Characterization, Epitope Confirmation, and Cross-reactivity Analysis of Parvalbumin from *Lateolabrax maculatus* by Multi-omics Technologies

Qing Liu [a], Zengying Sui [a], Nuan Feng [b], Yuhao Huang [c], Yonghong Li [c], Ishfaq Ahmed [c], Thimo Ruethers [d, e], Hui Liang [a], Zhenxing Li [c], Andreas L. Lopata [d, e], Lirui Sun [a*]

[a] *Department of Nutrition and Food Hygiene, School of Public Health, Qingdao University, Qingdao, 266071, China.*

[b] *Department of Nutrition, Qingdao Women and Children's Hospital, Qingdao, 266034, China.*

[c] *College of Food Science and Engineering, Ocean University of China, 266003, China.*

[d] *Molecular Allergy Research Laboratory, College of Public Health, Medical and Veterinary Sciences, Australian Institute of Tropical Health and Medicine, James Cook University, Townsville, QLD, Australia;*

[e] *Tropical Futures Institute, James Cook University, Singapore*

Corresponding author:

Dr. Lirui Sun

*E-mail: sunlirui@qdu.edu.cn

Tel.: +86-532-83812434

Abstract:

Spotted seabass (*Lateolabrax maculatus*) is the second largest maricultural fish species in China and the main trigger of food-related allergic reactions. Nevertheless, studies on allergens of *L. maculatus* are limited. In this study, two proteins of about 11 kDa were purified and confirmed as parvalbumins by mass spectrometry. The IgG and IgE-binding activities were evaluated through immunoblotting assays, respectively. The molecular characteristics of β-parvalbumin were investigated by combining proteomics, genomics, and immunoinformatics approaches. The results indicated that β-parvalbumin consists of 109 amino acids with a molecular weight of 11.5 kDa, and is the major allergen displaying strong IgE binding capacity. In silico analysis identified seven linear B cell epitopes distributed mainly on α-helixes and the calcium-binding loops. In addition, the cross-reactivity among 26 commonly consumed fish species was analyzed. The in-house generated anti-*L. maculatus* parvalbumin polyclonal antibody recognized 100% of the 26 fish species, demonstrating cross-reactivity and better binding capacity than the commercial anti-cod parvalbumin antibody. Together, this study provides an efficient protocol to characterize allergens with the aid of multi-omics methods and supports parvalbumin from *L. maculatus* as a candidate for fish allergen determination and allergy diagnosis.

Keywords: parvalbumin, epitopes, cross-reactivity, multi-omics, *Lateolabrax maculatus,* sea bass, fish allergy

1. Introduction

Food allergy, defined as an abnormal adverse immune response to a specific food protein, is now increasingly common in the population[1]. IgE-mediated allergic reactions usually appear within minutes to hours after ingestion of opposing food. It is estimated that up to 10% of the population is suffering from food allergy[2]. According to the big-9 food allergens proposed by FAO/WHO, seafood allergy primarily induced by fish and crustacean products is more common than others in Asia[3]. The prevalence of seafood allergy is increasing globally, ranging from 2.5% to 5% of the general population worldwide[3,4].

Fish is an essential dietary source of proteins, polyunsaturated fatty acids, lipids, and vitamins. However, it is also one of the most common sources of food allergy, causing diarrhea, vomiting, rhinitis, urticaria, and even life-threatening anaphylaxis[5]. Fish hypersensitivity is typically known to persist lifelong[6] and affects approximately 0.20% - 2.29% of the global population varying substantially across regions[7,8]. Furthermore, a higher prevalence of up to 8% has been reported among fish-processing workers [3,9]. Among countries with high seafood consumption, fish is a major cause of food allergy[10]. One meta-analysis on the prevalence of food allergy in China reported that the fish allergy accounts for 15%, while slightly below the shrimp allergy (16%)[11]. The advised precautionary measures for fish allergy is to ensure the total exclusion of fish and its

products[12,13]. Although multi-allergen labeling legislations had been established, partially ambiguous and precautionary labels still mislead consumers[14,15]. Besides, restricting fish intake leads to decreased nutritional diversity and lower quality of life. Voluntary Incidental Trace Allergen Labelling (VITAL) developed by Australian and New Zealand Foods provided a standardized and scientific reference dose of fish proteins at 27.3 mg, which likely provokes anaphylaxis in 10% of the allergic populations[15]. From the 2011–2018 data collection (completed data collection on a cohort of 227 patients), the minimum induced dose of fish protein ED01 was estimated at 2.6 mg, less than this dose may not cause allergic symptoms, and the estimated ED05 was 12.1 mg, but individual specificity remained[16]. The identification, characterization, and quantification of fish allergens are fundamentally carried out to guarantee the threshold dose application.

Parvalbumin (PV), tropomyosin (TM), aldolase A, collagen, triose phosphate isomerase (TIM), creatine kinase, Pyruvate kinase PKM-like, L-lactate dehydrogenase, Glucose 6-phosphate isomerase，Glyceraldehyde-3-phosphate dehydrogenase, and β-enolase have been reported as the fish allergy triggers by the World Health Organization (WHO)/International Union of Immunological Societies (IUIS) Allergen Nomenclature Sub Committee (www.allergen.org). PV is recognized as an pan-allergen that triggers at least 95% of IgE-mediated reactions in fish-

allergic individuals[17]. PV belongs to Ca^{2+} binding proteins of the EF-hand superfamily and is a small protein with a molecular weight of 10–15 kDa and an acidic pI (pH 3–5)[18,19]. Fish PVs have several isoforms (up to 5), varying in amino acids composition. For instance, four isoforms are recognized in *Gadus morhua* (Gad m1), three isoforms in *Paralichthys olivaceus*, and two isoforms in *Cyprinus carpio* (Cyp c1)[20,21]. Among them, β-PV in bony fish is the major isoform with stronger IgE binding capacity as compared to α-PV, primarily found in sharks, and plays a role in allergic cross-reactivity among different types of fish[22]. Hence, β-PV is commonly extracted, purified and expressed recombinant for fish allergy diagnosis and fish allergen detection.

β-PVs from various fish species share amino acid sequence identities ranging from 61% to 82%[5]. Discrepancy in epitopes limits the utilization of isolated PV in commercial applications. Fish PVs with similar epitopes show higher allergic cross-reactivity. Well-characterized cross-reactive epitopes between fish PVs will assist in developing a more accurate component-resolved diagnosis and contribute to the development of fish allergen determination. Limited studies have mapped epitopes in some fish PVs, such as Lat c1, Gad m1, Sal s1, and Cyp c1, which had been included in the Immune Epitope Database (IEDB)[23–26]. The most common B-cell epitopes were supposed to be located in highly conserved protein regions, especially near the ion-binding region[27,28]. However, the traditional epitope

mapping technologies, such as overlapping peptides and bioinformatics, didn’t consider the protease digestion of PVs.

Spotted seabass (*Lateolabrax maculatus*) belongs to the family of Lateolabracidae (order Perciformes) and is mainly farmed in East Asia[29]. Spotted seabass is the second largest maricultural fish species after the large yellow croaker (*Larimichthys crocea*) in China, with an annual production exceeding 218,000 tons according to the China Fishery Statistical Yearbook (2023)[30]. Nevertheless, there is a lack of understanding on the molecular characteristics, IgG/IgE binding epitopes, and cross-reactivity of spotted seabass allergens.

In this study, we aimed to purify and characterize PVs in *L. maculatus* with a multi-omics approach, assess the antigenicity and potential allergenicity, evaluate the cross-reactivity with other common fish species in China, and further map the linear B-cell epitopes of β-PV using bioinformatics prediction and subsequent confirmation using peptides in immunological assays.

