels in duplicate using sample applicator paper. IEF conditions were: 1500 V, 50 mA, 30 W until at least 4000 V·h were reached. A mixture of proteins standard in the 2.5–6.5 pH range (low pI standard from Amersham Biosciences) including human carbonic anhydrase B (pI: 6.55), bovine carbonic anhydrase B (pI: 5.85), β -lactoglobulin A (pI: 5.20), soybean trypsin inhibitor (pI: 4.55), glucose oxidase (pI: 4.15), methyl red (pI: 3.75), amyloglucosidase (pI: 3.5) and pepsinogen (pI: 2.8), was used. IEF strips corresponding to individual lanes were cut immediately after the run was completed and kept frozen at –80°C. A duplicate IEF strip was stained with 0.1% CBB (Sigma) according to the Amersham Biosciences staining protocol. Equilibration of pH 4–6.5 IEF gel strips was carried out at room temperature as described previously [16] and the second dimension was run using gradient (12–14%) SDS-PAGE precast polyacrylamide 245 × 180 × 0.5 mm gels (ExcelGel XL SDS, Amersham Biosciences) in Tris-glycine buffer [17]. A low molecular weight protein standard (14–67 kDa) was employed as reference. Gels were stained with silver using a protocol from Amersham Biosciences and the image analysis was carried out using the PDQuest software from Bio-Rad (Hercules, CA).

Preparative 2-DE gels for MS analysis, were performed using vertical SDS-PAGE (10%T and 3%C) gels with Tris-tricine buffer and stained with CBB (Amersham Bioscience) according to the manufacturer's instructions. PDQuest software was used for image analyses. These gels were specially prepared for further analysis by MS. Preparative 2-DE gels presented a low percent of acrylamide (10%) to allow trypsin to penetrate more easily inside the gel. A different electrophoresis buffer (Tris-tricine) was used to improve the resolution of proteins with low molecular weight. Gels were stained with CBB, which is more compatible with sample preparation for MS analysis.

2.4 Peptide mapping by MALDI-TOF MS

Parvalbumin spots were cut off from the preparative gels and subjected to *in-gel* digestion with trypsin (Promega, Madison, WI, USA). Triplicate spectra for every sample were taken between 400 and 4000 Da. Peptide masses obtained of each spot were compared with the theoretical peptide masses from all the parvalbumins accessible in the protein databases. The parameters used for the search were as follows: monoisotopic masses, two allowed missed cleavage sites for trypsin cutting and three variable modifications: acetyl N-terminal, carboxy-methylated cysteine and methionine sulfoxide. Spots were excised from the gel so as to maximize the protein-to-gel ratio: usually the most intensely stained region at the centre of the spot was used to

avoid extracting an excess of the gel matrix. Enzymatic digestion was performed overnight at 37°C according to Jensen *et al.* [39].

MALDI-TOF MS analysis was performed using an Autoflex instrument from Bruker (Bruker Daltonics, Billerica, MA, USA), equipped with delayed extraction and operating in the reflector mode. After digestion, peptide solution was acidified by the addition of TFA and dried down. Anchor-Chip plates (Bruker Daltonics) were used as MALDI targets. Matrix, 0.5 μ L of saturated DHB (Bruker Daltonics) in 30% ACN containing 0.1% TFA, was first applied to the target and allowed to dry. Aliquots (0.5 μ L) of the extracted peptides were applied later to the target and allowed to dry. External calibration was carried out using a set of synthetic peptides (Bruker Daltonics) added to the target in positions close to the samples. MASCOT (Matrix Sciences, London, UK) was used for database search (<http://www.matrixscience.com>).

3 Results and discussion

3.1 2-DE analyses of parvalbumins

Sarcoplasmic protein extracts were subjected to 2-DE using gradient gels (12–14%) which were silver stained. To ensure reproducibility, three individuals from each of the species (36 specimens) were analyzed, running gels in triplicate for each individual. Thus a total of 108 2-DE gels were processed and analyzed with the image software. In a preliminary survey, gels with a broad pH range (3–10) were used. Later, and to avoid overlapping of acidic parvalbumins, strips with a narrow pH range (pH 4–6.5) were chosen. These strips allowed a more uniform distribution, clear and well-defined, of parvalbumins in the gels improving reproducibility. Representative clusters of spots, corresponding to the parvalbumin fraction for the different hake species and both sub-species from grenadier are shown in Fig. S-1 in the Supplementary Material. The assignment and distribution of the electrophoretic spots is described for the different fish species.

Qualitative inter-specific differences were noticeable, as previously reported from five hake species [17], and high variations in pI, M_r and number of protein spots or isoforms can be observed. These are compiled in Table S-1 in the Supplementary Material. The species studied had at least two (*M. bilinearis*) major parvalbumin isoforms, but frequently three or four isoforms (*M. australis australis*) were present. Parvalbumin spots have been numbered with Arabic characters (P1–P36) from right to left according to the pI. This pattern of isoforms was consistently and reproducibly obtained when the parvalbumin fraction from different individuals of the same species was analyzed, suggesting that it could be used to differentiate between the different species. The differences in pI and M_r among these spots indicate the presence of amino acid substitutions in the parvalbumins sequences. Nevertheless, 2-DE patterns alone cannot reveal whether parvalbumins with the same pI and

M., are truly identical or not. For this reason, a further characterization of each one of the parvalbumin isoforms was performed by MS.

