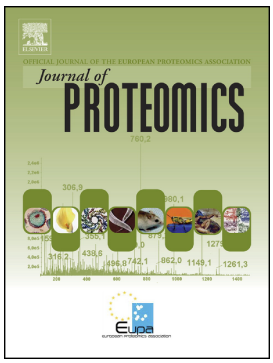
Accepted Manuscript



Molecular characterization of B-cell epitopes for the major fish allergen, parvalbumin, by shotgun proteomics, protein-based bioinformatics and IgE-reactive approaches

Mónica Carrera, África González-Fernández, Susana Magadán, Jesús Mateos, Lelia Pedrós, Isabel Medina, José M. Gallardo

PII: S1874-3919(19)30116-2
DOI: https://doi.org/10.1016/j.jprot.2019.04.005
Reference: JPROT 3355

To appear in: *Journal of Proteomics*

Received date: 16 November 2018
Revised date: 30 March 2019
Accepted date: 7 April 2019

Please cite this article as: M. Carrera, Á. González-Fernández, S. Magadán, et al., Molecular characterization of B-cell epitopes for the major fish allergen, parvalbumin, by shotgun proteomics, protein-based bioinformatics and IgE-reactive approaches, Journal of Proteomics, https://doi.org/10.1016/j.jprot.2019.04.005

Molecular Characterization of B-Cell Epitopes for the Major Fish Allergen, Parvalbumin, by Shotgun Proteomics, Protein-based Bioinformatics and IgE-reactive Approaches

Mónica Carrera,[1*] África González-Fernández,[2] Susana Magadán,[2] Jesús Mateos,[1] Lelia Pedrós,[3] Isabel Medina,[1] José M. Gallardo[1]

[1]Spanish National Research Council (CSIC), Marine Research Institute (IIM), Food Technology, Vigo, Pontevedra, Spain.

[2]Biomedical Research Center (CINBIO) (Centro Singular de Investigación de Galicia), University of Vigo, Immunology, Vigo, Pontevedra, Spain.

[3]Hospital Meixoeiro of Vigo, Galicia Sur Health Research Institute, Complexo Hospitalario Universitario de Vigo (CHUVI), SERGAS, Vigo, Pontevedra, Spain

[*]**CORRESPONDING AUTHOR:** Dr. Mónica Carrera

Spanish National Research Council (CSIC), Marine Research Institute (IIM), Food Technology, Eduardo Cabello 6, 36208 Vigo, Pontevedra, Spain. Phone: +34 986231930. Fax: +34 986292762.

E-mail: mcarrera@iim.csic.es

ABSTRACT

Parvalbumins beta (β-PRVBs) are the main fish allergens. The only proven and effective treatment for this type of hypersensitivity is to consume a diet free of fish. We present the molecular characterization of B-cell epitopes by shotgun proteomics of different β-PRVBs combined with protein-based bioinformatics and IgE-reactive approaches. The final goal of this work is to identify potential peptide vaccine candidates for fish allergy. Purified β-PRVBs from the main fifteen different fish species that cause allergy were analyzed by shotgun proteomics. Identified β-PRVBs peptide sequences and ninety-eight β-PRVB protein sequences from UniProtKB were combined, aligned and analyzed to determine B-cell epitopes using the Kolaskar and Tongaonkar algorithm. The highest rated predicted B-cell peptide epitopes were evaluated by ELISA using the corresponding synthetic peptides and sera from healthy and fish allergic patients. A total of 35 peptides were identified as B-cell epitopes. The top B-cell peptide epitopes (LKLFLQV, ACAHLCK, FAVLVKQ and LFLQNFV) that may induce protective immune responses were selected as potential peptide vaccine candidates. The 3D model of these peptides were located in the surface of the protein. This study provides the global characterization of B-cell epitopes for all β-PRVBs sequences that will facilitate the design of new potential immunotherapies.

KEYWORDS: shotgun proteomics; bioinformatics; parvalbumin; fish allergy; B-cell epitopes; food

SIGNIFICANCE:

This work provides the global characterization of B-cell epitopes for all β-PRVBs sequences by Shotgun Proteomics combined with Protein-based Bioinformatics and IgE-reactive approaches. This study will increase our understanding of the molecular mechanisms whereby fish allergens elicit allergic reactions and will facilitate the design of new potential peptide vaccine candidates.

1. INTRODUCTION

Food allergy is considered one of the major concerns in food safety. It is estimated to affect 6-8% of children and about 2-4% of adults [1]. The prevalence of this type of immunoglobulin E (IgE)-mediated allergy (Type-I hypersensitivity) has increased during the last two decades [2]. Currently the most frequent causes of food allergies are to peanuts, milk, eggs, shellfish and fish [1,3].

Parvalbumins beta (β-PRVBs) are considered the major fish allergens [4-6]. These proteins are present in elevated quantities in the sarcoplasmic fraction of white muscle of fish species. β-PRVBs have an acidic p*I* (3.0-5.0), a molecular weight of 10-12 kDa and three EF-hand motifs, two of which have a high affinity for Ca^{2+}. The first identified fish allergen was the β-PRVB of the cod *Gadus callarias*, also named Gad c1 protein or allergen M [4]. Proteomics data published by our group allowed the extensive *de novo* mass spectrometry (MS) sequencing of 41 new β-PRVB isoforms from the Merlucciidae family by a combination of Bottom-Up proteomics, accurate molecular mass measurement by FTICR-MS and selected MS/MS ion monitoring [7, 8]. Therefore, 98 β-PRVB protein sequences from all Teleostei species are currently available in the UniProtKB database (August 2018).

Fish-allergic individuals are frequently sensitive to several fish species [9, 10]. The symptomatology of this allergy arise within 60 minutes of ingestion and comprise rash, abdominal pain, vomiting, diarrhea and respiratory distress [11]. In the most serious cases, anaphylactic shock may occur [12]. The allergenic properties of β-PRVBs seem to be related to their resistance to heat and to certain gastrointestinal proteases [13, 14]. As a food control method, a new targeted proteomics strategy published by our group that is based on the monitoring of several β-PRVB peptide biomarkers using MS achieves the fast detection of this allergen in any food product in less than two hours [15]. This method covers all β-PRVB isoforms, is faster than the DNA-amplification techniques and avoid the cross-reactivity problems of the immunological techniques [16, 17].

Type-I food allergy involves two main phases: sensitization, followed by the effector phase [18, 19]. The sensitization phase occurs after interaction with the ingested allergen and comprises a succession of events (T-cell and B-cell activation) leading to the overproduction of allergen-specific IgE and its subsequent binding to the high-affinity IgE receptor FcεRI on the membrane of mast cells and basophils. In the effector phase, the allergen crosses the intestinal epithelium and cross-links with the IgE-FcεRI complexes, causing effector cell activation and the secretion of allergy factors.

The mechanism of sensitization of allergic individuals is initiated upon the presentation of the allergen-derived peptides by the antigen-presenting cells through the major histocompatibility class II molecule to naive T cells [18]. Then, these naive T cells are activated and differentiated into type 2 T helper cells (Th2). Cytokines, such as IL-4 and IL-13 released from Th2 cells, produce a class change in the allergen-specific B cells to secrete IgE antibodies directed against the precise allergen. These allergen-specific IgE antibodies bind to FcεRI receptors on the surface of basophils/mast cells. When an individual is exposed to the same allergen again, the allergen induces the cross-linking of two IgE molecules and

basophils and mast cells become activated to secrete inflammatory factors such as cytokines, histamine and leukotrienes, leading to an instantaneous allergic inflammation.

To date, the only proven and effective treatment for this type of hypersensitivity is to consume a diet free of the allergenic food and its derivatives. Additionally, allergen-specific immunotherapy (AIT), which is based on the administration of the allergenic extracts in the form of therapeutic vaccine to induce immunological tolerance, is a common practice. Mechanisms of AIT efficacy include the production of allergen-specific IgG antibodies, which may interfere with IgE recognition and suppress the allergic symptoms caused by IgE-allergen immune complexes to eventually induce tolerance [20]. However, the use of crude extracts has several disadvantages; these extracts may induce severe anaphylactic side reactions or sensitization towards new allergens present in the mixture [13, 21]. Furthermore, AIT generally requires complicated up-dosing schedules and repeated administrations, leading to poor patient compliance. Different strategies have been designed to try to overcome these negative effects, as the developed of hypoallergenic allergens, as the recently published hypoallergenic β-PRVB for AIT [22].

New developments in the characterization of epitopes are providing targets for the development of novel AIT [22-24]. In this sense, significant advances in B-cell epitope-based allergy vaccines have been recently achieved, including promising results applied in clinical trials [25, 26]. Limited studies have mapped epitopes in some fish species such as cod, salmon and trout using synthetic peptide immunoassays [27, 28]. However, to date, the global analysis of the B-cell epitopes for all β-PRVBs registered in the databases (98 β-PRVBs, UniProtKB) has not yet been carried out.

Thus, the emerging use of allergen-derived B-cell epitope databases has significantly assisted in the development of epitope-based immunotherapies that aim to modulate the immune responses of patients towards specific allergens. Different computer programs that

predict specific epitopes have been developed [29-32]. Currently B-cell epitopes are able to be predicted using different algorithms based on various physicochemical features of amino acids, such as hydrophilicity, flexibility, beta-turns and surface accessibility [33]. Recently, a computer strategy based on the characterization of B-cell epitopes was published for the vaccine design of a virus [34]. However, to our knowledge, no previous reports have been developed for the characterization by protein-based bioinformatics of B-cell epitopes for further development of new therapeutic treatments for fish β-PRVB allergens.

Therefore, in this study, we present for the first time the extensive characterization of B-cell epitopes for β-PRVBs. The strategy used comprises five consecutive steps: (*i*) shotgun proteomics analysis of β-PRVBs from the most allergenic 15 different fish species (ii) downloading all 98 β-PRVB protein sequences for all fish species included in the UniProtKB database, (*iii*) protein-based bioinformatics of B-cell epitopes, (*iv*) synthesizing the selected B-cell peptide epitopes and (*v*) performing of immunochemistry analyses using sera from healthy and allergic patients.

The findings from this study might provide a rationale repository of B-cell epitopes for the design of new specific immunotherapies for fish allergy.

2. MATERIALS AND METHODS

2.1 Fish species and β-PRVBs purification

Fifteen different fish species were used in this work (Table 1). These species were acquired from marketplaces in order to comprise the most consumed fish species in Europe containing the main species that cause fish allergy [24]. These species were genetically identified using the fishID Kit (Bionostra, Madrid, Spain).

Extraction of sarcoplasmic proteins was prepared as described [35]. Briefly, 0.5 g of fish white muscle were homogenized in 4 mL of 10 mM Tris-HCl buffer, pH 7.2, with 5 mM of PMSF, for 2 min using an Ultra-Turrax device (IKA Werke, Staufen, Germany). β-PRVBs were purified by taking advantage of their thermostability [7]. After centrifugation at 40000 g for 20 min (J221-M centrifuge, Beckman, Palo Alto, CA, USA), supernatants containing principally β-PRVBs were quantified by the BCA method (Sigma-Chemical Co., USA). Samples were analyzed in triplicate.