2 Materials and methods

2.1. Patient serum and antibodies

Sera samples from seven fish-allergic patients were obtained from Qingdao Women and Children's Hospital. Anti-fish PV IgE levels were determined using the ImmunoCAP (Thermo Fisher Scientific, Waltham, Massachusetts, USA) (Table S1). Sera from two crustacean-allergic, fish-

tolerant patients and three healthy individuals were used as the control. Sera was used for the immunoblotting, surf-blot, and dot-blotting assays. All participants provided the sera voluntarily with clinal consent forms and the medical ethics committee of Qingdao University approved the ethics approval for this study (QDU-HEC-2023252). All sera were stored at -80 °C until further use.

Rabbit anti-PV polyclonal IgG antibodies were prepared by ProteinGene Biotech Co., Ltd. (Wuhan, China). For each New Zealand white rabbit (BGI Co., Shenzhen, China), a total of six subcutaneous injections of purified PV from *L. maculatus* and *Gadus morhua* emulsified in Complete Freund's Adjuvant were applied on different days (injection quantity): 0 (400 μg), 28 (150 μg), 42 (150 μg), 47 (150 μg), 54 (150 μg), and 64 (150 μg). Rabbit sera were prepared and assessed by indirect ELISA with purified PVs. All rabbit antibodies were diluted to identical titers with phosphate-buffered saline (PBS) and stored at -80 °C until further use.

2.2. Purification and identification of parvalbumin from *L. maculatus*

Fresh *L. maculatus* used in this study was purchased from the fish market in Qingdao. The dorsal white muscle was cut into 1 cm pieces and stored at -80 °C. The extraction and purification of PV were referenced from Sun et al. with slight modifications[31]. Ten volumes of Tris-HCl buffer (0.1 M Tris buffer, 0.5 mM Glycine, and 1 mM DTT) were used to extract PVs. The increasing saturation of ammonium sulphate from 60% to 100%

was conducted to purify PVs. The MW cut-off membrane of 3.5 KDa (Millipore Corporation, Billerica, MA, USA) was adopted for the dialysis to remove the impurities with low molecular weight. The concentration of purified PV was measured using a Bicinchoninic acid Protein Assay Reagent Kit (Solarbio Co. Beijing, China).

The pure PVs were separated and analyzed with 15% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Gels were stained with Coomassie Brilliant Blue R-250, subsequently scanned and analyzed using Image Lab™ software (Bio-Rad, CA, USA) after destaining. The IgG and IgE-binding capacity of pure PVs were evaluated via immunoblotting with in-house polyclonal rabbit anti-cod PV IgG and human serum IgE, respectively. In brief, purified PVs were firstly transferred to polyvinylidene fluoride (PVDF) membranes with Bio-Rad semi-dry transfer cell at 20V for 10 min (Bio-Rad, California, USA). Membranes were blocked with 1% BSA/PBST (PBS with 1% tween 20). After three times washing with PBST, membranes were incubated with polyclonal rabbit anti-cod PV IgG (1:6000 dilution) and human serum IgE (1:20 dilution) at 4 °C overnight, respectively. After three times washing, membranes were subsequently incubated with HRP-labeled goat anti-rabbit IgG and anti-human IgE (1:6000 dilution, Thermo Fisher Scientific, MA, USA), respectively. Membranes were eventually scanned after incubation with the ECL chemiluminescent reagent. The density of target

protein bands was quantified using ImageJ software.

A surf-blot protocol was established to assess the IgE-binding of other proteins extracted using PBS alone and heating extracts at 90 °C for 10 min, respectively. PBS-extracted proteins were firstly transferred to PVDF membranes with Bio-Rad semi-dry transfer cell at 23V for 20 min (Bio-Rad, California, USA). Membranes were incubated with human serum (1:20 dilution) at 4 °C overnight using a Mini-PROTEAN® II multiscreen apparatus (Bio-Rad, California, USA). Membranes were incubated with HRP-labeled goat anti-human IgE antibody and analyzed after ECL chemiluminescent reagent adding.

Suspected PV bands at 10-15 kDa were from SDS-PAGE gels and analyzed using mass spectrometry after tryptic digestion as described previously[32]. Proteins were first recycled with acetonitrile and then digested by trypsin. Peptides were separated and identified using high-performance liquid chromatography coupled with the high-resolution mass spectrometer (HPLC-HRMS) (Q Exactive™ Plus, Thermo Fisher Scientific, MA, USA). Peptides mapping and visualization were performed with Mascot (v. 2.3.02) and MaxQuant (v.2.2.0.0) against the NCBI database (https://www.ncbi. nlm.nih.gov/protein/?term=parvalbumin, accessed on December 10, 2023). Since the information on *L. maculatus* PVs was not reported in the NCBI database, therefore the amino acid sequences were tentatively predicted based on the data of HPLC-HRMS.

178 **2.3. Sequencing the nucleic acid and amino acid of β-PV**

179 Besides the proteomics approach, genomic applications including
180 (RNA-Seq) and polymerase chain reaction (PCR) were used to confirm
181 nucleic acid sequence and amino acid sequence of β-PV from *L. maculatus*.

182 **2.3.1. Transcriptome resequencing**

183 Total RNA was extracted from fresh seabass muscle using TRIzol
184 reagent (Thermo Fisher Scientific, MA, USA). The Agilent 2100
185 Bioanalyzer system assessed extracted RNA's concentration, purity, and
186 integrity. The RNA-Seq was performed in conjunction with Biomarker
187 Technologies Co. (Beijing, China). Briefly, total mRNA was purified by
188 Oligo (dT) magnetic beads selectively attached to poly A, subsequently
189 fragmented into 200-300 bp in fragmentation buffer. First-strand cDNA
190 was synthesized with six base random primers and then used as the
191 template to synthesize the second-strand cDNA. After that, cDNA
192 fragments were enriched by PCR amplification, and library selection was
193 performed based on the preferred 450 bp in length. The quality, total and
194 effective concentration of the library were inspected and determined using
195 a 2100 Bioanalyzer system. The single-strand library was constructed via
196 alkali denaturation of the mixed library containing various index sequences.
197 Eventually, the sequencing library was sequenced using Next-Generation
198 Sequencing (NGS) on the Illumina Hiseq sequencing platform.

199 Data was analyzed by mapping to the reference genome. Raw data

was filtered to obtain the high-quality sequence (Clean Data). The clean data was mapped to the reference genome using HISAT2 software. The distribution of mapped Reads on different gene structural regions (CDS, intron, intergenic, and UTR) was statistically analyzed. The read count of each gene was calculated as the gene expression by HTSeq. The Fragments Per Kilobase of transcript per Million fragments mapped (FPKM) was adopted to standardize the gene expression levels. The analyses based on GO, Pfam, and Swiss-Prot databases were carried out with TBtools software.

Genes highly matched with PV were inspected due to FPKM values. The corresponding proteins were added and used for the new database construction. The raw MS data was re-analyzed by MaxQuant (v.2.2.0.0) against the database mentioned above. The mRNA of PV, which highly matched with PVII, was acquired and further used to identify nucleic acid and amino acid sequences.

2.3.2. Rapid amplification of cDNA ends (RACE) analysis

5′ and 3′-RACE experiments were performed to confirm the complete nucleic acid sequence using 3'-RACE Kit (Sangon Biotechnology Co., Shanghai, China) according to the manufacturer's instructions. In brief, 10 μM of 3'adaptor primer (5'-GCTGTCAACGATACGCTACGTAACGGC ATGACAGTGTTTTTTTTTTTTTTTTT-3') was designed and synthesized for the cDNA synthesis. The nested PCR method was

developed using at least two sets of primers. A pair of primers, including a forward primer (5′-GAGGAGGATGAGCTGAAGCTGT-3′) and a reverse primer (RPM, 5′-GCTGTCAACGATACGCTACGTAAC-3′), were designed as the external primers for the first round of nested PCR amplification. Another pair of primers, including a forward primer (5′-CTGAGACCAAGACGTTTCTCCAG-3′) and a reverse primer (5′-GCTACGTAACGGCATGACAGTG-3′), were used as the internal primers for the second round of amplification. The RACE PCR products were separated and analyzed on 1% agarose gel. DNA sequencing was performed with the help of Sangon Biotechnology Co., Ltd. (Shanghai, China). The amino acid sequence of β-PV was deduced from the above-confirmed DNA sequence by DNAman software.