3.2 MALDI-TOF MS analysis

After their separation in preparative 2-DE gels, all the spots corresponding to the parvalbumin isoforms were cut off, digested *in-gel* with trypsin and the digests analyzed by MALDI-TOF MS. Unfortunately, little information is available in the protein databases about the parvalbumins from the family Merlucciidae. In fact, although 40 parvalbumin sequences from Teleostei are actually present in the Swiss-Prot and TrEMBL databases, only two belong to the family Merlucciidae (PRVB_MERME, P02620 and PRVB_MERBI, P56503).

The results obtained by the analysis of the MALDI-TOF mass spectra are summarized in Table S-2 in the Supplementary Material and discussed there. It shows that in all the species except *M. bilinearis*, the parvalbumin isoforms with the highest *pI* (4.32–4.57) presented some degree of homology with PRVB_MERME; we have consequently named them as Parvalbumins A (PRVB_A). Another group of isoforms were found to display homology with PRVB_MERBI, and they were named Parvalbumins B (PRVB_B). Percent coverage is higher in the PRVB_A group as compared with the homology among the components of the PRVB_B group, indicating the existence of higher variability in the sequences of PRVB_B. This makes the latter group more suitable for the identification and authentication of members from the family Merlucciidae. Finally, we identified a third group of isoforms, named Parvalbumins C (PRVB_C), where no homologies with the described Merlucciidae parvalbumins can be found, but showing some degree of homology with the *Alaska pollock* β -parvalbumin sequence PRVB_THECH.

Our results demonstrate that peptide mapping by MS is crucial to differentiate spots that, by their electrophoretic mobility, may be assumed to be similar. As an example, parvalbumins P19 (*pI* 4.32; 11.20 kDa) and P23 (*pI* 4.30; 11.20 kDa), both from different subpopulations of *M. australis*, could be assumed by the 2-DE results as being the same parvalbumin isoform. Nevertheless, MALDI-TOF analysis demonstrated remarkable differences, showing that the isoform from *M. australis polylepsis* (P19) had a high degree of homology with PRVB_MERME, while that from *M. australis australis* (P23) belonged to the group showing homology to PRVB_MERBI.

3.3 Identification of differential masses

The most representative MALDI-TOF masses for each type of parvalbumin (PRVB_A, PRVB_B and PRVB_C) are shown in the Tables 1, 2 and 3, respectively. The complete list of masses can be found in the Supplementary Material (Tables S-3, S-4 and S-5).

In Table 1 there are 34 selected peptide peaks belonging to parvalbumins (PRVB_A) that present homology with PRVB_MERME. Only the peak with *m/z* at 1210.5 Da was shared by all of the PRVB_A isoforms (see also Fig. 1a). In addition, the spectra showed several peaks that were selective for each genus, of them, seven were present in all the species of the genus *Merluccius* and ten only in the *Macruronus* sub-species, (significant examples may be found in Fig. 1b and 1c, respectively). Finally, a peak at 865.32 Da is present in all hakes except in *M. australis polylepsis* (Fig. 1c). The information provided by these peaks is enough to clearly differentiate between hakes and grenadiers using only the MALDI spectra from the parvalbumin isoform with the higher *pI*.

Furthermore, some differences were also obtained among the *Merluccius* species, which may allow, in the first instance, the tentative division of hakes into three groups according to the variations in the PRVB_A isoform sequence. *M. merluccius* and *M. capensis* spectra show a peak at 1193.50 Da which is not present in any of the other *Merluccius* species (Fig. 1d). Instead, the rest of hakes present a peak at 1179.50 Da which is absent in *M. merluccius* and *M. capensis*. The difference of 14 Da between these two masses might be justified by a single amino acid substitution in the PRVB_MERME sequence. In Table 1 it is also shown that the spectrum for *M. australis polylepsis* contains a peak at 1471.54 Da that is specific for this sub-species, and is absent in the spectra of all the other species (Fig. 1e).

Table 2 lists the peaks, together with the tentative peptide sequences, from PRVB_B, the Merlucciidae isoforms presenting homology with PRVB_MERBI sequence (P56503). Four peaks were common to all the species of hakes and grenadiers, one of them (*m/z* 1082.15 Da) is shown in Fig. 1f. Additionally, the spectra showed peaks which were present in all the *Merluccius* species, but not in the *Macruronus* sub-species for example the peak at 1208.14 Da (Fig. 1g). Conversely, other peaks (Fig. 1h) were exclusively present in *Macruronus*. Some other specific peptides, which may be used to differentiate between hake species, are highlighted in Table 2. A peak, represented with 1275.17 Da, is only present in *M. hubbsi*, *M. gayi*, *M. australis polylepsis*, P24 from *M. australis australis*, *M. bilinearis* and in both sub-species of *Macruronus*. However, a specific peak at 1230.97 Da is present in the rest of the species: *M. merluccius*, *M. capensis*, *M. polli*, *M. paradoxus*, *M. productus* and P23 from *M. australis australis*. A unique specific peptide at 1588.29 Da which is present in *M. bilinearis* may be used to differentiate this species from the rest, as can be seen in the Fig. 1i. Peptides that seem to be homologous to this one may be found in hakes (1602.25 Da) and grenadiers (1619.80 Da). Some other interesting differences may be observed at masses at 1906.98 Da, 1934.55 Da, 1945.26 Da, 1959.23 Da.

Finally, in Table 3, a list of masses from parvalbumin isoforms presenting homology with β -parvalbumin from *Alaska pollock* (PRVB_THECH) is presented. Three specific masses were found to be shared by both sub-species of the genus *Macruronus*. Besides, there is only one peak at

Table 1. MALDI-TOF MS analysis of PRVB_A.