2.2 Shotgun proteomics of β-PRVBs

Proteins were in-solution digested with trypsin [36]. A total of 100 μg of proteins were denatured in 8 M urea and reduced with 5 mM TCEP for 30 min at 37ºC. After the alkylation

with 50 mM iodoacetamide for 60 min, samples were diluted 4-fold with 25 mM ammonium bicarbonate pH 8.25. Proteins were digested with trypsin (Promega) (1:100 protease to protein) overnight at 37ºC.

Peptide digests were acidified with formic acid, cleaned on C18 MicroSpin™ columns (The Nest Group, South-borough, MA) and analyzed by LC-MS/MS using a Proxeon EASY-nLC II LC system (Thermo Fisher Scientific, San Jose, CA) coupled to an LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific). Separation of peptides (1 µg) was performed on a RP column (75 µm x 10 cm) packed with C18 resin (Magic C18 AQ 3 µm; Michrom BioResources, Auburn, CA) using 0.1% formic acid and 98% ACN in 0.1% formic acid as mobile phases A and B, respectively. A linear 60 min gradient from 5 to 35% B was used at a flow rate of 300 nL/min. The spray voltage was 1.95 kV and the capillary temperature 230 °C. Peptides were analyzed from 400 to 1600 amu (1 µscan), followed by four data-dependent MS/MS scans (1 µscans), using an isolation width of 3 amu and a normalized collision energy of 35%. Fragmented masses were set in dynamic exclusion for 30 s after the second fragmentation event. Singly charged ions were excluded from analysis.

MS/MS spectra were searched using SEQUEST (Proteome Discoverer 2.1, Thermo Fisher Scientific), against the Teleostei UniProt/TrEMBL database (release 2017_12; 158.545 entries). The following restrictions were used: semi-tryptic cleavage with up to two missed cleavage sites and tolerances 1.2 Da for precursor ions and 0.5 Da for fragments ions. The variable modifications allowed were carbamidomethylation of cysteine, methionine oxidation and acetylation of the N-terminus of the protein. Results were subjected to statistical analysis with the PeptideProphet algorithm. The FDR was kept below 1%.

2.3 Bioinformatics analysis of β-PRVB sequences and B-cell epitopes

All β-PRVB peptide sequences identified by shotgun proteomics were analyzed using BLASTp algorithm to determine homologies with protein sequences registered in the NCBI database [37]. Additionally, 98 β-PRVB protein sequences (58 Swiss-Prot and 40 TrEMBL) for all Teleostei species that are registered in the UniProtKB database were downloaded in FASTA format. All β-PRVB sequences were aligned with Clustal W (BioEdit, http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

The method reported by Kolaskar and Tongaonkar on the Immune Epitope Database and Analysis Resource (IEDB) (http://tools.immuneepitope.org/bcell/) was applied to compute linear B-cell epitopes for each of the β-PRVB protein sequences [33]. The method predicts epitopes with approximately 75% accuracy [33]. According to the Kolaskar and Tongaonkar method a window size of 5 to 7 amino acids was selected as appropriate for finding regions that may potentially be antigenic [33]. The features of the B-cell epitopes include surface accessibility, flexibility and hydrophilicity [38, 39].

2.4 Collection of sera from allergic patients and healthy controls

Serum samples from 12 allergic patients with a clinical history of allergic reactions to fish were obtained from the Hospital Meixoeiro of Vigo (Spain) and from the Hospital Carlos III of Madrid (Table 2). The diagnosis of IgE-mediated fish allergy was confirmed by measuring cod-specific IgE antibodies. Sera from three healthy donors with a negative clinical history for any Type-I allergy were used as negative controls. Informed consent from all subjects were provided and all protocols were approved by the board of the local institutional authority (CEIC of Galicia, Spain).

2.5 Synthetic peptide epitopes

Selected B-cell peptide epitopes from β-PRVB were chemically synthetized (Track Peptide Libraries, JPT Peptide Technologies GmbH, Berlin, Germany) (Table 3). The lyophilized peptides were individually reconstituted with 1x PBS and stored at -80°C. Additionally, as negative control peptides were used synthetic peptides with the same length (7 residues) but with a shuffled amino acid sequence (ELLINVK) belonging to the alcohol dehydrogenase protein of *Sacharomyces cerevisiae* and a purity of 91.2% (AQUA Peptides, Sigma-Aldrich, S.L., Madrid, Spain).

2.6 ELISA

A 20 µg/mL solution of each B-cell synthetic peptide or negative control peptides were used to coat 96-well high-binding plates (50 µL/well) (Costar Corning Inc., New York, USA). Plates were incubated overnight at 4°C and then blocked with PBS/1% BSA (w/v) (200 µL/well) for 2 h at 37°C. After three washes with PBS/0.05% Tween 20, plates were incubated with 50 µL of the sera from different fish allergic patients and healthy donors (1:5 in PBS/1% BSA) for 1.5 h at 37°C. After three washes, bound IgE antibodies were detected using 50 µL per well of goat anti-human IgE-HRP (1:1000 in PBS/1% BSA) (Sigma-Aldrich, St. Louis, MO, USA) for 1.5 h at 37°C. After washing, the colorimetric reaction was developed following the addition of 50 µL of o-phenylene-diamine (Sigma), and stopped by the addition of 50 µL of 3 M sulfuric acid. The optical density (OD) was measured at 492 nm using an ELISA Microplate Reader (Multiskan^{TM} GO, Thermo Fisher Scientific). Analyses were performed in triplicate. The OD value obtained for each B-cell synthetic peptide was corrected by subtracting the OD value determined in negative control peptide samples. Finally, the mean intensities for each peptide epitope are reported.

2.7 Euclidean hierarchical clustering and box-plot analysis

ELISA data were evaluated using the statistical package R version 3.4.1 (http://www.r-project.org). Heat map analysis and hierarchical clustering were conducted using the function *heatmap.2()* on R, using the Euclidean distance and the complete linkage method. Box-plot graphics were performed using the *boxplot()* command showing the outliers on R package.

2.8 3D structural modeling

The Swiss-Model server (http://swissmodel.expasy.org) was used to create the 3D models and the Swiss-PdbViewer 4.1 to visualize the 3D structure and to find the sequence of interest.

3. RESULTS

3.1 Shotgun proteomics analysis of β-PRVBs

A shotgun proteomics approach was used to generate a reference dataset of identified β-PRVBs proteins and peptides for each of the different 15 fish species (Table 1). Shotgun proteomics refers to the analysis of a mixture of proteins that are digested with a protease, and the resulting mixture of peptides are then analyzed by LC-MS/MS. The spectra obtained are assigned to peptide sequences using database searching algorithms, and the identification of these peptides allows the identification of the proteins present in the complex mixture. For that, β-PRVBs for each of the different fish species, were purified from the sarcoplasmic extracts taking advantage of their thermostability (7). Complete list of sarcoplasmic proteins and peptides for each of the different species, before and after the treatment with heat, are present in the Table S1 (original) and Table S2 (heated) in the Supplementary repository. The majority of identified peptides in the heated samples corresponded to β-PRVBs (30-90%). Table S3 in the Supplementary repository presents the complete list of non-redundant β-PRVBs peptides for each of the different 15 fish species.

Table 4 summarizes the list of identified β-PRVBs proteins and non-redundant peptides analyzed by LC-MS/MS and identified by SEQUEST for each of the different 15 fish species. A total of 243 non-redundant β-PRVBs peptides corresponding to a total of 7674 spectra counts were identified. Additionally, taking into account the availability of entries in the NCBI database, the results obtained by BLASTp algorithm allowed determining homologies with the β-PRVBs registered in the protein databases. Thus, the results showed a sequence coverage with β-PRVBs between 35%-82% (corresponding to samples S12 and S8).

To our knowledge, this is the most comprehensive dataset of β-PRVBs peptides from different fish species identified to date. This valuable repository will add new and significant information to the universal public protein databases.

3.2 β-PRVB sequences

All β-PRVB peptide sequences identified by shotgun proteomics and analyzed using BLASTp were aligned by Clustal W (Figure 1). Additionally, 98 β-PRVB protein sequences (58 Swiss-Prot and 40 TrEMBL) for all Teleostei species registered in the UniProtKB database were included in the analysis (Figure 2). All these protein sequences correspond to 41 different teleost fish species belonging to 15 different families, the majority of which are included in the human diet (Figure 3).

Figure 4A shows in detail the multiple sequence alignment of all β-PRVB protein sequences generated using Clustal W. β-PRVBs represent one of the 32 subclasses within the EF-hand superfamily, a group of proteins with several highly conserved helix-loop-helix (EF-hand) motifs that bind Ca^{2+} [40]. β-PRVBs include three EF-hand motifs, named as AB, CD and EF. Among these regions, only two (CD and EF) are functional in chelating Ca^{2+}. Based on the results of the sequence alignment, the most conserved region corresponded to the sequence between amino acids 46 to 77 (Figure 4A). This sequence contains one EF-hand motif (CD domain) [40], which is composed of a central 12-residue Ca^{2+}-binding loop flanked by two α-helixes positioned perpendicular to each other [41]. The Ca^{2+} ion is coordinated by the conserved amino acids located in positions (x: Asp53, y: Asp55, z: Ser57, -x: Phe59, -y: Glu61, and -z: Glu64) (Figure 4A) [41].

However to date, information about B-cell peptide epitopes for all β-PRVB proteins is not available in the current databases.

3.3 Protein-based bioinformatics characterization of B-cell epitopes

The Kolaskar and Tongaonkar method available on the IEDB (http://tools.immuneepitope.org/bcell/) was used to predict linear B-cell epitopes for all β-PRVB protein sequences [33]. This method is based on previous experimental data that determined that

the hydrophobic residues Val, Leu and Cys, are more likely to be included in antigenic sites if they occur on the protein surface. Flexibility, surface accessibility, and hydrophilicity among amino acid sequence regions are properties used to predict potential B-cell epitopes. Thus, this semi-empirical method, using the physicochemical features of amino acids and their frequencies on experimentally known protein epitopes, was employed to predict antigenic determinants on proteins with approximately 75% accuracy [33].

Figure 4B shows several representative examples of the linear B-cell epitope regions identified for the β-PRVBs. The EF-hand motifs are hydrophobic regions that function to chelate Ca^{2+} (CD and EF domains), and thus they are predicted to be B-cell epitopes.