2.4. Modeling of the secondary and three-dimensional structure

The secondary structure of PVs was predicted from the PSIPRED online server (http://bioinf.cs.ucl.ac.uk/psipred/). The three-dimensional structure (3D) was simulated via iterative threading assembly refinement (I-TASSER), Robetta, SWISS-MODEL, and AlphaFold 2[33]. The stimulated model was assessed by SAVE (https://saves.mbi.ucla.edu/). Since PV belongs to the calcium-binding protein group, the binding sites and molecular parameters were computed via ScanProsite (https://prosite.expasy.org/scanprosite/) and ProtParam (https://web.expasy.org/protparam/), respectively.

2.5. Immunoinformatic prediction of linear epitopes

Multi bioinformatics tools including BepiPred 2.0 server (http://tools.iedb.org/bcell/), IEDB (http://tools.immuneepitope.org/bcell/), DNAStar Protean system (Version 7.1.0), and Predicted Antigenic Peptides (PAP) (http://imed.med.ucm.es/Tools/antigenic.pl) were used for the prediction of B-cell linear epitopes of β-PV. Therein, BepiPred 2.0 server and DNAStar Protean system filter epitopes based on proteins' physiochemical properties such as hydrophilicity, flexibility, β-turn, and accessibility. In addition, the confirmed linear epitopes of Cyp c1, Gad m1, and Sal s1 from IEDB were also selected for comparison. The final candidate epitopes were produced with at least 90% purity in Sangon Biotechnology Co., Ltd.

The predicted linear epitopes were primarily validated by inhibition dot blotting assay using three patients' sera (QDF 1, 9, and 15) according to the method described previously[34]. In brief, 2 μg of purified *L. maculatus* PVs were spotted on the damp-dry nitrocellulose membranes and blocked for 2 h in 1% BSA/PBST. The predicted peptides (10 μg) were pre-incubated with human serums (1:40 dilution) for 1 h at 37°C, and further coated with the membranes for 1.5 h. After incubating with HRP-labeled secondary antibodies, membranes were visualized by ECL reagent as above. The density of target protein bands was quantified using ImageJ software.

2.6. Cross-reactivity assessment of PVs from *L. maculatus*

To assess the cross-reactivity of rabbit IgG specific to parvalbumins from *L. maculatus*, immunoblotting and inhibition ELISA were performed among 26 common fish species commonly consumed in China (Table S3). Fish proteins were extracted using PBS and thermally treated for the preliminary purification of PVs. 5 μg of protein was separated using *SurePAGE*, Bis-Tris, and 4–20% gel (Genscript Biotech Co., Nanjing, China) and then transferred onto a PVDF membrane, followed by incubation with rabbit anti-*L. maculatus* PV polyclonal IgG antibody (1:10000 dilution). The intensity of binding was quantified using ImageJ software after visualization with ECL reagent. For inhibition ELISA, 100 ng of purified seabass and cod PV dissolved in 0.1M carbonate-bicarbonate buffer were pre-coated on the ELISA plates overnight at 4 °C, respectively. After washed three times by PBST, plates were blocked by 1% skim milk and 5 % BSA. As the inhibitors, the extracted fish proteins (50 μg) were pre-incubated with 50 μL of rabbit anti-*L. maculatus* and *G. morhua* PV polyclonal IgG antibodies (1:100000 dilution). After three times washing with PBST, PV was detected with HRP-coupled goat anti-rabbit IgG antibody in PBST (1:1000 dilution). The color reaction was conducted by adding TMB substrate then terminated with 2 M sulfuric acid. Plates were scanned at 405 nm via a microplate reader (CMax Plus, CA, USA). Bovine serum albumin (BSA) was used as the control. The inhibition rate was

calculated based on the OD values. All determinations were performed with three duplicates and expressed as mean values.

2.7. Statistical analysis

All the quantitative results were analyzed in triplicates and presented as means ± standard deviation (SD). Statistical analysis was performed by one-way analysis of variance (ANOVA) via GraphPad prism 8.

3. Results and discussion

3.1. Purification and identification of *L. maculatus* PVs

Spotted seabass (*L. maculatus*) is a major edible fish in Asia due to its high nutritional value and high yield, but it can also cause severe anaphylaxis in sensitized patients. Strikingly, no allergic proteins in *L. maculatus* have been reported. In this study, the major allergen PV from *L. maculatus* was extracted by the Tris buffer and further purified by heating and ammonium sulfate fractionation. Fish allergens such as PV, tropomyosin, and collagen are mostly thermo-stable proteins, and therefore concentrated in the protein extract through heat treatment to remove the heat-labile proteins[35]. We combined gradient ammonium sulfate fractionation and dialysis for the subsequent purification of PVs. In contrast to the frequently-used column chromatography method, salt fractionation has the advantages of time-saving, low cost, and high efficiency, but is limited to separating the protein isoforms completely[36]. SDS-PAGE analysis showed that target proteins between 10 and 15 kDa

were extracted by Tris buffer (Figure 1A, lane 1) and exhibited thermostability after heat processing (Figure 1A, lane 2). Eventually, the purified proteins displaying two adjacent bands were successfully obtained with a purity higher than 95% (Figure 1A, lane 3). The total yield was about 20 mg from 100 g of white muscle. As shown in Figure 1B, both two bands were detected by the rabbit anti-cod PV polyclonal IgG antibody, which indicated that PVs with two isoforms, namely PVI and PVII, had been successfully extracted and purified from *L. maculatus*. PVI and PVII exhibited the binding ability with IgE from fish-allergic individuals (Figure 1C). Although the sIgE levels from fish-allergic subjects differ little, while the IgE binding capacity among them showed a significant difference. There was no binding with shrimp-allergic patients' and healthy individuals' sera. Consistent with PV-specific antibody binding, PVII showed stronger IgE binding capacity than PVI. Notably, PVI was 1.8-fold more abundant as compared to PVII but exhibited an approximately 2.4-fold higher signal with the anti-PV antibody and at least 1.5-fold higher signal with human serum. It has been reported that the allergenicity of PV is closely associated with its content[37]. However, native PV isoforms are complex to be separated and purified because of their similar physicochemical properties, so they are rarely further investigated in comparison to isoforms' allergenicity. Huang et al. recombined the α-PV of turbot (*Scophthalmus maximus*) and compared the allergenicity with

natural PVs by KU812 cells release assay[38]. The recombinant α-PV showed lower IgG binding activity and allergenicity than natural PVs, possibly due to the Ca^{2+} depletion[19]. To reveal the relationship between the allergenicity of PVs and their contents, the determination of natural PVs extracted from various fish species using the same antibody needs to be further conducted.