O. mass ^{a)}	C. mass ^{a)}	Delta ^{b)}	Peptide	Position	Modification	Spot Number											
						<i>Merluccius</i>						<i>Macrur.</i>					
						P1	P4	P7	P10	P13	P16	P19	P22	P26	P31	P34	
603.77	604.33	−0.56	SAADIK	39–44		■				■				■			
614.88	614.79	0.09	IGLK GK	33–38			■				■						
634.11	634.35	−0.24	LFLQN	65–69		■			■			■					
637.93	638.31	−0.38	AEGSFK	20–25		■											■
678.10																	■
766.23																	■
781.10	781.00	−0.10	LFLQNF	65–70		■					■						
865.32	865.42	−0.04	HGEFFTK	26–32		■						■					
877.40																	■
917.37	917.54	−0.17	GKSAADIKK	37–45					■								
1034.50	1034.55	−0.05	VFGIIDQDK	46–54		■											■
1048.50																	■
1095.49	1095.54	−0.05	TDAETATFLK	78–87		■											
1162.55	1162.64	−0.04	KVFGIIDQDK	45–54		■											
1167.61																	■
1176.60																	■
1179.50																	■
1193.50	1193.62	−0.10	IGVEEFAAMVK	97–107		■											
1208.61																	■
1210.50	1210.54	−0.04	SDFVEEDELK	55–64		■											■
1223.60	1223.65	−0.05	LFLQNFSAGAR	65–75		■											■
1266.60																	■
1276.81	1276.70	0.11	HGEFFTKIGLK	26–36		■											
1280.60	1280.67	−0.07	ALTDAETATFLK	76–87		■											
1471.54																	
1484.85	1484.71	0.14	AEGSFKHGEFFTK	20–32		■											
1891.30																	■
1904.80	1904.95	−0.15	AFAGILADADITAAALAACK	1–19	ACET; CysCAM	■											■
1907.10																	■
1995.80	1995.93	−0.04	AGDSGDGKIGVEEFAAMVK	88–107		■											
2052.90	2052.95	−0.14	AGDSGDGKIGVEEFAAMVKG	88–108		■											
2226.02	2226.08	−0.06	VFGIIDQDKSDFVEEDELK	46–64													
2354.21	2354.17	0.04	KVFGIIDQDKSDFVEEDELK	45–64		■											
2415.10	2415.18	−0.08	SDFVEEDELKFLQNFSAGAR	55–75		■											

a) Peaks represent monoisotopic mass in Da.

b) Observed mass-calculated mass

(■) Denotes the presence of a peak and (□) the absence; O. mass (observed mass) and C. mass (calculated mass), ACET: acetyl N-terminal; CysCAM: carboxy-methylated cysteine.

Macrur. genus *Macruronus*; Spot numbers corresponding to: (P1) *M. merluccius*, (P4) *M. capensis*, (P7) *M. polli*, (P10) *M. paradoxus*, (P13) *M. hubbsi* (P16) *M. gayi*, (P19) *M. australis polylepsis*, (P22) *M. australis australis*, (P26) *M. productus*, (P31) *Ma. nov. nov.*, (34) *Ma. nov. magellanicus*;

541.81 Da which is common for all the species studied, except for spots P3, P6, P9, and P25. There is a second peak at 1529.49 Da specific for *M. hubbsi*, *M. gayi*, *M. australis polylepsis*, *M. productus*, *M. bilinearis* and *Ma. nov. nov.* Some other differential peaks, which cannot be assigned to any sequence, are present in certain species.

Figure 2 presents a schematic flow diagram that can be used for a systematic differentiation. As shown, there are a set of peptides that are specific for *Merluccius*, being absent in

Macruronus. Conversely, there are also peaks always present in grenadiers, but which are never present in hakes. Thus, unequivocal genus differentiation is clearly possible, although further distinction among the grenadier sub-species studied was not possible with the present data. *Merluccius* species may be further classified into two groups, including American hakes and Euro-African hakes. Within these groups the presence or absence of certain peptide masses makes it possible to discriminate among the different

Table 2. MALDI-TOF MS analysis of PRVB_B.

O. mass ^{a)}	C. mass ^{a)}	Delta ^{b)}	Peptide	Position	Modification	Spot Number												
						<i>Merluccius</i>										<i>Macrur.</i>		
						P2	P5	P8	P11	P14	P17	P20	P23	P24	P27	P29	P32	P35
478.28	478.30	−0.02	AFLK	84–87		■						■						
582.70	583.32	−0.62	AFFAK	28–32							■							
633.69						■												
765.84	766.45	−0.61	LFLQVF	65–70					■									
776.19	776.41	−0.22	SADDIKK	39–45		■												
803.77						■												
847.91	848.43	−0.52	ALTDAETK	76–83														
1082.15	1082.55	−0.40	AFFVIDQDK	46–54		■												
1166.24	1166.55	−0.31	SGFIEDELK	55–64		■												
1208.14	1208.67	−0.53	LFLQVFSAGAR	65–75		■												
1210.01	1210.64	−0.63	KAFFVIDQDK	45–54		■												
1230.97						■												
1232.65	1232.64	0.01	KAFFVIDQDK	45–54	Na	■												
1235.85						■												
1275.17	1275.50	−0.33	ACEAADSFNYK	17–27	CysCAM													
1588.29																		
1602.25						■												
1619.80						■												
1906.98																		
1934.55																		
1945.26			AGDSGDGAIGVDEWAALVK	88–107														
1959.23						■												
1967.33																		■

a) Peaks represent monoisotopic mass in Da.

b) Observed mass-calculated mass.