Table 5 summarizes the linear B-cell epitopes identified for the β-PRVBs using the Kolaskar and Tongaonkar method. According to this algorithm, a window size of five to seven amino acids is appropriate for finding regions that may potentially be antigenic [33]. Thus, a total of 35 different B-cell epitopes with seven amino acids were identified (Table 5). The highest binding score and occurrence corresponded to the sequences ACAHLCK (residues 1 to 7) and LKLFLQV (residues 65 to 71). The first B-cell epitope (ACAHLCK) is located in the N-terminus of the molecule and is close to the AB motif (Figures 4A and 4B). This domain does not chelate Ca^{2+} because it houses two amino acid deletions in the loop region. The second more relevant B-cell epitope (LKLFLQV) is located adjacent to the CD motif in a hydrophilic region of the molecule (Figures 4A and 4B). Finally, we must emphasize that part of this last B-cell epitope (LKLFLQ) was recently recognized as an immunologically reactive site with amyloid properties [42, 43]. Amyloids are protein aggregated of cross-linked β-sheet-rich states that confer protease resistance to molecules. Thus, the formation of these amyloid fibers and their resistance to gastrointestinal protease digestion may be one of the reasons for the allergenicity of β-PRVBs.

3.4 Validation of B-cell peptide epitopes

The 12 B-cell epitopes with the highest percentage score determined using the Kolaskar and Tongaonkar method were synthesized with >70% purity to validate the predicted B-cell peptide epitopes (Table 3). Additionally, a synthetic peptide with the same length and a shuffled amino acid sequence was used as negative control peptide for the analysis (see 2.5 and 2.6 sections of Materials and Methods). These 12 B-cell peptides were tested against sera from 12 fish allergic patients and sera from 3 healthy volunteers using an indirect ELISA (Table 2). The results of the specific IgE responses for each individual are presented in Table 6.

Comparing with the healthy donors, Figure 5A shows that the sera of the fish allergic patients recognized mainly all the selected β-PRVB synthetic peptides except the sequences Acetyl-AFAGVLA/AFASVLK (residues 1 to 7). It is important to emphasize, that the peptides LKLFLQV (residues 65 to 71) located in a region adjacent to the CD motif; ACAHLCK (residues 1 to 7) located adjacent to the N-terminus region of the sequence; FAVLVKQ (residues 104 to 110) located at the C-terminus of the molecule adjacent to the EF motif; and LFLQNFV (residues 67-73) located adjacent to the CD motif are the major IgE-binding epitopes that are detected by the sera of the fish allergic patients (Table 6 and Figure 5A).

Figure 5B shows the heatmap that groups according to similarity the results of the recognition pattern of 12 fish allergic patients and 3 healthy donors (rows) against the 12 β-PRVB synthetic peptides (columns). The legend color bar indicates the ELISA values, showing high-recognition (red) or no/low-recognition (green). The R program estimated a threshold of 0.31. Based on the results of the heatmap on the x-axis were identified three different clusters. Cluster 1 (***) groups the high-recognized peptides (LKLFLQV, ACAHLCK, FAVLVKQ and LFLQNFV). Partial or longer sequence of some of these high-recognized B-cell epitopes, as ACAHLCKE and LKLFLQNF have been previously identified by other authors as the major IgE-binding epitopes (40-42, 44-45). Cluster 2 (**) is involved by the medium-recognized

peptides (CAHLCKE, FAALVKA, VKKAFFV, AALAACK and AAALAAC). Finally, cluster 3 (*) groups the no/low-recognized peptides (Acetyl-AFASVLK, Acetyl-AFAGVLA and AAALEAC). In the y-axis, were differentiated three different clusters (cluster A, B, and C). The cluster A groups those fish-allergic patients that are more sensitive to the β-PRVB synthetic peptides. Cluster B groups those patients with less sensitivity to the β-PRVB synthetic peptides. Finally, as expected, the triple healthy control samples that showed no recognition to the β-PRVB synthetic peptides were grouped together in the cluster C.

Figure 6 shows the box-plot graphic of the 12 fish allergic patients against the 12 β-PRVB synthetic peptides. The same threshold value (0.31) that was used before in the heatmap was employed for the box-plot graphic. Thus, with a threshold of >0.31 are the high-recognized peptides (LKLFLQV, ACAHLCK, FAVLVKQ, and LFLQNFV) that were also identified in the heatmap in cluster 1 (***). With a threshold of ~0.31 are the medium-recognized peptides (CAHLCKE, FAALVKA, VKKAFFV, AALAACK and AAALAAC) that previously were identified in the heatmap in cluster 2 (**). Finally, with a threshold of <0.31 are the no/low-recognized peptides (Acetyl-AFASVLK, Acetyl-AFAGVLA and AAALEAC) that previously were identified in the heatmap as cluster 3 (*). Therefore, the immunoassay results obtained in this work corroborated the data previously obtained using the computational epitope-based Kolaskar and Tongaonkar software.

3.5 Potential peptide vaccine candidates

Table 7 summarizes the list of four potential peptide vaccine candidates selected for fish allergy. The list was created using the top four B-cell peptide epitopes based on the highest percentage score determined using the Kolaskar and Tongaonkar algorithm and the IgE immunoassay results (LKLFLQV, ACAHLCK, FAVLVKQ and LFLQNFV) (***) (Table 5 and 6; Figures 5A, 5B and 6). As is presented here, the present bioinformatics approach offers a

good and rapid strategy to select potential B-cell peptide vaccine candidates that further needs to be validated using new experiments with cell lines and animal models to obtain an efficient and safe immunotherapy for fish allergy.

3.6 3D structure of peptide vaccine candidates

Figure 7 shows the 3D structure of the four potential peptide vaccine candidates (LKLFLQV, ACAHLCK, FAVLVKQ and LFLQNFV) on the β-PRVBs protein structure. As protein templates were used, for LKLFLQV (P56503, PRVB_MERBI); for ACAHLCK (Q91482, PRVB1_SALSA), for FAVLVKQ (Q91482, PRVB1_SALSA) and for LFLQNFV (P86741, PRVB2_MACMG). All of these four epitopes were localized in the surface of the β-PRVBs. As was described previously, the hydrophobic residues of Val, Leu and Cys, are more likely to be included in antigenic sites if they occur on the protein surface [33].

4. DISCUSSION

In this study, we presented for the first time the extensive characterization of B-cell epitopes for all β-PRVBs, using a shotgun proteomics approach of β-PRVBs from the most allergenic 15 fish species, combined with the protein-based bioinformatics of all 98 β-PRVB protein sequences registered in the protein databases. Then, all β-PRVB protein sequences were analyzed using epitope-based bioinformatics tools in order to identify new prospective peptide vaccines. All protein sequences were screened for B-cell epitopes using the Kolaskar and Tongaonkar algorithm available in the IEDB database. The top B-cell peptide epitopes (LKLFLQV, ACAHLCK, FAVLVKQ and LFLQNFV) that may induce protective immune responses were selected as potential peptide vaccines candidates (Table 7). These B-cell peptide epitopes are good candidates because are localized on the protein surface of the β-PRVBs. All of these peptide candidates will be used in future investigations using cell lines and animal models.

With the goal of generating a vaccine for fish allergy based on B-cell epitopes, the top B-cell peptide epitopes selected in this work (Table 7) could be synthetized and administered orally/sublingually by AIT as a mixture of peptides combined with specific adjuvants in order to induce oral tolerance or anergy [44]. The goals of B-cell epitope-based allergy vaccines are the reduction of allergenic proprieties of immunogens while retaining their immunogenicity to guarantee the induction of allergen-specific IgG antibodies that block the interaction between IgE and the allergen. In addition, the absence of side effects due to the activation of allergen-specific T-cells is also expected. B-cell epitopes identified in this work and validated by ELISA could be chemically coupled to a carrier protein or be produced as recombinant fusion proteins and administered subcutaneously to achieve these objectives. One classical protein carrier is the keyhole limpet hemocyanin, which induces allergen-specific IgG against allergens that induce poor IgG immune responses [45]. Moreover, the inclusion of peptides from the IgE-binding sites

would focus allergen-specific IgG responses to the β-PRVB IgE-epitopes, potentially resulting in better protection and a reduced number of vaccine administrations. Another interesting option is the production of recombinant fusion proteins in which allergen peptides are linked to viral proteins such as the VP1 coat protein from human rhinovirus or the PreS domain from hepatitis B virus (HBV); this strategy has been applied to different allergens [26]. The best example is BM32, the recombinant B cell epitope-based for grass pollen allergy vaccine, in which peptides from the IgE binding sites of the allergen were fused to PreS-HBV and produced as recombinant fusion proteins in *Escherichia coli* [46]. The results of different works, including clinical trials, confirmed a good safety profile for this vaccine and its capacity of inducing a highly selective allergen-specific IgG response that does not boost allergen-specific IgE responses [47]. In addition, advances in nanotechnology and polymer design have enabled the design of new delivery systems for immunization therapy that are also suitable for mucosal and oral administration routes [48, 49].

Therefore, the results of this study using shotgun proteomics, protein-based bioinformatics and Ig-E reactive approaches, provide for the first time the global characterization of B-cell epitopes for the major fish allergen β-PRVB. Comparing the different strategies discussed here, this repository of B-cell epitopes will be very useful for further development of new therapeutic treatments based on peptide vaccines for fish allergy.

REFERENCES

[1] Renz H, Allen KJ, Sicherer SH, Sampson HA, Lack G, Beyer K, Oettgen HC. Food allergy. Nat Rev Dis Primers. 2018;4:17098.

[2] Ben-Shoshan M, Turnbull E, Clarke A, Food allergy: temporal trends and determinants. Curr Allergy Asthma Rep. 2012;12:346-72.

[3] Venter C, Arshad SH. Epidemiology of food allergy. Pediatr Clin North Am. 2011;58:327-49.

[4] Elsayed S, Bennich H. The primary structure of allergen M from cod. Scand J Immunol. 1975;4:203-8.

[5] Swoboda I, Bugajska-Schretter A, Verdino P, Keller W, et al. Recombinant carp parvalbumin, the major cross-reactive fish allergen: a tool for diagnosis and therapy of fish allergy. J Immunol. 2002;168:4576-84.

[6] Ruethers T, Taki AC, Johnston EB, Nugraha R, Le TTK, Kalic T, McLean TR, Kamath SD, Lopata AL. Seafood allergy: A comprehensive review of fish and shellfish allergens. Mol Immunol. 2018;100:28-57.

[7] Carrera M, Cañas B, Vázquez J, Gallardo JM. Extensive de novo sequencing of new parvalbumin isoforms using a novel combination of bottom-up proteomics, accurate molecular mass measurement by FTICR and selected MS/MS ion monitoring. J Proteome Res. 2010;9: 4393-4406.

[8] Carrera M, Cañas B, Gallardo JM. Advanced proteomics and systems biology applied to study food allergy. Curr Opin Food Sci. 2018;22: 9-16.

[9] Hansen TK, Bindslev-Jensen C, Skov PS, Poulsen LK. Codfish allergy in adults: IgE cross-reactivity among fish species. Ann Allergy Asthma & Immunol. 1997;78:187-194.

[10] Perez-Gordo M, Cuesta-Herranz J, Maroto AS, Cases B, et al. Identification of sole parvalbumin as a major allergen: study of cross-reactivity between parvalbumins in a Spanish fish-allergic population. Clin Exp Allergy. 2011;41:750-8.

[11] Jeebhay MF, Robins TG, Lehrer SB, Lopata AL. Occupational seafood allergy: a review. Occup Environ Med. 2001;58:553-62.