Enolase, aldolase, and triosephosphate isomerase in salmon raw PBS-raw extracts and PVs, tropomyosin, and collagen in PBS-heated extracts had been reported as fish allergens based on the surf-blot analyses[39]. To investigate the IgE-binding capacity of other proteins, a surf-blot system was established in this study. The PBS-raw and heated protein extracts were incubated with seven fish-allergic patients' sera, three shrimp-allergic patients, and two healthy individuals. As shown in Figure S1, only one band at about 12 kDa showed a robust IgE-binding capacity with sera from six fish-allergic patients, without binding with other negative sera. Comparing the grave value of bands from raw and heated extracts, the IgE binding capacity was directly proportional to the PV concentration. In addition, there was no binding on other fish allergens except PV in this study. Tropomyosin was the second most crucial fish allergen, after PV, which showed a high abundance and could be recognized by up to 36% of patients' sera. Some patients' recognized tropomyosin but not PV and tropomyosin is a critical component in Skin Prick Test extracts for allergy diagnosis[39,40]. This is a challenge for fish allergy component-resolved

diagnosis and fish allergen detection. In this study, *L. maculatus* tropomyosin with a molecular weight of 36 kDa was found on SDS-PAGE but could not be identified by the sera from fish-allergic patients and even shrimp-allergic patients. Hence, the characteristics and cross-reactivity of fish allergens, such as tropomyosin, collagen, triosephosphate isomerase, etc., merit further research.

Besides SDS-PAGE and immunoblotting, HRMS is an accurate method commonly applied for allergen identification by mapping peptides against the NCBI database containing amino acid sequences. The matched proteins were supposed to contain at least one unique peptide. The PVs were cut and extracted from gels (Figure 1A), and then digested to multi-peptides. The peptides with various MWs and polarities were separated by a C18 column with gradient elution. Because *L. maculatus* PV information is unknown in NCBI and Uniprot database, peptides match in this study was against fish PV database (https://www.ncbi.nlm.nih.gov/protein/?term=parvalbumin, accessed on December 10, 2023). As shown in Figure 2A, PV II displayed high homology with multi proteins consisting of XP_037546693.1, XP_042253152.1, and WP_237742019.1 et al., ranking in order of decreasing match score. Most of them are named PV beta which has been recognized as the major fish allergen. The highly matched peptides are displayed and marked in red in Figure 2B, thereinto, peptide of ALTDAETK, which is the specific peptide of fish β-PV, had

been detected with a higher intensity (Figure 2C)[32]. In addition, PVI also exhibited high identity with fish PVs (XP_037546693.1, XP_053196838.1, and WP_237742019.1). In contrast to PVII, the peptide of SGFIEEDELK in PVI showed a high intensity. The amino acid sequence of PVII was tentatively predicted based on the matched score and intensity by MaxQuant (v.2.2.0.0). Besides the sequence of N-terminal amino acids, the amino acid sequence of β-PV containing SGFLEEDELK and ALTDAETK with higher intensity had been listed and marked in red, as shown in Figure 2D. However, further transcriptomics and genomics methods should be used to validate this sequence.

3.2. Identification of PVs' nucleic acid and amino acid sequence

To explore the amino acid sequence, the nucleic acid sequence should be first confirmed by the RNA-Seq. RNA-Seq is an efficient and high-throughput technology that can sequence cDNA libraries transcribed from all RNAs extracted from tissues and cells on the Illumina platform[41]. The transcripts were mapped onto the existing genomic sequences to obtain comprehensive genetic information like transcription localization. RNA-Seq has been widely applied in medical and clinical research, but is rarely used in food allergens confirmation[42]. The total RNA sequence of *L. maculatus* had been determined and mapped onto the reference genome (assembly ASM402354v1, https://ngdc.cncb.ac.cn, accessed on December 10, 2023). Approximately 48 million raw reads were analyzed with at least

94.45% high-quality reads over a Phred quality score (Q score) Q30, indicating the deep-sequencing data denotes an accuracy of 99.9%. The RNAs of PV in fish muscle were suspected to be the high expression genes. Thus, the normalized (FPKM) was employed to evaluate the gene expression level and used for the cluster analysis. The thirty FPKM values' highest-ranked genes were listed and displayed on the heatmap shown in Figure 3A. The factors containing the molecular function of calcium ion binding from the GO database, EF-hand domain pair from the Pfam database, and matched PV from the Swiss-Prot database were considered for further screening. Two genes, evm. TU. scaffold_90.59 and evm. TU. scaffold_135.39, were finally obtained and translated into protein, as shown in Figures 3B and 3C. Proteins were named P-PVI and P-PVII, respectively. P-PVI had a number of 109 amino acids, but P-PVII had a number of 184 amino acids which exceeded the normal range. P-PVII had the peptides of ALTDAETK and SGFLEEDELK that matched the MS-predicted data and conformed to the characteristics of β-PV. It is noteworthy that the amino acid 98 to 109 is IGCDAPDSFCYK, which has a meager match rate in the fish PV database (MS data). Compared with multi-fish PVs, the amino acid residues at 102 and 103 are most often E and F, respectively, indicating that the sequence P-PVII is partially untrue. Furthermore, the amino acid sequences of P-PVI and P-PVII were imported into the established fish PV database. Peptides of purified PVs

were reanalyzed and matched to the new database by MaxQuant (v.2.2.0.0). The α-PV highly matched to the sequence of P-PVI with 8 unique peptides and the highest unique sequence coverage (53.2%). The β-PV showed a highly significant match to P-PVII with eight unique peptides. Upon combination of MS prediction (Figure 2D) and RNA-seq (Figure 3C), the amino acid 98 to 109 (IGVDEFAAMIKG) predicted from MS results was used. Thus, the amino acid sequence of β-PV was preliminarily inferred as follows: MAFAGLLNDA DITAALAACQ AADSFKHKEF FAKVGLAAKT PDDIKKAFAV IDQDKSGFIE EDELKLFLQN FSAGARALTD AETKTFLQAG DSDGDGKIGV DEFAAMIKG. The sequence was further verified by PCR assay.

Full-length amino acid sequences of food allergens are commonly acquired by PCR with the primers designed depending on the known RNA sequence or the sequence shared relatively high homology[33,43,44]. Forward and reverse primers had been designed for P-PVI (P1/P2) and P-PVII (PV1/PV2) (Table S2). The sequence of P-PVI has been successfully assembled and validated, which is identical to the result shown in Figure 2B. However, the RNA strand of P-PVII could not be synthesized due to the wrong 3'RNA sequence (294 bp to 327 bp). To obtain the full-length RNA sequence of β-PV and the 3' end of the RNA genomes, the 3'RACE method was conducted based on the sequence of P-PVII. 3'RACE method is an expansively utilizable molecular tool to explore the 3' end sequence

by setting the naturally occurring poly(A) tail of mRNA as the typical sequence, and usually performed with nested PCR assay[45]. cDNA strand of P-PVII was firstly synthesized using reverse transcriptase and 3'adaptor primer. To verify the accuracy of the target sequence, the specific cDNA was then amplified using primers of specific genes located in the middle position as follows: GAGGAGGATGAGCTGAAGCT and CTGAGACCAAGACGTTTCTCCAG. The corresponding amino sequences are EEDELKL and LTDAETKTFLQ, respectively, located in the conserved regions of fish PVs[43]. The primers are listed in Table S2 (MRS1/MRS2). During the nested PCR assay, specific primers of 5' known genes were designed and used for the specific cDNA synthesis. The PCR products at approximately 400 bp with at least 98% purity had been visualized by 1% agarose gel electrophoresis (Figure 4A). After clone sequencing, the RNA sequence and corresponding amino acid sequence of β-PV are displayed in Figure 4B.

The protein sequences obtained by MS, RNA-seq, and 3'RACE methods were contrasted by the UniprotKB Align system (https://www.uniprot.org/align), and the homology amino acids were marked with a yellow shade. As shown in Figure 4C, the percent identity of sequence MS predicted is 79.17% and 87.50%, to RNA-seq and 3'RACE methods, respectively. Although each peptide has the highest mascot score (or the highest match for multiple peptides), the MS-matched

method alone could not confirm the N-terminal peptide sequences and identify the individual amino acid residue. Sequence identified by RNA-seq shared 79.17% identity with that of MS, and 92.66% identity with the 3'RACE results. However, RNA-seq supplied the wrong C-terminal peptide sequences (marked in blue shade), but could accurately predict the sequence in combination with the MS method. RACE assay is the most efficient way to obtain and validate the full-length RNA sequence in this study. Whereas, the method cannot work without the known partial or specific nucleic acid sequence which could be acquired by MS prediction and RNA-seq. Therefore, the advantages of the multi-omics method for allergen amino acid sequence confirmation are apparent.