■ Denotes the presence of a peak and □ the absence; O. mass (observed mass) and C. mass (calculated mass); Na: sodium adduct; CysCAM: carboxy-methylated cysteine.

Macrur. genus *Macruronus*. Spot numbers corresponding to: (P2) *M. merluccius*, (P5) *M. capensis*, (P8) *M. polli*, (P11) *M. paradoxus*, (P14) *M. hubbsi*, (P17) *M. gayi*, (P20) *M. australis polylepsis*, (P23) *M. australis australis*, (P24) *M. australis australis*, (P27) *M. productus*, (P29) *M. bilinearis*, (P32) *Ma. nov. nov.*, (P35) *Ma. nov. magellanicus*.

Table 3. MALDI-TOF MS analysis of PRVB_C.

O. mass ^{a)}	C. mass ^{a)}	Delta ^{b)}	Peptide	Position	Modification	Spot Number											
						<i>Merluccius</i>										<i>Macrur.</i>	
						P3	P6	P9	P12	P15	P18	P21	P25	P28	P30	P33	P36
541.81	542.04	−0.23	TFFK	28–31													
767.97	768.05	−0.08	LFLQTF	65–70													
781.11																	
1082.23	1082.55	−0.32	AFFVIDQDK	46–54		■											
1166.25	1166.55	−0.30	SGFIEDELK	55–64		■											
1223.56						■											
1529.49	1529.70	−0.21	AALAGCAAADSFNYK	13–27	CysCAM												
2230.01	2230.10	−0.09	AFFVIDQDKSGFIEDELK	46–64													

a) Peaks represent monoisotopic mass in Da.

b) Observed mass-calculated mass.

■ Denotes the presence of a peak and □ the absence; O. mass (observed mass) and C. mass (calculated mass); CysCAM: carboxy-methylated cysteine

Macrur. genus *Macruronus*; Spot numbers corresponding to: (P3) *M. merluccius*, (P6) *M. capensis*, (P9) *M. polli*, (P12) *M. paradoxus*, (P15) *M. hubbsi*, (P18) *M. gayi*, (P21) *M. australis polylepsis*, (P25) *M. australis australis*, (P28) *M. productus*, (P30) *M. bilinearis*, (P33) *Ma. nov. nov.*, (P36) *Ma. nov. magellanicus*.