[12] Sampson HA. Anaphylaxis and emergency treatment. Pediatrics. 2003;111:1601-8.

[13] Astwood JD, Leach JN, Fuchs RL. Stability of food allergens to digestion in vitro. Nat Biotechnol. 1996;14:1269-73.

[14] Arif SH, Jabeen M, Hasnain A. Biochemical characterization and thermostable capacity of parvalbumins: the major fish-food allergens. J Food Biochem. 2007;31:121-37.

[15] Carrera M, Cañas B, Gallardo JM. Rapid direct detection of the major fish allergen, parvalbumin, by selected MS/MS ion monitoring mass spectrometry. J Proteomics. 2012;75: 3211-20.

[16] Aranishi F, Okimoto T. PCR-based detection of allergenic mackerel ingredients in seafood. J Genet. 2004;83:193-5.

[17] Gajewski KG, Hsieh YN. Monoclonal antibody specific to a major fish allergen: parvalbumin. J Food Prot. 2009;72:818-25.

[18] Berin MC. Pathogenesis of IgE-mediated food allergy. Clin Exp Allergy. 2015;45:1483-96.

[19] Sicherer SH, Sampson HA. Food allergy: Epidemiology, pathogenesis, diagnosis, and treatment. J Allergy Clin Immunol. 2014;133:291-307.

[20] Reisacher WR, Davison W. Immunotherapy for food allergy. Curr Opin Otolaryngol. Head Neck Surg. 2017;25:235-41.

[21] Bannon GA. What makes a food protein an allergen? Curr Allergy Asthma Rep. 2004;4: 43-6.

[22] Swoboda I, Balic N, Klung C, Focke M. et al. A general strategy for the generation of hypoallergenic molecules for the immunotherapy of fish allergy. J Allergy Clin Immunol. 2013;132:979-81.

[23] Bernhiesel-Broadbent J, Scanlon SM, Sampson H. Fish hypersensitivity. I. In vitro and oral challenge results in fish-allergic patients. J Allergy Clin Immunol. 1992;89:730-7.

[24] Tanaka R, Ichikawa K, Hamano K. Clinical characteristics of seafood allergy and classification of 10 seafood allergens by cluster analysis. Arerugi. 2000;49:479-86.

[25] Zieglmayer P, Focke-Tejkl M, Schmutz R, Lemell P, Zieglmayer R, Weber M, Kiss R, Blatt K, Valent P, Stolz F, Huber H, Neubauer A, Knoll A, Horak F, Henning R, Valenta R. Mechanisms, safety and efficacy of a B cell epitope-based vaccine for immunotherapy of grass pollen allergy. EBioMedicine. 2016;11:43-57.

[26] Valenta R, Campana R, Niederberger V. Recombinant allergy vaccines based on allergen-derived B cell epitopes. Immunol Lett. 2017;S0165-2478:30180-3.

[27] Perez-Gordo M, Pastor-Vargas C, Lin J, Bardina L, et al. Epitope mapping of the major allergen from Atlantic cod in Spanish population reveals different IgE-binding patters. Mol Nutr Food Res. 2013;57:1283-90.

[28] Kobayashi A, Ichimura A, Kobayashi Y, Shiomi K. IgE-binding epitopes of various fish parvalbumins exist in a stereoscopic conformation maintained by Ca(2+) binding. Allergol Int. 2016;65:345-8.

[29] Salimi N, Fleri W, Peters B, Sette A. Design and utilization of epitope-based databases and predictive tools. Immunogenetics. 2010;62:185-96.

[30] Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. Bioinformatics. 2001;17:1236-7.

[31] McClain S. Bioinformatics screening and detection of allergen cross-reactive IgE-binding epitopes. Mol Nutr Food Res. 2017;61:1600676.

[32] Jandrlić DR, Lazić GM, Mitić NS, Pavlović MD. Software tools for simultaneous data visualization and T cell epitopes and disorder prediction in proteins. J Biom Inform. 2016;60:120-31.

[33] Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Letters. 1990;1-2:172-4.

[34] Tambunan US, Sipahutar FR, Parikesit AA, Kerami D. Vaccine design for H5N1 based on B- and T-cell epitope predictions. Bioinform Biol Insights. 2016;10:27-35.

[35] Carrera M, Cañas B, Piñeiro C, Vázquez J, Gallardo JM. De novo mass spectrometry sequencing and characterization of species specific peptides from nucleoside diphosphate kinase B for the classification of commercial fish species belonging to the family Merlucciidae. J Proteome Res. 2007;6:3070-80.

[36] Carrera M, Cañas B, Gallardo JM. The sarcoplasmic fish proteome: pathways, metabolic networks and potential bioactive peptides for nutritional inferences. J Proteomics. 2013;78:211-20.

[37] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403-10.

[38] Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis a virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol. 1985;55:836-9.

[39] Karplus PA, Schulz GE. Prediction of chain flexibility in proteins – A tool for the selection of peptide antigens. Naturwissenschafren. 1985;72:212-3.

[40] Ikura M. Calcium binding conformational response in EFhand proteins. Trends Biochem Sci. 1996;21:14-7.

[41] Kawasaki H, Kretsinger RH. Calcium-binding proteins. 1: EFhands. Protein Profile. 1994;1:343-517.

[42] Martínez J, Sánchez R, Castellanos M, Fernández-Escamillac AM, et al. Fish β-parvalbumin acquires allergenic properties by amyloid assembly. Swiss Med Wkly. 2015;145:1-11.

[43] Sánchez R, Martínez J, Castro A, Pedrosa M., et al. The amyloid fold of Gad m 1 epitopes governs IgE binding. Scientific Reports. 2016;6:32801.

[44] Akdis CA. Therapies for allergic inflammation: refining strategies to induce tolerance. Nat Med. 2012;18:736-49.

[45] Klimek L, Schmidt-Weber CB, Kramer MF, Skinner MA, Heath MD. Clinical use of adjuvants in allergen-immunotherapy. Expert Rev Clin Immunol. 2017;13:599-610.

[46] Focke-Tejkl M, Weber M, Niespodziana K, Neubauer A, et al. Development and characterization of a recombinant, hypoallergenic, peptide-based vaccine for grass pollen allergy. J Allergy Clin Immunol. 2015;135:1207-17.

[47] Zieglmayer P, Focke-Tejkl M, Schmutz R, Lemell P, et al. Mechanisms, safety and efficacy of a B cell epitope-based vaccine for immunotherapy of grass pollen allergy. EBioMedicine. 2016;11:43-57.

[48] Pfaar O, Barth C, Jaschke C. Hörmann, K., Klimek, L., Sublingual allergen-specific immunotherapy adjuvanted with monophosphoryl lipid A: a phase I/IIa study. Int Arch Allergy Immunol. 2011;154:336-44.

[49] Scheiblhofer S, Machado Y, Feinle A, Thalhamer J, et al. Potential of nanoparticles for allergen-specific immunotherapy - use of silica nanoparticles as vaccination platform. Expert Opin Drug Deliv. 2016;12:1777-88.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE: This study was carried out in agreement with the guidelines and procedures established by the CEIC (Ethical Committee for Clinical Research) of Galicia Spain, with a written informed consent from all subjects. This informed consent was in accordance with the Declaration of Helsinki. The protocol was approved by the CEIC of Galicia, Spain.

CONSENT FOR PUBLICATION: Informed consent from all subjects for publication were provided and all protocols were approved by the board of the local institutional authority (CEIC of Galicia, Spain).

AVAILABILITY OF DATA AND MATERIALS: Data sharing not applicable to this article as no data-sets were generated or analyzed during the current study

ACKNOWLEDGEMENTS: The authors also wish to express their gratitude to Lorena Barros for her excellent technical assistance and her support in the collection of fish species used in this study. The authors also wish to express their gratitude to med. Dr. Miguel González from Hospital Carlos III of Madrid, Spain for his support in collecting the serum samples from healthy and fish allergic patients used in the present study and to Dr. Luiz Stark from the Immunology group at the University of Vigo for his excellent scientific recommendations. This work was supported by grants from the EU Marie Curie actions (FP7-PEOPLE-2012-IEF, ref. 332274), by the Ramón Areces Foundation (XVII National grant) and by the GAIN-Xunta de Galicia project (IN607D 2017/01). MC is supported by the Ramón y Cajal contract (MINECO).