3.3. Secondary structure prediction and three-dimensional modeling of PVs

The structures of PVs have been well characterized using bioinformatics systems, X-ray crystallography, and NMR spectroscopy. Fish PVs with several isoforms shared similar structures with EF-hand motifs for Ca^{2+} binding at the C-terminus, while the N-terminal EF-hand site is a non-functional domain[20,46]. Ca^{2+} in PVs plays essential roles in muscle relaxation and other biological processes[47]. In addition, PVs are found to be conserved in thermal stability, chemical denaturation, enzymatic digestion, and allergenicity, which are closely associated with structures[31,47–49]. To explore the relationship between protein function and

structure, the secondary and 3D structures of PVs were predicted based on the confirmed amino acid sequences with multi-bioinformatics systems. As shown in Figure 5A and Figure S2A, the secondary structure prediction via the PSIPRED server showed that PVs from *L. maculatus* mainly consist of six α-helixes regions, six random coils, and PVII has one strand. These elements are identical to those of the other PV structures[20]. The IgE reactivity was stable with the heating process at 90 °C, however, it was completely lost at 140 °C, probably due to the destruction of its secondary structure, while IgE-binding to PV in highly processed canned fish products has been reported [50–52]. Li et al. elucidated that ohmic heating of PV could reduce its allergenicity by approximately 65% through increasing β-sheets and unordered structures and decreasing α-helixes[53]. As the heating temperature increases, protein will be aggregated and further disrupt secondary and tertiary structures, ultimately leading to allergenicity reduction by modifying linear or conformational epitopes[54,55]. Four systems containing I-TASSER, Robetta, SWISS-MODEL, and AlphaFold 2 were used for tertiary structure modeling, and all data were evaluated with SAVE requirements. Besides Robetta, the stimulated models of PVII predicted by the I-TASSER, SWISS-MODEL, and AlphaFold 2 had successfully passed the VERIFY 3D assessments with the ERRAT overall quality factors at 97.00%, 94.90%, and 97.03%, respectively. Higher factors denote more accurate structure modeling. Since the models I-

TASSER and AlphaFold 2 obtained showed a minor variation in quality factor, the Ramachandran plots were further analyzed and compared. As shown in Figure S3, the amino acid residues in the most favored regions (A, B, L) modeled by the AlphaFold 2 accounted for 92.9%, much higher than that of I-TASSER (85.7%). Therefore, the AlphaFold 2 was finally selected for the PVII modeling, and its outcomes of ERRAT and Ramachandran plot were exhibited in Figures 5B and 5C. The 3D structure modeling of PVI was also conducted on AlphaFold 2 with the highest ERRAT overall quality factors (Figures S2B). 3D structure models were displayed by PyMOL (Figure 5D and Figure S2D). Similar to the predicted secondary structure, PVII was composed mainly of α-helix and random coil. The predicted molecular weights of PVI and PVII are 11595 and 11521Da with a pI of 4.60 and 4.53, respectively. The results above revealed that the PVI and PVII purified from *L. maculatus* were α-PV and β-PV, respectively.

Two calcium-binding domains were predicted in the amino acid sequence of 52-64 and 91-103 which were located on the loop regions flanked by the first two helix pairs (Figure 5D), further indicating that PVs are in good agreement with calcium-binding EF-hand superfamily properties. In the calcium-binding domains, aspartic acid is the significant binding site. The calcium binding maintains the allergenicity of PVs[19]. The recombinant PV or apo-protein without calcium binding showed a stronger

IgG binding, however, the reduced IgE-binding capacity and the weaker effector cells release ability than the natural PV[31,56,57]. Accordingly, to decrease the risk of fish PV, the food-producing technology towards destroying PV structures and Ca^{2+} depletion is worth looking forward to in the future.

3.4. Identification and validation of IgE epitopes of β-PV

The characterization of epitopes is beneficial to the development of food allergy component-resolved diagnosis and allergen-specific immunotherapy. Vaccines produced based on the B-cell linear epitopes have been applied to allergen-specific immunotherapy in clinical trials[58,59]. Various approaches including pepscan, minotopes, X-ray diffraction, mutagenesis, mass spectrometry, and computational prediction have been developed for linear epitope identification[60]. The utilization of computational and experimental methods based on known databases is fairly straightforward epitopes and can essentially manage and interpret large amounts of data. The linear B-cell epitopes of Gad m1, Sal s1, and Cyp c1 had been explored in the IEDB (Table 1). To obtain the representative B-cell epitopes of β-PVs, Carrera M et al. collected and compared 98 β-PVs, and further predicted the epitopes by Kolaskar and Tongaonkar based on the IEDB, ultimately reported four epitopes (position) containing LKLFLQV (AAs 65-71), ACAHLCK (AAs 1-7), FAVLVKQ (AAs 104-110) and LFLQNFV (AAs 67-73)[27]. However, the above

552 experiments did not consider epitopes protease hydrolysis process.

553 With the help of three immunoinformatic tools and IEDB, the epitopes
554 of β-PV were predicted as shown in Table 1. Epitope prediction was
555 supposed to consider multiple parameters including hydrophilicity, surface
556 probability, flexibility, and allergenic index (Figure S4A). Partial amino
557 acids at a range of 1-10 performed with low hydrophilicity but showed
558 flexible and allergenic indexes were also selected as the potential epitopes
559 (Figures S4A and S4B). Epitopes predicted by different methods
560 overlapped in some regions which are the vital reference. Lysine (K) is the
561 protease cleavage site, then set as the epitopes' last amino acid. The final
562 seven candidate epitopes (P1-7) were obtained as shown in Table 2 and
563 further synthesized to be validated by inhibition dot blotting assay.
564 Thereinto, P3 was designed as the final protean digested product of P2 to
565 elaborate the effect of digestion on epitope activity. IgE-binding activities
566 of predicted epitopes were assessed using three fish-allergic patients’ sera.
567 PV and BSA added groups were set as the positive and negative control
568 groups, respectively. As shown in Figure 6A, all epitopes performed the
569 IgE binding activities with different degrees. The inhibition rates were
570 calculated as shown in Figure 6B. The average inhibition rates of all
571 epitopes had exceeded 50%, but P4, P5, and P7 showed a wide variation,
572 possibly due to the serum IgE titer of patient QDF15. To validate whether
573 the predicted epitopes can be used for allergen detection, the inhibition dot

blotting assay using rabbit anti-*L. maculatus* PV IgG antibodies (1:8000 dilution) were conducted. Besides P7 and P4, other epitopes showed significant IgG-binding activities. P3 exhibited the most effective suppression, whereas P2 and P3 also showed similar suppression on IgE-binding capacity, but had a significant gap in IgG-binding capacity, indicating that protean digestion factors had a minor influence on the epitope IgE binding capacity but greatly affected IgG-binding. P1 and P2 had weak hydrophilicity and had to be dissolved in the dimethyl sulphoxide solution. After knocking out the last three amino acids (P3), the peptides turned out to be hydrophilic and might have a strong IgG-binding capacity.