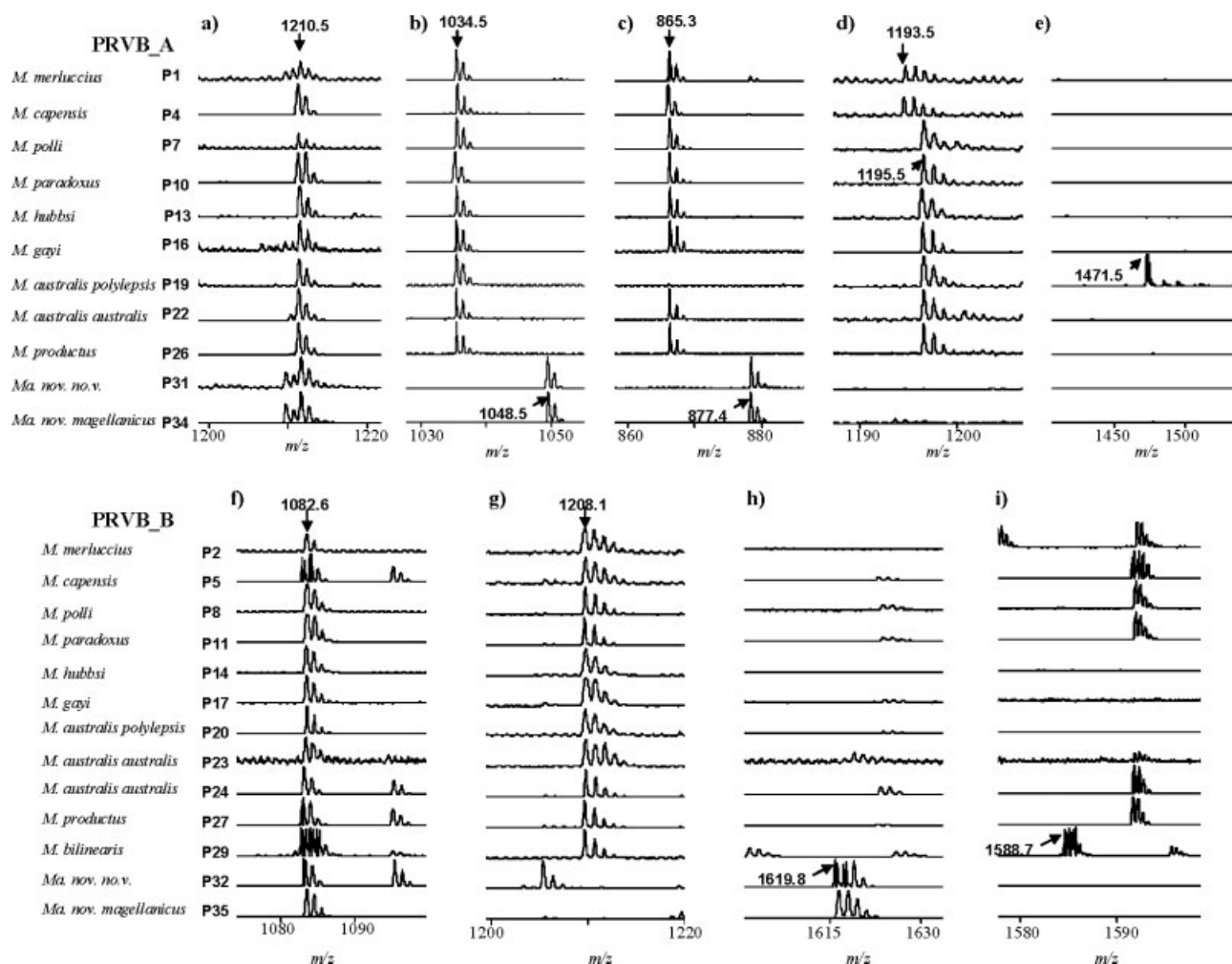


Figure 1. Informative regions of the MALDI-TOF MS peptide maps from the tryptic in-gel digestion of parvalbumin spots showing either, common or specific peaks. There are shown peaks which are: a) present in all PRVB_A isoforms, b) common to all PRVB_A isoforms from genus *Merluccius* and from genus *Macruronus* (1034.5 Da and 1048.5 Da, respectively), c) present in all the PRVB_A isoforms from the genus *Merluccius* except for *M. australis australis* (P19) (865.3 Da) and specific for the genus *Macruronus* (877.4 Da), d) specific for the PRVB_A isoforms from *M. merluccius* and *M. capensis* (1193.5 Da.) and for the rest of hakes (1195.5 Da.), e) specific for the PRVB_A isoform from *M. australis australis* (1471.5 Da.), f) common for all the PRVB_B isoforms, g) common for PRVB_B isoforms from all the hakes, h) common for PRVB_B isoforms from both grenadiers, and i) specific for PRVB_B from *M. bilinearis* (1588.7 Da.).

species, except between *M. merluccius* and *M. capensis* and between *M. hubbsi* and *M. gayi*; these last two species may, however, be distinguished by their pI after 2-DE analysis (Table S-1 in the Supplementary Material).

DNA-based techniques are actually the most required techniques for organism identification purposes, and permit the unambiguous identification of species due to their high sensitivity, robustness and by the resistance to high temperatures of the DNA molecule. However, the proteomics methodologies may be a new complementary method for this purpose. The experiences accumulated in our lab over the last ten years [15–17] have demonstrated the suitability of parvalbumin isoforms, which are fish species-specific, as a forensic species-defining markers for the hake species. Their

high concentration (up to 5 mg/g wet weight) in white muscle of many species, resistance to heating and high inter-specific variability, makes them very adequate for hake species differentiation. Now a further step has been made finding the specific peptides that are responsible for their capability as biomarkers. The sensitivity and selectivity of the MS techniques coupled with the recognition of a serial of specific peptide masses, provide a new strategy fast, straightforward, simple, sensitive and effective means of conducting fish species identification from tissue samples. For instance, specific peptide analysis by selected reaction monitoring, a scan mode available in a wide array of mass spectrometers like triple quadrupoles or ion traps, could allow very rapid hake identification by MS. The potential use of the peptide