Figure 1

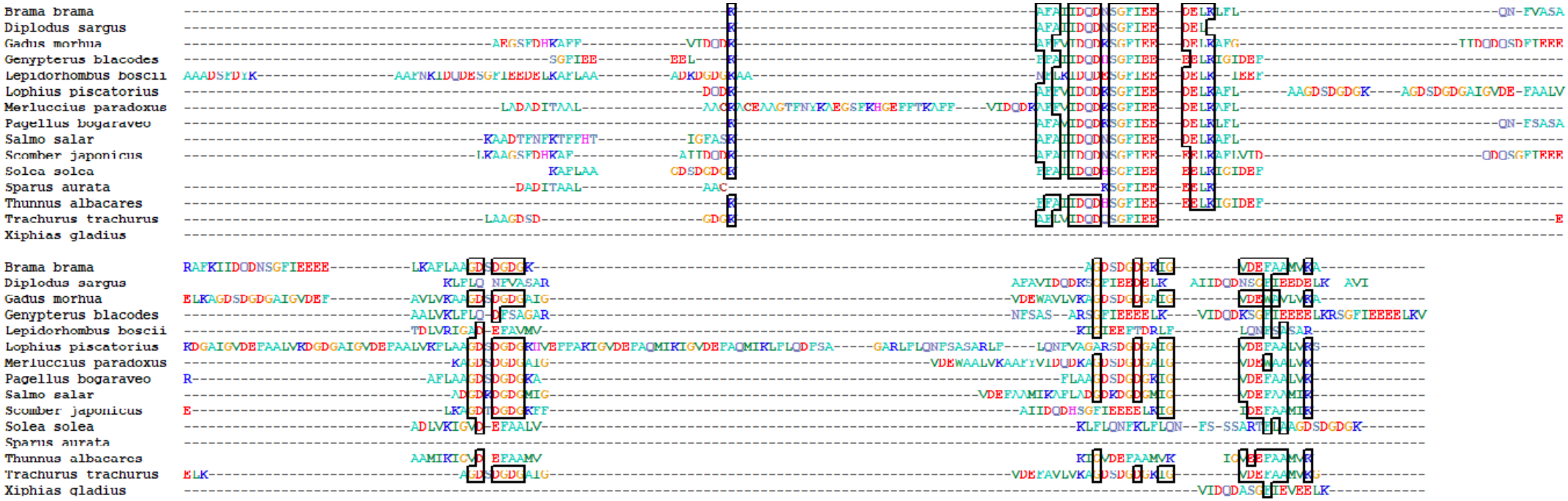


Figure 2

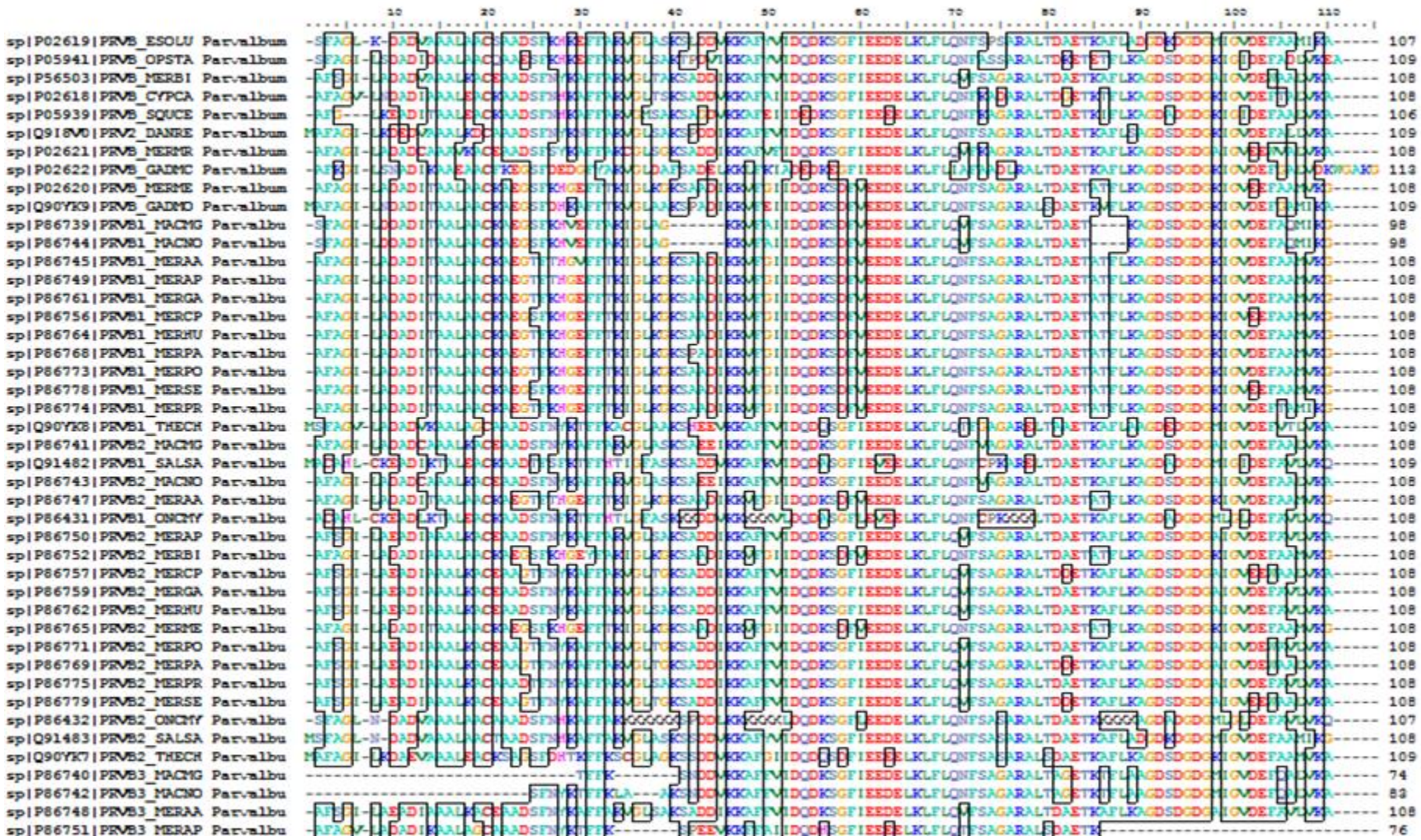


Figure 2. Continuation.