Furthermore, seven B-cell linear epitopes of β-PV were mapped on the 3D structure by PyMOL software (Figure 6C). All epitopes were located on the surface of β-PV, and P1-P7 were mainly distributed on the α-helixes. Both P5 and P7 were located on the calcium-binding loops. The most conserved amino acid region of fish PV was reported to be located on the first calcium-binding site (containing P5), and the most common B-cell epitopes were located between positions 25 and 45 (containing P3 and P4)[26]. In contrast to what Carrera M et al. reported, the representative fish B-cell epitopes were outside the scope of what we predicted using multi-immunoinformatic tools. This difference might result from different scoring criteria of immunoinformatic tools. Therefore, well-recognized epitope prediction methods based on computer learning were supposed to

be developed.

3.5. Cross-reactivity analysis among common fish species

It is widely accepted that one fish-allergic individual has an approximately 50% risk of being allergic to other fish species, which can be attributed to the cross-reactivity of fish PVs[61]. Hence, numerous methods based on anti-PV antibodies have been established for fish allergen detection and fish allergy diagnosis. In contrast to other food allergens, methods targeting PVs are challenging because of their high biochemical and immunological variability among the different fish species, even among PV isoforms[26]. Gad m 1 from Atlantic cod (*G. morhua*) had been commonly used in commercial ELISA kits for fish allergen detection. However, the ELISA kit based on the Gad m 1 showed poor detection capacity among 57 bony fish, and could not determine cartilaginous fish samples[61]. Hence, several polyclonal or monoclonal antibodies produced in rabbits and mice against natural PVs purified from different fish species (Asian seabass, carp, cod, pilchard, and salmon) or recombinant PV (carp) had recently been successfully applied in fish allergen detection[23,39,62]. Moreover, the polyclonal IgG raised against Asian seabass PV was the most cross-reactive compared with anti-pilchard, anti-seabass, and anti-salmon[62]. As with Asian seabass, *L. maculatus* belongs to the perciformes. The cross-reactivity of *L. maculatus* PVs deserves to be investigated for its further utilization in fish allergen detection.

618 Fish species commonly used in cross-reactivity assessment are
619 frequently consumed in Europe and North America. The commonly
620 consumed fish species in China were less studied. Therefore, the cross-
621 reactivity of *L. maculatus* PV toward 26 Chinese common fish species
622 (Table S3) was evaluated by immunoblotting and inhibition ELISA. As
623 shown in Figure 7, PBS heated extracts of fish had been separated by SDS-
624 PAGE and PVs identified by rabbit anti-*L. maculatus* PV polyclonal IgG
625 antibody. The recombinant Gad m1 (No. 27) and purified *L. maculatus* PVs
626 (No. 28) were also analyzed for the comparison. As shown in Figure 7B,
627 92% of the 26 fish species had been detected by the target antibody. The
628 antibody can identify not only β-PVs but also α-PVs from bony fish, but it
629 cannot bind the PV from cartilaginous fish (*Triakis scyllia*, No.26). The
630 biophysical properties and allergenicity of α-PV from cartilaginous fish
631 were different with PV from bony fish probably due to the various amino
632 acid sequence[22]. There are few successful examples of detecting α-PV from
633 cartilaginous fish and other cross-reacting vertebrates using cross-reactive
634 antibodies[48,62,63]. Two factors govern the allergenicity of allergens, one is
635 the type of allergen (in this case, PV), and another is the content of the
636 target allergen in the tissue[37]. To explore the association of the PV contents
637 and their IgG binding capacity, the intensities of bands on SDS-PAGE and
638 immunoblotting were analyzed by Image J (Figure S5). The binding
639 capacity exhibited a considerable fish-specific variation regardless of PV

contents. For instance, the contents of fish PVs from grouper (No.19), *Lateolabrax maculatus* (No.28), *Paralichthys olivaceus* (No.12), *Ctenopharyngodon idellus* (No.1), and *Late scalcarifer* (No.18) ranked as the top five among all the fish species. However, the IgG binding capacity ranked top five in *Ctenopharyngodon idellus* (No.1), *Hypophthalmichthys molitrix* (No.2), *Cyprinus carpio* (No.4), *Carassius carassius* (No.5), and *Lateolabrax maculatus* (No.17), respectively. It is noteworthy that PVs from freshwater fishes exhibited stronger IgG binding capacity than that of marine fishes. We inferred that PVs from freshwater fishes might be a better alternative for detecting fish allergens and diagnosing fish allergy.

Whether polyclonal anti-*L. maculatus* PV IgG antibody will be a good substitute for fish allergen determination remains to be investigated. Both the anti-*L. maculatus* and *G. mohua* PVs IgG antibodies were prepared for the inhibition ELISA toward 26 fish species, recombinant Gad m1, and purified *L. maculatus* PVs. As shown in Figure 8, the inhibition ELISA exhibited a higher sensitivity than the immunoblotting method as PVs in *Thunnus alba* (No.20) and even in cartilaginous fish (*Triakis scyllia*, No.26) were recognized. All samples had been detected and the overall inhibition rate of anti-*L. maculatus* PV IgG antibody (Figure 8A) was higher than the anti-cod PV antibody (Figure 8B), indicating that the polyclonal antibody can be used for fish allergen detection. The anti-*L. maculatus* antibody showed a more robust binding capacity on PV from Cypriniformes and

Perciformes. Inhibition rates of two antibodies were higher in freshwater fishes than marine fishes which is consistent with the results of immunoblotting assay (Figure 7B). Both Asian seabass and spotted seabass belong to the Perciformes, and their PVs have been well characterized on the molecular features, and then exhibit cross-reactivity toward various fish species[62]. Thus, the corresponding antibodies used in fish allergen detection are very promising.

In summary, as one of the most commonly consumed maricultural fish species, *L. maculatus* allergens have not previously been intensively investigated. To our knowledge, this is the first study that characterized the molecular properties of major allergen-PVs from *L. maculatus* with the help of multi-omics assays. Moreover, the epitopes and structure of β-PV were obtained using immunoinformatics approaches. The cross-reactivity toward common Asia-Pacific fish species was assessed and revealed that anti-*L. maculatus* PV antibody could be regarded as a replacement for traditional anti-cod antibodies in prospective applications. This study will thus enhance the understanding of fish PVs and provide a theoretical basis for fish allergen determination and further promote the development of fish allergen component-resolved allergy diagnosis. Further studies should focus on the comparative allergenicity of recombinant fish PVs and their application in allergen detection and allergy diagnosis.

Acknowledgments

This work was funded by the grants of the National Natural Science Foundation of China (nos. 32102090) and Shandong Postdoctoral Innovation Foundation (nos. SDCX-ZG-202202009).

Notes

The authors declare no competing financial interest.

Supporting Information

Clinical characteristics of patients; primers designed for 3'Race; characterization of 26 fish species; surf-blot analysis of fish proteins; structural characteristics of PVI; ramachandran plot analysis; prediction of linear B-cell epitopes; and SDS-PAGE and immunoblotting analysis.

Abbreviations used

CDC, Centers for Disease Control; VITAL, Voluntary Incidental Trace Allergen Labelling; PV, Parvalbumin; TM, Tropomyosin; IEDB, Immune Epitope Database; CFA, Complete Freund's Adjuvant; BCA, Bicinchoninic acid; SDS-PAGE, Sodium dodecyl sulphate-polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; PVDF, polyvinylidene fluoride; HPLC-HRMS, High-resolution mass spectrometer; NGS, Next-Generation Sequencing; CDS, Gene structural regions; FPKM, Fragments Per Kilobase of transcript per Million fragments mapped; RACE, Rapid amplification of cDNA ends.

References

(1) Warren, C. M.; Jiang, J.; Gupta, R. S. Epidemiology and Burden of Food Allergy. *Curr. Allergy Asthma Rep.* **2020**, *20* (2), 6.

(2) Lopes, J. P.; Sicherer, S. Food Allergy: Epidemiology, Pathogenesis, Diagnosis, Prevention, and Treatment. *Curr. Opin. Immunol.* **2020**, *66*, 57–64.