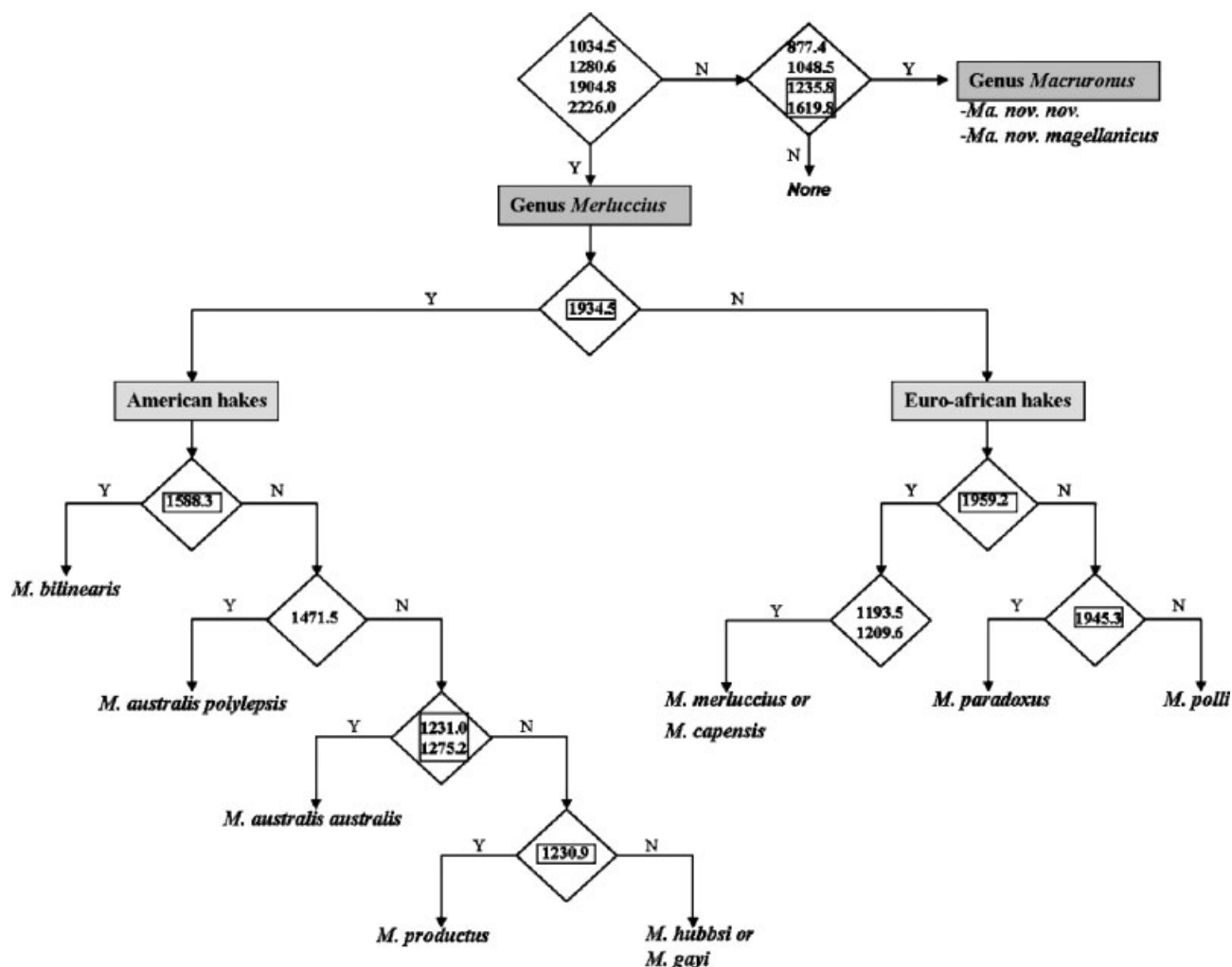


Figure 2. Flow diagram for a systematic identification of hakes and grenadiers using MALDI-TOF data. Specific peak masses from PRVB_B are shown in the boxes. Y denotes presence and N absence of a particular peak.

sequence of hake parvalbumins for identification purposes opens the way to the application of proteomics technology as a new and reproducible strategy for fish authentication. In addition, this research is the first step for the generation of monoclonal anti-peptide antibodies, which may be used in immunological techniques or in protein chips.

3.4 Alignment of the parvalbumins

In Fig. 3, the amino acid sequences for the parvalbumin isoforms from different hakes and grenadiers, as deduced from MALDI-TOF MS data, were aligned with the homologous sequences present in the protein databases: PRVB_MERME, PRVB_MERBI and PRVB_THECH.

Parvalbumins have two EF-hand repeats which bind calcium tightly (EF-1 and EF-2 in Fig. 3) and contain conserved amino acid regions. An invariant aspartic acid residue at position 61 is common to all the parvalbumin sequences

studied. Furthermore, a conserved cysteine at position 18 and an arginine at position 75 are characteristics of the β lineage type [40].

The parvalbumins from the group PRVB_A, present a high degree of identity among the different species and sub-species that were studied. As it can be seen in Fig. 3a), a sequence, SDFVEEDELK, running between amino acids 55 and 62 is common to all the Merlucciidae family. Additionally, there are two sequence regions in PRVB_A that are identical for all the species belonging to the genus *Merluccius*: the acetylated N-terminal sequence (1–19), and the long string that runs between amino acids 39 and 96, and that includes the first calcium binding region (EF-1) and part of the second (EF-2). The sequence HGEFFTK (26–32) is also conserved in the PRVB_A isoform for all the species belonging to the *Merluccius* genus, except for *M. australis polylepsis*. The degree of identity between *Macruronus* and *Merluccius* is limited to the sequences AEGSFK (20–25) and