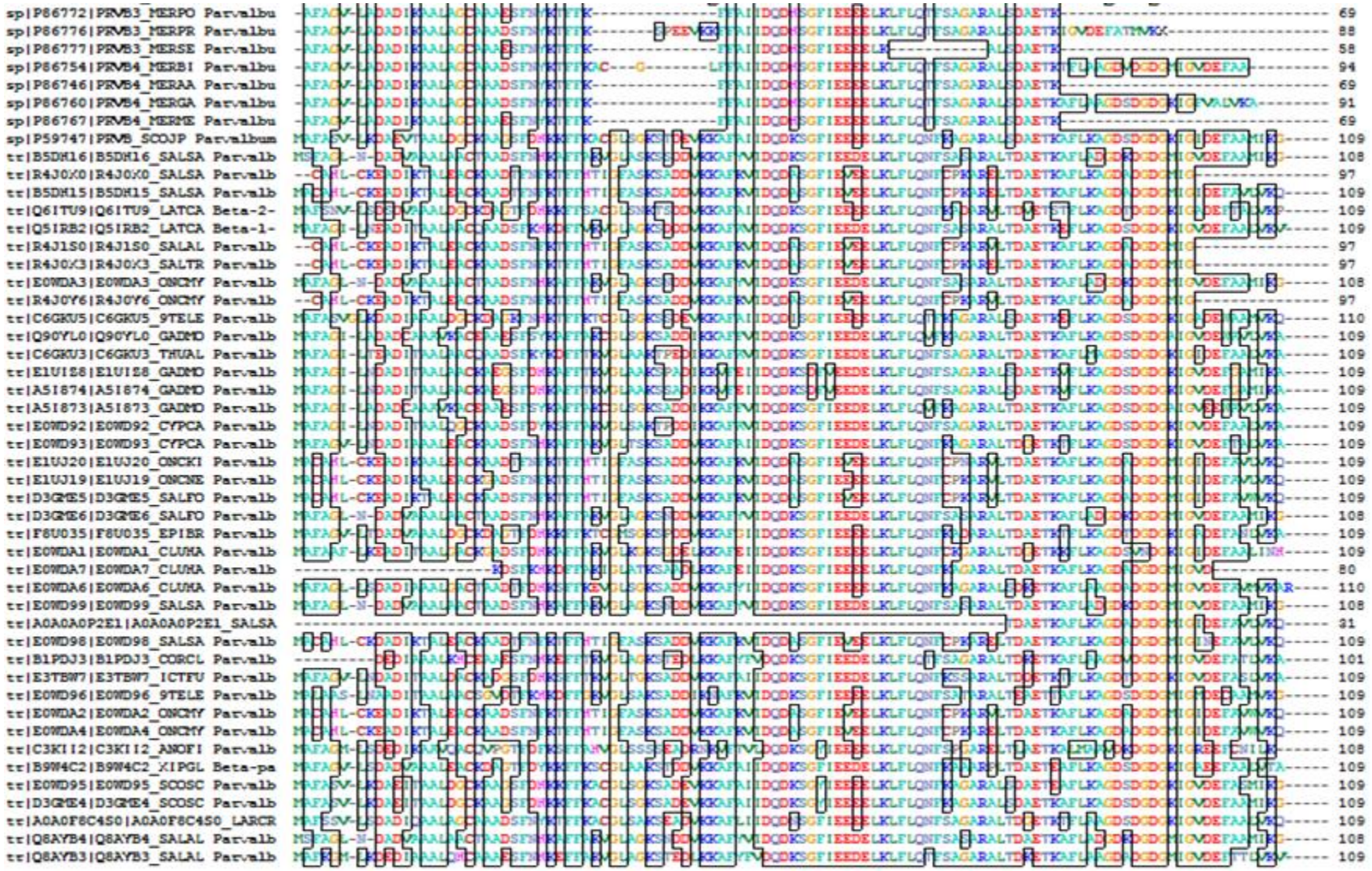


Figure 3

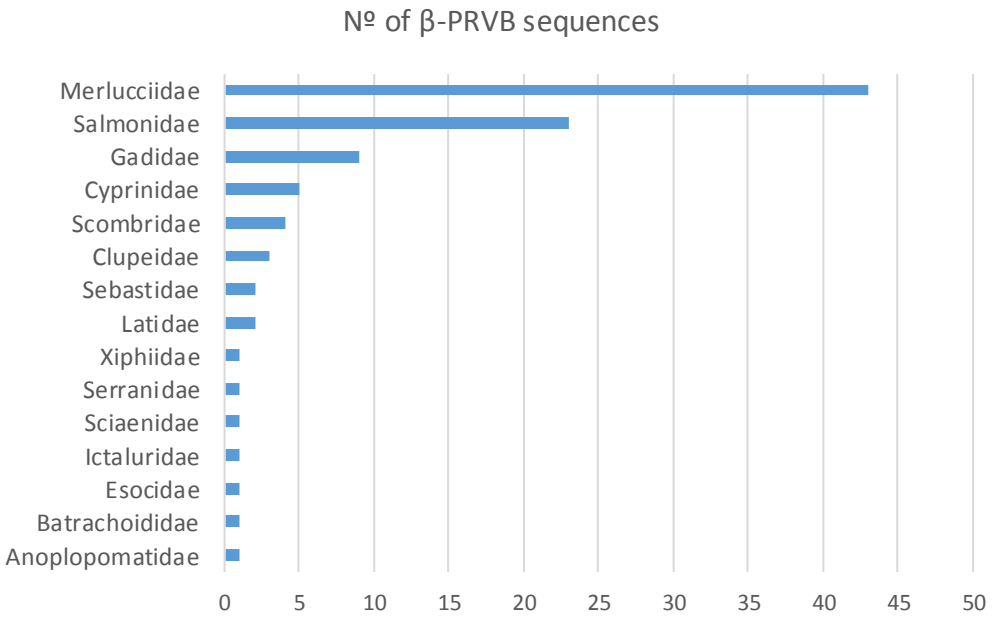


Figure 4

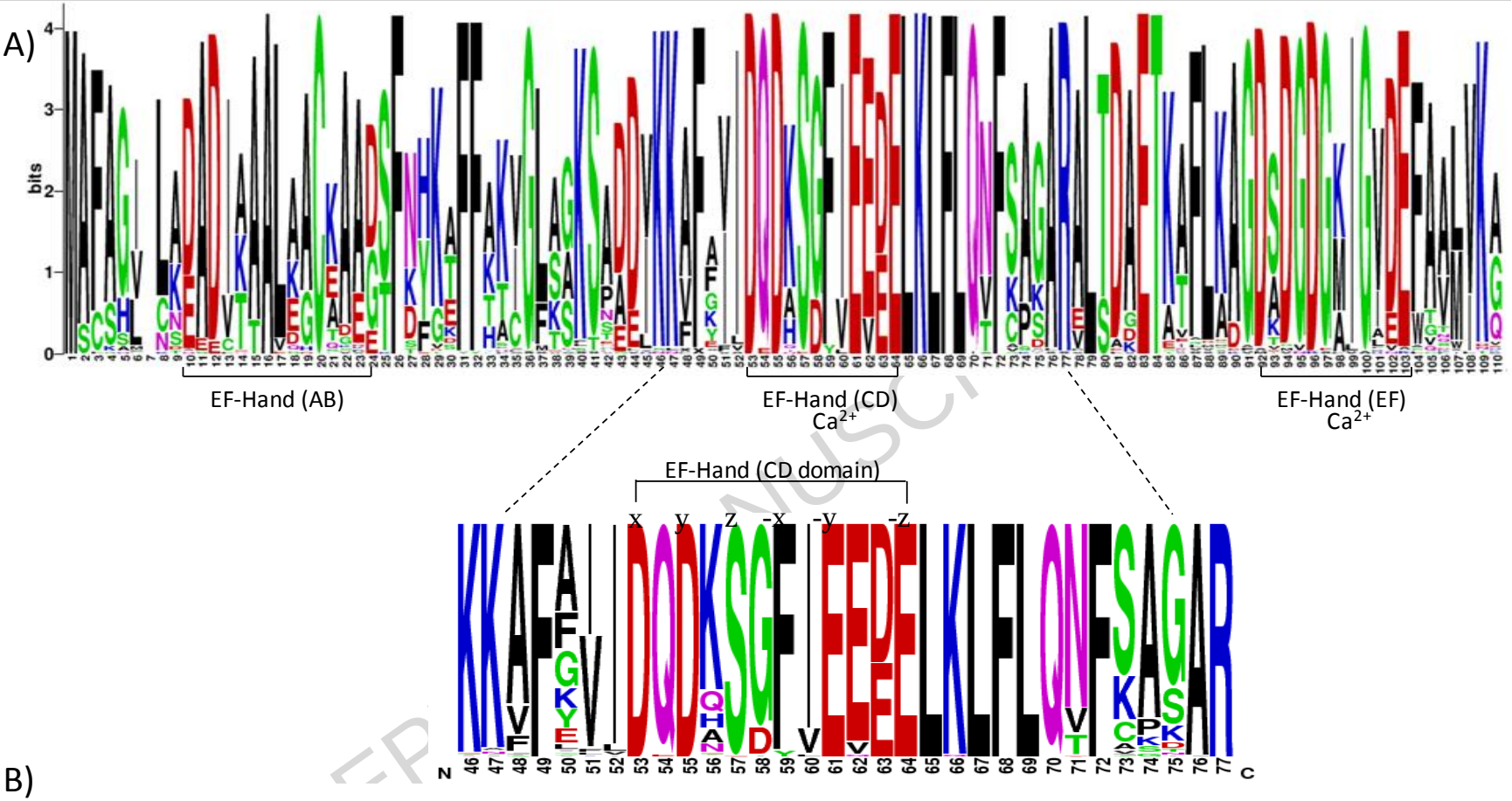


Figure 5

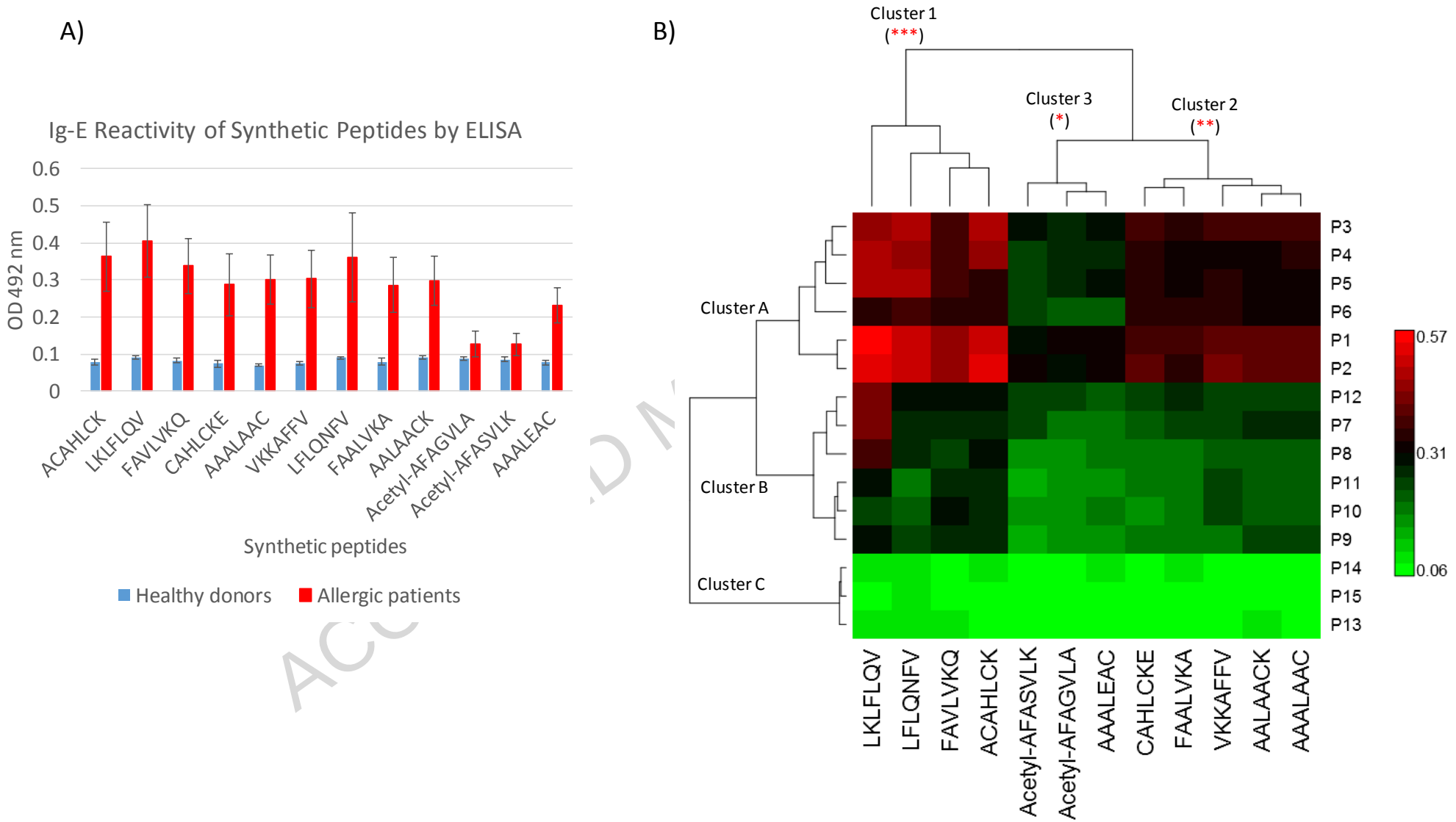


Figure 6

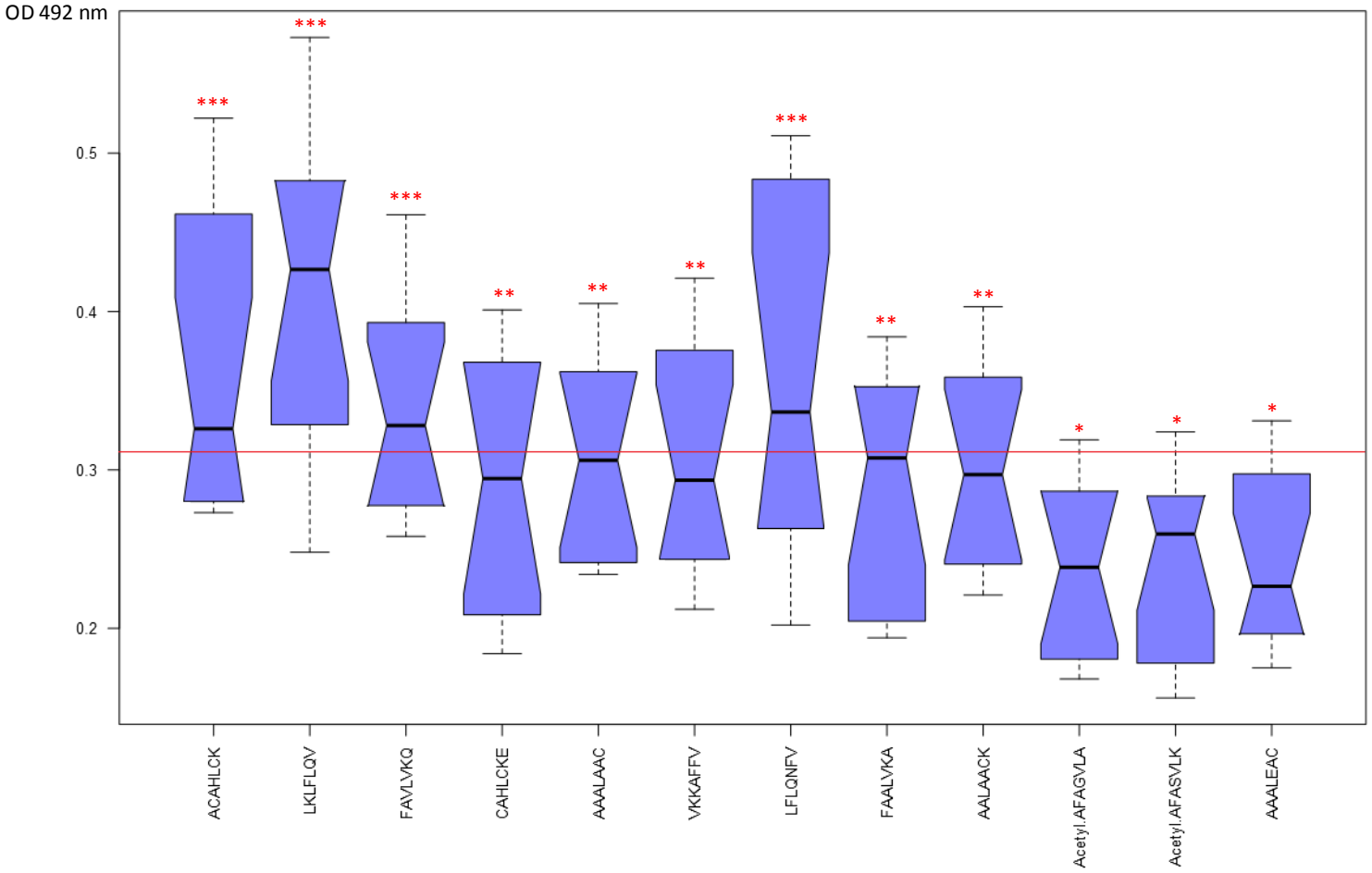


Figure 7

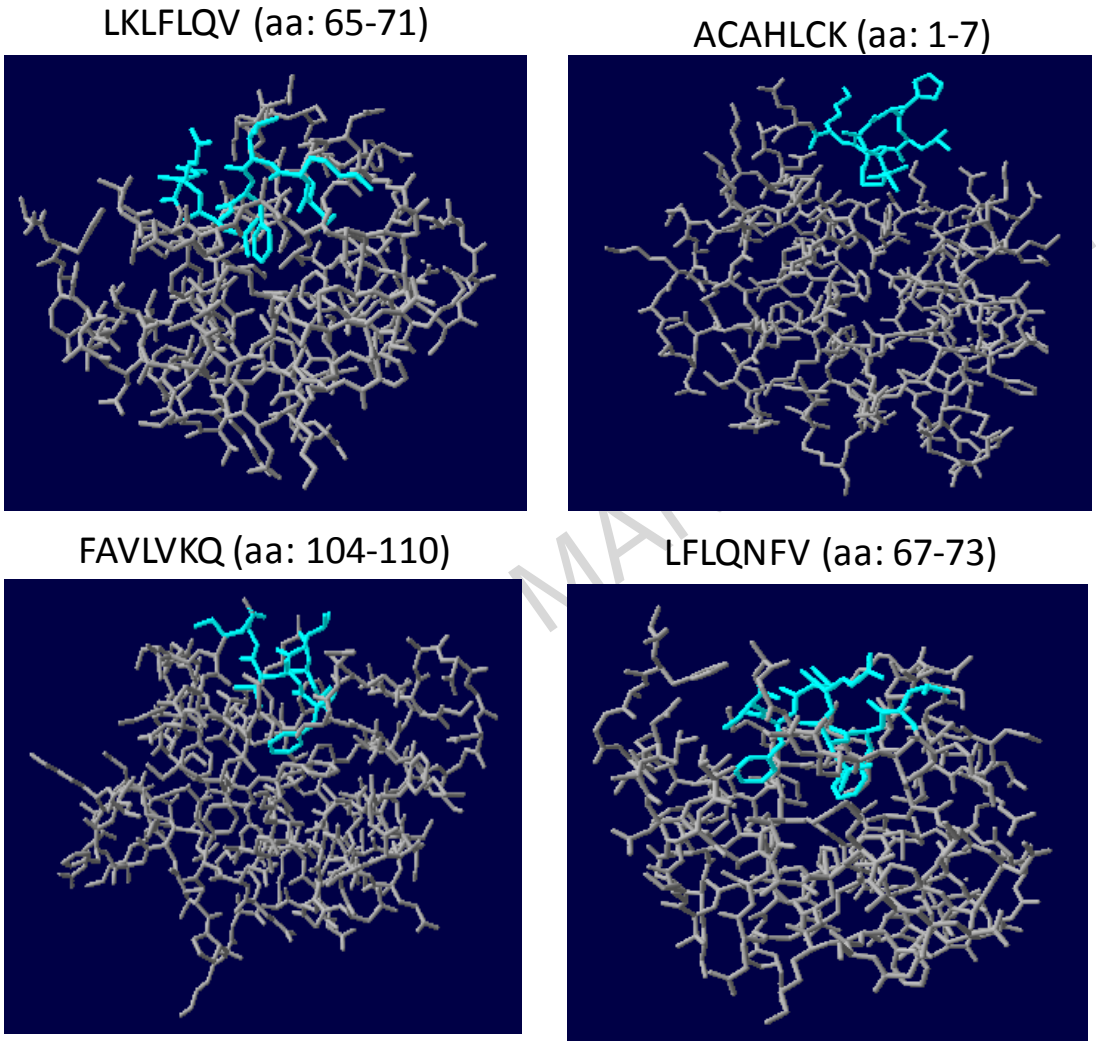


FIGURE CAPTIONS:

Figure 1: Alignment by Clustal W of all β-PRVBs peptides identified by shotgun proteomics for each of the 15 fish species analyzed.

Figure 2: Alignment by Clustal W of all β-PRVBs registered in the UniProtKB database (n= 98).

Figure 3: Number of β-PRVBs sequences analyzed in the present work according to the Teleostei family.

Figure 4: A) Highly conserved regions of the all Teleostei β-PRVB sequences obtained by shotgun proteomics and downloaded from the UniProtKB database. B) Linear B-cell epitope regions identified for several β-PRVBs using the Kolaskar and Tongaonkar algorithm on the IEDB database.

Figure 5: A) Analysis by indirect ELISA of the reactivity of the Ig-E of 12 fish allergic patients and 3 healthy donors against 12 synthetic β-PRVB peptides. B) Heatmap of the Euclidean hierarchical clustering of the ELISA reactivity of the Ig-E of 12 fish allergic patients and 3 healthy donors (rows) against 12 synthetic β-PRVB peptides (columns).

Figure 6: Box-plot of the ELISA reactivity of the Ig-E of 12 fish allergic patients against 12 synthetic β-PRVB peptides.

Figure 7: 3D model structure of β-PRVBs to represent the four B-cell peptide vaccine candidates (LKLFLQV, ACAHLCK, FAVLVKQ and LFLQNFV) using the Swiss-Model and Swiss-PdbViewer programs. As protein templates were used, for LKLFLQV (P56503, PRVB_MERBI); for ACAHLCK (Q91482, PRVB1_SALSA), for FAVLVKQ (Q91482, PRVB1_SALSA) and for LFLQNFV (P86741, PRVB2_MACMG).

Table 1. Fish species considered in the study.

Sample	Species (Order)	Common name
S1	*Brama brama* (Perciformes)	Ray's bream
S2	*Diplodus sargus* (Perciformes)	White seabream
S3	*Gadus morhua* (Gadiformes)	Cod
S4	*Genypterus blacodes* (Ophidiiformes)	Pink cusk-eel
S5	*Lepidorhombus boscii* (Pleuronectiformes)	Four-spot megrim
S6	*Lophius piscatorius* (Lophiiformes)	Angler
S7	*Merluccius paradoxus* (Gadiformes)	Deep-cape hake
S8	*Pagellus bogaraveo* (Perciformes)	Common seabream
S9	*Salmo salar* (Salmoniformes)	Salmon
S10	*Scomber japonicus* (Perciformes)	Club mackerel
S11	*Solea solea* (Pleuronectiformes)	Common sole
S12	*Sparus aurata* (Perciformes)	Gilthead seabream
S13	*Thunnus albacares* (Perciformes)	Yellowfin tuna
S14	*Trachurus trachurus* (Perciformes)	Horse mackerel
S15	*Xiphias gladius* (Perciformes)	Swordfish