(3) Moonesinghe, H.; Mackenzie, H.; Venter, C.; Kilburn, S.; Turner, P.; Weir, K.; Dean, T. Prevalence of Fish and Shellfish Allergy: A Systematic Review. *Ann. Allergy. Asthma. Immunol.* **2016**, *117* (3), 264-272.

(4) Ruethers, T.; Taki, A. C.; Johnston, E. B.; Nugraha, R.; Le, T. T. K.; Kalic, T.; McLean, T. R.; Kamath, S. D.; Lopata, A. L. Seafood Allergy: A Comprehensive Review of Fish and Shellfish Allergens. *Mol. Immunol.* **2018**, *100*, 28–57.

(5) Kalic, T.; Radauer, C.; Lopata, A. L.; Breiteneder, H.; Hafner, C. Fish Allergy Around the World—Precise Diagnosis to Facilitate Patient Management. *Front. Allergy* **2021**, *2*, 732178.

(6) Savage, J.; Sicherer, S.; Wood, R. The Natural History of Food Allergy. *J. Allergy Clin. Immunol. Pract.* **2016**, *4* (2), 196–203.

(7) Seth, D.; Poowutikul, P.; Pansare, M.; Kamat, D. Food Allergy: A Review. *Pediatr. Ann.* **2020**, *49* (1), 50-58.

(8) Sharp, M. F.; Lopata, A. L. Fish Allergy: In Review. *Clin. Rev. Allergy*

Immunol. **2014**, *46* (3), 258–271.

(9) Lopata, A. L.; Jeebhay, M. F. Airborne Seafood Allergens as a Cause of Occupational Allergy and Asthma. *Curr. Allergy Asthma Rep.* **2013**, *13* (3), 288–297.

(10) Le, T. T. K.; Nguyen, D. H.; Vu, A. T. L.; Ruethers, T.; Taki, A. C.; Lopata, A. L. A Cross-sectional, Population-based Study on the Prevalence of Food Allergies among Children in Two Different Socio-economic Regions of Vietnam. *Pediatr. Allergy Immunol.* **2019**, *30* (3), 348–355.

(11) Luo, J.; Zhang, Q.; Gu, Y.; Wang, J.; Liu, G.; He, T.; Che, H. Meta-Analysis: Prevalence of Food Allergy and Food Allergens — China, 2000−2021. *China CDC Weekly.* **2022**, *4* (34).

(12) Dijkema, D.; Emons, J. A. M.; Van De Ven, A. A. J. M.; Oude Elberink, J. N. G. Fish Allergy: Fishing for Novel Diagnostic and Therapeutic Options. *Clin. Rev. Allergy Immunol.* **2022**, *62* (1), 64–71.

(13) Davis, C. M.; Gupta, R. S.; Aktas, O. N.; Diaz, V.; Kamath, S. D.; Lopata, A. L. Clinical Management of Seafood Allergy. *J. Allergy Clin. Immunol. Pract.* **2020**, *8* (1), 37–44.

(14) Zhu, L.; Li, S.; Sun, L.; Zhao, J.; Huang, J.; Jiang, Y.; Wan, S.; Pavase, T. R.; Li, Z. Development and Validation of a Specific Sandwich ELISA for Determination of Soybean Allergens and Its Application in Processed Foods. *Process Biochem.* **2022**, *117*, 134–141.

(15) Holzhauser, T.; Johnson, P.; Hindley, J. P.; O'Connor, G.; Chan, C.-H.; Costa, J.; Fæste, C. K.; Hirst, B. J.; Lambertini, F.; Miani, M.; Robert, M.-C.; Röder, M.; Ronsmans, S.; Bugyi, Z.; Tömösközi, S.; Flanagan, S. D. Are Current Analytical Methods Suitable to Verify VITAL® 2.0/3.0 Allergen Reference Doses for EU Allergens in Foods? *Food Chem. Toxicol.* **2020**, *145*, 111709.

(16) Remington, B. C.; Westerhout, J.; Meima, M. Y.; Blom, W. M.; Kruizinga, A. G.; Wheeler, M. W.; Taylor, S. L.; Houben, G. F.; Baumert, J. L. Updated Population Minimal Eliciting Dose Distributions for Use in Risk Assessment of 14 Priority Food Allergens. *Food Chem. Toxicol.* **2020**, *139*, 111259.

(17) Dramburg S, Hilger C, Santos A, et al. EAACI Molecular Allergology User's Guide 2.0. *Pediatr Allergy Immu*. **2023**, *34*, 13854.

(18) Haiech, J.; Moreau, M.; Leclerc, C.; Kilhoffer, M.-C. Facts and Conjectures on Calmodulin and Its Cousin Proteins, Parvalbumin and Troponin C. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2019**, *1866* (7), 1046–1053.

(19) Costa, J.; Villa, C.; Verhoeckx, K.; Cirkovic-Velickovic, T.; Schrama, D.; Roncada, P.; Rodrigues, P. M.; Piras, C.; Martín-Pedraza, L.; Monaci, L.; Molina, E.; Mazzucchelli, G.; Mafra, I.; Lupi, R.; Lozano-Ojalvo, D.; Larré, C.; Klueber, J.; Gelencser, E.; Bueno-Diaz, C.; Diaz-Perales, A.; Benedé, S.; Bavaro, S. L.; Kuehn, A.; Hoffmann-

770 Sommergruber, K.; Holzhauser, T. Are Physicochemical Properties
771 Shaping the Allergenic Potency of Animal Allergens? *Clin. Rev.*
772 *Allergy Immunol.* **2022**, *62* (1), 1–36
773 (20) Kumeta, H.; Nakayama, H.; Ogura, K. Solution Structure of the Major
774 Fish Allergen Parvalbumin Sco j 1 Derived from the Pacific Mackerel.
775 *Sci. Rep.* **2017**, *7* (1), 17160.
776 (21) Kuehn, A.; Swoboda, I.; Arumugam, K.; Hilger, C.; Hentges, F. Fish
777 Allergens at a Glance: Variable Allergenicity of Parvalbumins, the
778 Major Fish Allergens. *Front. Immunol.* **2014**, *5*, 179.
779 (22) Stephen, J. N.; Sharp, M. F.; Ruethers, T.; Taki, A.; Campbell, D. E.;
780 Lopata, A. L. Allergenicity of Bony and Cartilaginous Fish –
781 Molecular and Immunological Properties. *Clin. Exp. Allergy* **2017**, *47*
782 (3), 300–312.
783 (23) Sližienė, A.; Plečkaitytė, M.; Rudokas, V.; Juškaitė, K.; Žvirblis, G.;
784 Žvirblienė, A. Cross-Reactive Monoclonal Antibodies against Fish
785 Parvalbumins as a Tool for Studying Antigenic Similarity of Different
786 Parvalbumins and Analysis of Fish Extracts. *Mol. Immunol.* **2023**, *154*,
787 80–95.
788 (24) Freidl, R.; Gstoettner, A.; Baranyi, U.; Swoboda, I.; Stolz, F.; Focke-
789 Tejkl, M.; Wekerle, T.; Van Ree, R.; Valenta, R.; Linhart, B. Blocking
790 Antibodies Induced by Immunization with a Hypoallergenic
791 Parvalbumin Mutant Reduce Allergic Symptoms in a Mouse Model of