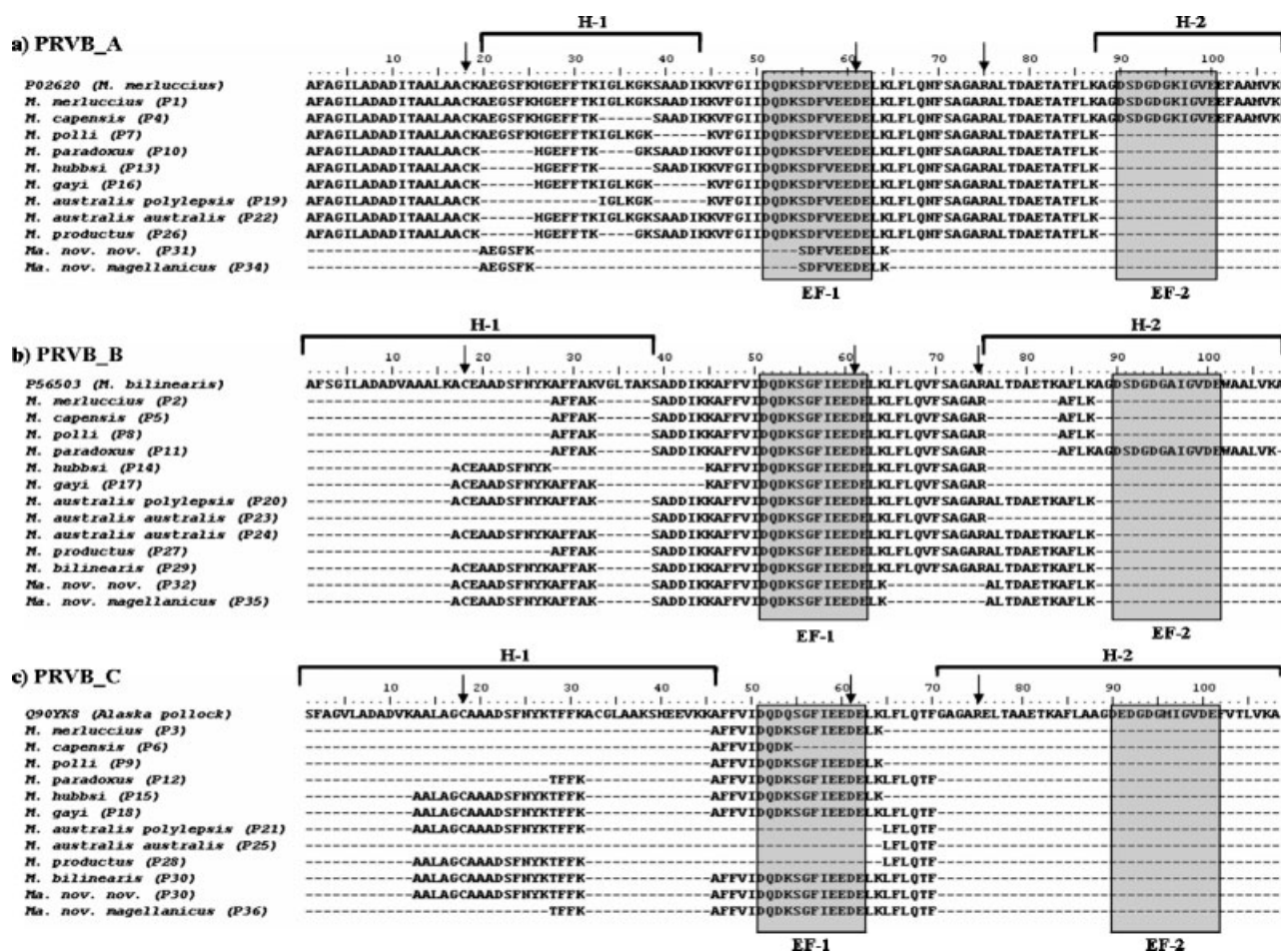


Figure 3. Alignment of the deduced amino acid of parvalbumin sequences, as obtained by MALDI-TOF MS analysis, for all the hakes or grenadiers studied. The sequences were aligned in function of their homology with the sequences present in the databases: a) PRVB_MERME, b) PRVB_MERBI, and c) PRVB_THECH. EF-1 and EF-2 indicate both EF-hand calcium-binding domains which are present in all parvalbumins. Arrows indicate conserved amino acid residues characteristic of the β lineage type. H-1 and H-2 show the amino acid sequence regions presenting a higher theoretical variability.

SDFVEEDELK (55–62). In this group, there are two regions (H-1 and H-2) which display the highest degree of heterogeneity, and that includes part of the second EF domain.

In the PRVB_B sequences (Fig. 3b), there is only a region common to all the species, that is, the peptide KAFFVIDQDKSGFIEEDELK (45–64), which contains the first EF domain. This identity extends to position 75 in the species belonging to the genus *Merluccius*. Regions where the differences with PRVB_MERBI were more noticeable included the two sequences marked as H-1 (1–38) and H-2 (76–108). H-2 comprises the second calcium binding domain, although it does not necessarily mean that this calcium binding domain must be heterogeneous. Long peptides including this domain may be heterogeneous in the last part of the sequence that is not included in the EF-2 domain.

The sequence alignment for PRVB_C isoforms is much poorer, since the closest sequence available in databases, PRVB_THECH, belongs to a fish from a different family,

Alaska pollock (Fig. 3c). There is one common sequence, TFFK (28–31), with a mass coincidence for all the species except for three. Two more sequences, comprising the amino acids between positions 13–31 and 46–64, that are shared by several members of the *Merlucciidae* family, were also detected.

4 Concluding remarks

Proteomics methodology, using the resolution power of 2-DE in combination with fast protein identification by MALDI-TOF MS analysis, was used to find similarities and differences among all the parvalbumin isoforms from nine different hakes and two grenadiers sub-species. The results obtained by 2-DE demonstrated the pI and M_r heterogeneity for parvalbumin spots, which is highly reproducible for each of the species tested. Besides, specific parvalbumin patterns

may be useful for differentiation among members of the family Merlucciidae, providing in some cases an unequivocal identification.

On the other hand, MALDI-TOF MS analysis defined a set of molecular fish authentication markers, relying in the presence or absence of species-specific peptide masses, providing: (i) a selective differentiation between the genus *Merluccius* and *Macruronus*, (ii) a classification of the hake species in two groups according to their geographic precedence, American or Euro-African hakes, and (iii) an unequivocal identification of several hake species.

The authors wish to express their thanks to Lorena Barros and Alberto Jorge García for their excellent technical assistance and to Daniel López-Ferrer for his helpful assistance and suggestions. This work was supported by the Comisión Interministerial de Ciencia y Tecnología (CICYT) (Project AGL2000-0440-P4-02).