Table 2. Serum samples from fish allergic patients and healthy donors.

Sample	Sex	Age (years)	History of a Type-I Fish Allergy	Anti-cod IgE levels (kU/L)
Allergic patients				
P1	f	11	+	46.6
P2	m	3	+	32.4
P3	m	2	+	28.4
P4	m	1	+	23.8
P5	m	3	+	22.7
P6	m	34	+	22.4
P7	m	2	+	19.3
P8	m	9	+	14.3
P9	m	1	+	5.38
P10	m	8	+	4.99
P11	f	9	+	4.88
P12	f	10	+	3.49
Healthy donors				
P13	f	36	-	-
P14	f	28	-	-
P15	m	38	-	-

kU/L: kilo units/liter; f: female; m: male

Table 3. Synthetic B-cell peptide epitopes.

Peptide epitope	Sequence	% Highest IEDB score for binding to all β-PRVBs	Amount (mg)	Purity (%)
1	ACAHLCK	1.177	2.3	76.5
2	LKLFLQV	1.167	2.2	79.4
3	FAVLVKQ	1.159	2.1	76.9
4	CAHLCKE	1.146	2.2	73.3
5	AAALAAC	1.140	2.5	69.8
6	VKKAFFV	1.125	2.4	86.2
7	LFLQNFV	1.122	2.0	70.1
8	FAALVKA	1.121	2.1	89.1
9	AALAACK	1.121	2.1	81.5
10	Acetyl-AFAGVLA	1.113	2.6	83.9
11	Acetyl-AFASVLK	1.113	2.1	76.7
12	AAALEAC	1.110	2.4	72.4

Table 4. Shotgun proteomics results of β-PRVBs proteins and peptides identified for each of the different 15 fish species.

Sample	Fish species	Protein Description	UniProtKB Accession	Gene	Peptide Counts	Non-redundant Peptides	Sequence Coverage (%)
S1	*Brama brama*	Parvalbumin-like protein [*Sparus aurata*]	Q4QY67	N/A	70	19	69
S2	*Diplodus sargus*	Parvalbumin-like protein [*Sparus aurata*]	Q4QY67	N/A	141	14	60
S3	*Gadus morhua*	parvalbumin beta, allergen Gad c 1 [*Gadus morhua*]	P02622	N/A	1331	33	52
S4	*Genypterus blacodes*	Parvalbumin beta 3 [*Merluccius hubbsi*]	P86763	N/A	484	9	65
S5	*Lepidorhombus boscii*	parvalbumin [*Lepidorhombus whiffiagonis*]	B5WX08	pvalb	558	16	57
S6	*Lophius piscatorius*	probable calcium-binding protein CML20 [*Esox lucius*]	N/A	N/A	213	17	40
S7	*Merluccius paradoxus*	Parvalbumin beta 1 [*Merluccius paradoxus*]	P86768	N/A	817	104	54
S8	*Pagellus bogaraveo*	parvalbumin beta [*Kryptolebias marmoratus*]	N/A	N/A	414	46	82
S9	*Salmo salar*	parvalbumin beta 1 [*Salmo salar*]	B5DH15	PRVB1	879	35	59
S10	*Scomber japonicus*	major fish allergen parvalbumin Scoj1 [*Scomber japonicus*]	P59747	N/A	403	14	49
S11	*Solea solea*	parvalbumin [*Astyanax mexicanus*]	N/A	N/A	214	8	49
S12	*Sparus aurata*	Parvalbumin [*Sparus aurata*]	D0VB96	N/A	31	2	35

S13	*Thunnus albacares*	Parvalbumin beta [*Thunnus albacares*]	C6GKU3	pvalb1	57	10	48
S14	*Trachurus trachurus*	Parvalbumin [*Fundulus similis*]	C0LEL7	N/A	896	21	64
S15	*Xiphias gladius*	parvalbumin beta 1-like [*Xiphias gladius*]	B9W4C2	pvalb	163	1	75

Table 5. Predicted B-cell epitopes after the application of the Kolaskar and Tongaonkar algorithm on the IEDB database.

β-PRVB	B-cell Epitope	% Highest score for binding to all β-PRVBs	β-PRVB	B-cell Epitope	% Highest score for binding to all β-PRVBs
P02619|PRVB_ESOLU	AAALAC	1.14	P86748|PRVB3_MERAA	LKLFLQV	1.167
P05941|PRVB_OPSTA	AALAACQ	1.133	P86751|PRVB3_MERAP	AFAGVLA	1.113
P56503|PRVB_MERBI	LKLFLQV	1.167	P86753|PRVB3_MERBI	LKLFLQV	1.167
P02618|PRVB_CYPCA	AAALEAC	1.11	P86755|PRVB3_MERCP	AFAGVLA	1.113
P05939|PRVB_SQUCE	FAALVKA	1.121	P86758|PRVB3_MERGA	LKLFLQV	1.167
Q9I8V0|PRV2_DANRE	FALLVKA	1.147	P86763|PRVB3_MERHU	KVALVKA	1.143
P02621|PRVB_MERMR	LKLFLQV	1.167	P86766|PRVB3_MERME	LKLFLQV	1.167
P02622|PRVB_GADMC	LKLFLIA	1.141	P86770|PRVB3_MERPA	FVALVKA	1.166
P02620|PRVB_MERME	AALAACK	1.121	P86772|PRVB3_MERPO	AFAGVLA	1.113
Q90YK9|PRVB_GADMO	AALAACK	1.121	P86776|PRVB3_MERPR	AFAGVLA	1.113
P86739|PRVB1_MACMG	LKLFLQV	1.167	P86777|PRVB3_MERSE	AFAGVLA	1.113
P86744|PRVB1_MACNO	LKLFLQV	1.167	P86754|PRVB4_MERBI	CGLFFAI	1.133
P86745|PRVB1_MERAA	AALAACK	1.121	P86746|PRVB4_MERAA	AFAGVLA	1.113
P86749|PRVB1_MERAP	AALAACK	1.121	P86760|PRVB4_MERGA	IGFVALV	1.171
P86761|PRVB1_MERGA	AALAACK	1.121	P86767|PRVB4_MERME	AFAGVLA	1.113
P86756|PRVB1_MERCP	AALAACK	1.121	P59747|PRVB_SCOJP	AFASVLK	1.113
P86764|PRVB1_MERHU	AALAACK	1.121	B5DH16|B5DH16_SALSA	AAALAAC	1.14
P86768|PRVB1_MERPA	AALAACK	1.121	R4J0X0|R4J0X0_SALSA	CAHLCKE	1.146
P86773|PRVB1_MERPO	AALAACK	1.121	B5DH15|B5DH15_SALSA	ACAHLCK	1.177
P86778|PRVB1_MERSE	AALAACK	1.121	Q6ITU9|Q6ITU9_LATCA	FFSACGL	1.113
P86774|PRVB1_MERPR	AALAACK	1.121	Q5IRB2|Q5IRB2_LATCA	FAALVKV	1.166
Q90YK8|PRVB1_THECH	FVTLVKA	1.144	R4J1S0|R4J1S0_SALAL	CAHLCKE	1.146

P86741|PRVB2_MACMG	LFLQNFV	1.122		R4J0X3|R4J0X3_SALTR	CAHLCKE	1.146
Q91482|PRVB1_SALSA	ACAHLCK	1.177		E0WDA3|E0WDA3_ONCMY	AAALAAC	1.14
P86743|PRVB2_MACNO	LFLQNFV	1.122		R4J0Y6|R4J0Y6_ONCMY	CAHLCKE	1.146
P86747|PRVB2_MERAA	AALAACK	1.121		C6GKU5|C6GKU5_9TELE	AFASVGL	1.105
P86431|PRVB1_ONCMY	ACAHLCK	1.177		Q90YL0|Q90YL0_GADMO	LKLFLQV	1.167
P86750|PRVB2_MERAP	LKLFLQV	1.167		C6GKU3|C6GKU3_THUAL	AALAACQ	1.133
P86752|PRVB2_MERBI	AAALAAC	1.14		E1UIZ8|E1UIZ8_GADMO	AALAACK	1.121
P86757|PRVB2_MERCP	LKLFLQV	1.167		A5I874|A5I874_GADMO	AALAACK	1.121
P86759|PRVB2_MERGA	LKLFLQV	1.167		A5I873|A5I873_GADMO	LKLFLQV	1.167
P86762|PRVB2_MERHU	LKLFLQV	1.167		E0WD92|E0WD92_CYPCA	FAALVKA	1.121
P86765|PRVB2_MERME	AALAACK	1.121		E0WD93|E0WD93_CYPCA	AAALEAC	1.11
P86771|PRVB2_MERPO	LKLFLQV	1.167		E1UJ20|E1UJ20_ONCKI	ACAHLCK	1.177
P86769|PRVB2_MERPA	LKLFLQV	1.167		E1UJ19|E1UJ19_ONCNE	ACAHLCK	1.177
P86775|PRVB2_MERPR	LKLFLQV	1.167		D3GME5|D3GME5_SALFO	ACAHLCK	1.177
P86779|PRVB2_MERSE	LKLFLQV	1.167		D3GME6|D3GME6_SALFO	AAALAAC	1.14
P86432|PRVB2_ONCMY	FAVLVKQ	1.159		F8U035|F8U035_EPIBR	ADVAAAL	1.108
Q91483|PRVB2_SALSA	AAALAAC	1.14		E0WDA1|E0WDA1_CLUHA	LFLQNFC	1.126
Q90YK7|PRVB2_THECH	AAALEAC	1.11		E0WDA7|E0WDA7_CLUHA	ELKLFLQ	1.091
P86740|PRVB3_MACMG	VKKAFFV	1.125		E0WDA6|E0WDA6_CLUHA	AAALGAC	1.113
P86742|PRVB3_MACNO	VKKAFFV	1.125		E0WD99|E0WD99_SALSA	AAALAAC	1.14
E3TBW7|E3TBW7_ICTFU	FASLVKA	1.113		A0A0A0P2E1_SALSA	FAVLVKQ	1.159
E0WD96|E0WD96_9TELE	LAACSGV	1.151		E0WD98|E0WD98_SALSA	ACAHLCK	1.177
E0WDA2|E0WDA2_ONCMY	ACAHLCK	1.177		B1PDJ3|B1PDJ3_CORCL	AAALKHC	1.127
E0WDA4|E0WDA4_ONCMY	ACAHLCK	1.177		C3KII2|C3KII2_ANOFI	AVQACQV	1.191
E0WD95|E0WD95_SCOSC	AFASVLK	1.113		B9W4C2|B9W4C2_XIPGL	VLSDADV	1.118
A0A0F8C4S0_LARCR	AFSSVLS	1.118		D3GME4|D3GME4_SCOSC	AFASVLK	1.113
Q8AYB4|Q8AYB4_SALAL	AAALAAC	1.14		Q8AYB3|Q8AYB3_SALAL	AAALQHC	1.139

Table 6. Validation of B-cell epitopes using ELISA assay against sera (IgE-binding) from healthy and fish allergic patients

			Allergic patients (ELISA OD 492 nm)													
B-cell Peptide Epitope	**Sequence**	**% Highest IEDB to all β-PRVBs**	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**Mean**	**Stand. Desv. (SD)**
1	ACAHLCK	1.177	0.514	0.494	0.464	0.440	0.368	0.386	0.302	0.286	0.276	0.258	0.288	0.289	0.3637	0.0934
2	LKLFLQV	1.167	0.568	0.515	0.457	0.463	0.486	0.362	0.409	0.347	0.304	0.242	0.286	0.423	0.4051	0.0989
3	FAVLVKQ	1.159	0.456	0.447	0.402	0.354	0.397	0.361	0.264	0.246	0.265	0.287	0.289	0.291	0.3382	0.0742
4	CAHLCKE	1.146	0.378	0.398	0.367	0.361	0.348	0.340	0.234	0.209	0.198	0.183	0.196	0.235	0.2872	0.0840
5	AAALAAC	1.140	0.392	0.398	0.368	0.351	0.323	0.321	0.278	0.231	0.243	0.232	0.230	0.238	0.3004	0.0661
6	VKKAFFV	1.125	0.405	0.417	0.386	0.322	0.353	0.352	0.257	0.221	0.208	0.247	0.236	0.237	0.3034	0.0771
7	LFLQNFV	1.122	0.521	0.512	0.486	0.462	0.438	0.376	0.269	0.298	0.250	0.223	0.198	0.298	0.3609	0.1185
8	FAALVKA	1.121	0.382	0.361	0.352	0.340	0.331	0.348	0.266	0.201	0.199	0.196	0.192	0.267	0.2862	0.0741
9	AALAACK	1.121	0.397	0.392	0.376	0.329	0.320	0.316	0.267	0.223	0.245	0.222	0.226	0.254	0.2972	0.0665
10	Acetyl-AFAGVLA	1.113	0.128	0.146	0.183	0.183	0.155	0.133	0.100	0.097	0.098	0.083	0.096	0.121	0.1269	0.0341
11	Acetyl-AFASVLK	1.113	0.114	0.128	0.202	0.165	0.128	0.126	0.125	0.084	0.102	0.123	0.107	0.106	0.1258	0.0309
12	AAALEAC	1.110	0.283	0.291	0.281	0.276	0.285	0.231	0.205	0.187	0.182	0.185	0.168	0.202	0.2313	0.0483

		Healthy donors (ELISA OD 492 nm)					
B-cell Peptide Epitope	**Sequence**	**% Highest IEDB to all β-PRVBs**	**P13**	**P14**	**P15**	**Mean**	**SD**
1	ACAHLCK	1.177	0.072	0.086	0.074	0.0773	0.0075
2	LKLFLQV	1.167	0.096	0.092	0.085	0.0910	0.0055
3	FAVLVKQ	1.159	0.087	0.084	0.076	0.0823	0.0056
4	CAHLCKE	1.146	0.083	0.064	0.074	0.0736	0.0095
5	AAALAAC	1.140	0.065	0.072	0.071	0.0693	0.0037
6	VKKAFFV	1.125	0.073	0.080	0.072	0.0750	0.0043
7	LFLQNFV	1.122	0.094	0.087	0.089	0.0900	0.0036

8	FAALVKA	1.121	0.075	0.090	0.072	0.0790	0.0096
9	AALAACK	1.121	0.091	0.093	0.084	0.0893	0.0047
10	Acetyl-AFAGVLA	1.113	0.087	0.085	0.093	0.0883	0.0041
11	Acetyl-AFASVLK	1.113	0.092	0.081	0.083	0.0853	0.0058
12	AAALEAC	1.110	0.074	0.084	0.074	0.0773	0.0057

Table 7. Potential B-cell peptide vaccine candidates for fish allergy.

B-cell Epitope	Residues (aa to aa)	% Highest IEDB score for binding to all β-PRVBs	ELISA (mean of Ig-E binding of 12 allergic patients)	Heatmap of ELISA reactivity of 12 allergic patients	Boxplot of ELISA reactivity of 12 allergic patients
LKLFLQV	65-71	1.167	0.4051	(***)	(***)
ACAHLCK	1-7	1.177	0.3637	(***)	(***)
FAVLVKQ	104-110	1.159	0.3382	(***)	(***)
LFLQNFV	67-73	1.121	0.3609	(***)	(***)

Mónica Carrera,[1*] África González-Fernández,[2] Susana Magadán,[2] Jesús Mateos, [1] Lelia Pedrós, [3] Isabel Medina, [1] José M. Gallardo,[1]

[1]Spanish National Research Council (CSIC), Marine Research Institute (IIM), Food Technology, Vigo, Pontevedra, Spain.

[2]Biomedical Research Center (CINBIO) (Centro Singular de Investigación de Galicia), University of Vigo, Immunology, Vigo, Pontevedra, Spain.

[3]Hospital Meixoeiro of Vigo, Galicia Sur Health Research Institute, Complexo Hospitalario Universitario de Vigo (CHUVI), SERGAS, Vigo, Pontevedra, Spain

DECLARATION BY THE AUTHORS:

CONFLICT OF INTEREST: The authors declare that they have no conflict of interests

HIGLIGHTS:

- Shotgun Proteomics and Protein-based Bioinformatics for B-cell epitopes of the major fish allergen
- IgE-reactivity of B-cell epitopes using synthetic peptides and sera of healthy and allergic patients.
- Design of potential peptide vaccine for fish allergy