5 References

- [1] Lloris, D., Matallanas, J., Oliver, P., *Merluzas del mundo (Familia Merlucciidae) Catálogo comentado e ilustrado de las merluzas conocidas*, FAO Catálogo de especies para los fines de la pesca N°2., Roma, 2003.
- [2] Law 121/2004 23th January, Ministry of Agriculture, Fisheries and Food, Spain.
- [3] Perez, M., Vieites, J. M., Presa, P., *J. Agric. Food Chem.* 2005, 53, 5239–5247.
- [4] Perez, J., Garcia-Vazquez, E., *J. Food Prot.* 2004, 67, 2792–2796.
- [5] Quinteiro, J., Vidal, R., Izquierdo, M., Sotelo, C. G. *et al.*, *J. Agric. Food Chem.* 2001, 49, 5108–5114.
- [6] Ross, C., Hevener, S., Clark, R., Hartmann, J. X., Mari, F., *Prep. Biochem. Biotechnol.* 1998, 28, 49–60.
- [7] Gallardo, J. M., Sotelo, C. G., Piñeiro, C., Pérez-Martín, R. I., *J. Agric. Food Chem.* 1995, 43, 1238–1244.
- [8] Armstrong, S. G., Leach, D. N., Wyllie, S. G., *Food Chem.* 1992, 44, 147–155.
- [9] Kjærsgård, I. V., Nørrelykke, M. R., Jessen, F., *Proteomics* 2006, 6, 1606–1618.
- [10] Morzel, M., Chambon, C., Lefèvre, F., Paboeuf, G., Laville, E., *J. Agric. Food Chem.* 2006, 54, 2997–3001.
- [11] Martinez, I., Solberg, C., Lauritzen, C., Ofstad, R., *Appl. Theor. Electrophor.* 1992, 2, 201–206.
- [12] Martínez, I., Friis, T. J., *Proteomics* 2004, 4, 347–354.
- [13] Piñeiro, C., Barros-Velázquez, J., Vázquez, J., Figueras, A., Gallardo, J. M., *J. Proteome Res.* 2003, 2, 127–135.
- [14] López, J. L., Marina, A., Álvarez, G., Vázquez, J., *Proteomics* 2002, 2, 1658–1665.
- [15] Piñeiro, C., Barros-Velázquez, J., Sotelo, C. G., Gallardo, J. M., *Z. Lebensm. Unters. Forsch. A.* 1999, 208, 342–348.
- [16] Piñeiro, C., Barros-Velázquez, J., Sotelo, C. G., Pérez-Martín, R. I., Gallardo, J. M., *J. Agric. Food Chem.* 1998, 46, 3991–3997.
- [17] Piñeiro, C., Vázquez, J., Marina, A. I., Barros-Velázquez, J., Gallardo, J. M., *Electrophoresis* 2001, 22, 1545–1552.
- [18] Krestinger, R. H., Nelson, D. J., *Coord. Chem. Rev.* 1976, 18, 29–124.
- [19] Krestinger, R. H., in: Carafoli, E. (Eds.), *Hypothesis: Calcium Modulated Proteins Contain EF-hands in Calcium Transport in Contraction and Secretion*, North-Holland Publishing, Amsterdam, 1975, pp. 469.
- [20] Krestinger, R. H., Nockolds, C. E., *J. Biol. Chem.* 1973, 248, 3313–3326.
- [21] Heizmann, C. W., Berchtold, M. W., Rowlerson, A. M., *Proc. Natl. Acad. Sci. USA* 1982, 79, 7243.
- [22] Van Do, T., Hordvik, I., Endresen, C., Elsayed, S., *Mol. Immunol.* 2003, 39, 595–602.
- [23] Jiang, Y., Johnson, J. D., Rall, J. A., *Am. J. Physiol.* 1996, 270, C411–C417.
- [24] Gillis, J. M., Thomason, D., Lefevre, J., Krestsinger, R. H., *J. Muscle Res. Cell. Motil.* 1982, 3, 377–398.
- [25] Bardina, L., Vukic, E. J., Sampson, H. A., Beyer, K., *J. Allergy Clinical Immunol.* 2004, 113, 237.
- [26] Swoboda, I., Bugajska-Schretter, A., Verdino, P., Keller, W. *et al.*, *J. Immunol.* 2002, 168, 4576–4584.
- [27] Bugajska-Schretter, A., Grote, M., Vangelista, L., Valent, P. *et al.*, *Gut* 2000, 46, 661–669.
- [28] Van Do, T., Hordvik, I., Endresen, C., Elsayed, S., *Scand. J. Immunol.* 1999, 50, 619–625.
- [29] Lindstrom, C. D. V., Van Do, T., Hordvik, I., Endresen, C., Elsayed, S., *Scand. J. Immunol.* 1996, 44, 335–344.
- [30] Capony, J. P., Ryden, L., Demaille, J., Pechere, J. F., *Eur. J. Biochem.* 1973, 32, 97–108.
- [31] Revett, S. P., King, G., Shabanowitz, J., Hunt, D. F. *et al.*, *Protein Sci.* 1997, 6, 2397–2408.
- [32] Cox, J., Milos, M., McManus, J. P., *J. Biol. Chem.* 1990, 265, 6633–6637.
- [33] Williams, T. C., Corson, D. C., Oikawa, K., McCubbin, W. D. *et al.*, *Biochemistry* 1986, 29, 1829–1840.
- [34] Tsalkova, T. N., Privalov, P. L., *J. Mol. Biol.* 1985, 181, 533–544.
- [35] Filimonov, V. V., Pfeil, W., Tsalkova, T. N., Privalov, P., *Bio-phys. Chem.* 1978, 8, 117–122.
- [36] Pechere, J. F., in: Wasserman, R. H., Corradino, R., Carafoli, E. *et al.* (Eds.), *The Significance of Parvalbumin Among Muscular Calcium Proteins in Calcium-Binding Proteins and Calcium*, Elsevier, Amsterdam 1997, pp. 213.
- [37] Rehbein, H., Kundiger, R., Piñeiro, C., Pérez-Martín, R. I., *Electrophoresis* 2000, 21, 1458–1463.
- [38] Piñeiro, C., Barros-Velázquez, J., Pérez-Martín, R. I., Gallardo, J. M., *Food Chem.* 2000, 70, 241–245.
- [39] Jensen, O. N., Wilm, M., Shevchenko, A., Mann, M., *Methods Mol. Biol.* 1999, 112, 513–530.
- [40] Krestsinger, R. H., *CRC Crit. Rev. Biochem.* 1980, 8, 119–174.