792 Fish Allergy. *J. Allergy Clin. Immunol.* **2017**, *139* (6), 1897-1905.
793 (25) Perez-Gordo, M.; Lin, J.; Bardina, L.; Pastor-Vargas, C.; Cases, B.;
794 Vivanco, F.; Cuesta-Herranz, J.; Sampson, H. A. Epitope Mapping of
795 Atlantic Salmon Major Allergen by Peptide Microarray Immunoassay.
796 *Int. Arch. Allergy Immunol.* **2012**, *157* (1), 31–40.
797 (26) Sharp, M. F.; Taki, A. C.; Ruethers, T.; Stephen, J. N.; Daly, N. L.;
798 Lopata, A. L.; Kamath, S. D. IgE and IgG4 Epitopes Revealed on the
799 Major Fish Allergen Lat c 1. *Mol. Immunol.* **2021**, *131*, 155–163.
800 (27) Carrera, M.; González-Fernández, Á.; Magadán, S.; Mateos, J.; Pedrós,
801 L.; Medina, I.; Gallardo, J. M. Molecular Characterization of B-Cell
802 Epitopes for the Major Fish Allergen, Parvalbumin, by Shotgun
803 Proteomics, Protein-Based Bioinformatics and IgE-Reactive
804 Approaches. *J. Proteomics* **2019**, *200*, 123–133.
805 (28) Klueber, J.; Schrama, D.; Rodrigues, P.; Dickel, H.; Kuehn, A. Fish
806 Allergy Management: From Component-Resolved Diagnosis to
807 Unmet Diagnostic Needs. *Curr. Treat. Options Allergy* **2019**, *6* (4),
808 322–337.
809 (29) Cai, L.-S.; Wang, L.; Song, K.; Lu, K.-L.; Zhang, C.-X.; Rahimnejad,
810 S. Evaluation of Protein Requirement of Spotted Seabass (*Lateolabrax*
811 *Maculatus*) under Two Temperatures, and the Liver Transcriptome
812 Response to Thermal Stress. *Aquaculture* **2020**, *516*, 734615.
813 (30) Fisheries Administration. Ministry of Agriculture and Rural Affairs of

814 the People's Republic of China: China Agriculture Press: Beijing,
815 China, *China Fisheries Statistical Yearbook* **2023**.
816 (31) Sun, L.; Xu, L.; Huang, Y.; Lin, H.; Ahmed, I.; Li, Z. Identification
817 and Comparison of Allergenicity of Native and Recombinant Fish
818 Major Allergen Parvalbumins from Japanese Flounder (*Paralichthys*
819 *Olivaceus*). *Food Funct.* **2019**, *10* (10), 6615–6623.
820 (32) Sun, L.; Lin, H.; Li, Z.; Sun, W.; Wang, J.; Wu, H.; Ge, M.; Ahmed, I.;
821 Pavase, T. R. Development of a Method for the Quantification of Fish
822 Major Allergen Parvalbumin in Food Matrix via Liquid
823 Chromatography-Tandem Mass Spectrometry with Multiple Reaction
824 Monitoring. *Food Chem.* **2019**, *276*, 358–365.
825 (33) Zhu, W.; Zhao, J.; Huang, Y.; Ahmed, I.; Wang, H.; Zhang, Z.; Lin, H.;
826 Li, Z. Purification, Expression, and Characterization of Sarcoplasmic
827 Calcium Binding Protein: A Novel Allergen of *Portunus*
828 *Trituberculatus*. *J. Agric. Food Chem.* **2023**, *71* (28), 10773–10786.
829 (34) Xu, L. L.; Chen, J.; Sun, L. R.; Gao, X.; Lin, H.; Ahmed, I.; Pramod,
830 S. N.; Li, Z. X. Analysis of the Allergenicity and B Cell Epitopes in
831 Tropomyosin of Shrimp (*Litopenaeus Vannamei*) and Correlation to
832 Cross-Reactivity Based on Epitopes with Fish (*Larimichthys Crocea*)
833 and Clam (*Ruditapes Philippinarum*). *Food Chem.* **2020**, *323*, 126763.
834 (35) Kalic, T.; Kamath, S. D.; Ruethers, T.; Taki, A. C.; Nugraha, R.; Le, T.
835 T. K.; Humeniuk, P.; Williamson, N. A.; Hira, D.; Rolland, J. M.;

836 O'Hehir, R. E.; Dai, D.; Campbell, D. E.; Breiteneder, H.; Lopata, A.
837 L. Collagen—An Important Fish Allergen for Improved Diagnosis. *J.*
838 *Allergy Clin. Immunol. Pract.* **2020**, *8* (9), 3084-3092.
839 (36) Kobayashi, Y.; Yang, T.; Yu, C.-T.; Ume, C.; Kubota, H.; Shimakura,
840 K.; Shiomi, K.; Hamada-Sato, N. Quantification of Major Allergen
841 Parvalbumin in 22 Species of Fish by SDS–PAGE. *Food Chem.* **2016**,
842 *194*, 345–353.
843 (37) Kobayashi, A.; Kobayashi, Y.; Shiomi, K. Fish Allergy in Patients with
844 Parvalbumin-Specific Immunoglobulin E Depends on Parvalbumin
845 Content Rather than Molecular Differences in the Protein among Fish
846 Species. *Biosci. Biotechnol. Biochem.* **2016**, *80* (10), 2018–2021.
847 (38) Huang, Y.; Li, Z.; Wu, Y.; Guo, Y.; Pavase, T. R.; Chen, G.; Zhang, Z.;
848 Lin, H. Comparison of Immunological Properties of Recombinant and
849 Natural Turbot (Scophthalmus Maximus) Parvalbumin. *Eur. Food Res.*
850 *Technol.* **2021**, *247* (8), 2053–2065.
851 (39) Ruethers, T.; Taki, A. C.; Karnaneedi, S.; Nie, S.; Kalic, T.; Dai, D.;
852 Daduang, S.; Leeming, M.; Williamson, N. A.; Breiteneder, H.; Mehr,
853 S. S.; Kamath, S. D.; Campbell, D. E.; Lopata, A. L. Expanding the
854 Allergen Repertoire of Salmon and Catfish. *Allergy* **2021**, *76* (5),
855 1443–1453.
856 (40) Ruethers, T.; Taki, A. C.; Nugraha, R.; Cao, T. T.; Koeberl, M.; Kamath,
857 S. D.; Williamson, N. A.; O'Callaghan, S.; Nie, S.; Mehr, S. S.;

858 Campbell, D. E.; Lopata, A. L. Variability of Allergens in Commercial
859 Fish Extracts for Skin Prick Testing. *Allergy* **2019**, *74* (7), 1352–1363.
860 (41) Xu, Z.; Zhang, T.; Chen, H.; Zhu, Y.; Lv, Y.; Zhang, S.; Chen, J.; Chen,
861 H.; Yang, L.; Jiang, W.; Ni, S.; Lu, F.; Wang, Z.; Yang, H.; Dong, L.;
862 Chen, F.; Zhang, H.; Chen, Y.; Liu, J.; Zhang, D.; Fan, L.; Guo, G.;
863 Wang, Y. High-Throughput Single Nucleus Total RNA Sequencing of
864 Formalin-Fixed Paraffin-Embedded Tissues by snRandom-Seq. *Nat.*
865 *Commun.* **2023**, *14* (1), 2734.
866 (42) Withanage MHH, Liang H, Zeng E. RNA-Seq Experiment and Data
867 Analysis. *Methods Mol Biol*. **2022**, 2418, 405-424.
868 (43) Carrera, M.; Cañas, B.; Gallardo, J. M. Rapid Direct Detection of the
869 Major Fish Allergen, Parvalbumin, by Selected MS/MS Ion
870 Monitoring Mass Spectrometry. *J. Proteomics* **2012**, *75* (11), 3211–
871 3220.
872 (44) Yu, C.; Gao, X.; Lin, H.; Xu, L.; Ahmed, I.; Khan, M. U.; Xu, M.;
873 Chen, Y.; Li, Z. Purification, Characterization, and Three-Dimensional
874 Structure Prediction of Paramyosin, a Novel Allergen of *Rapana*
875 *Venosa*. *J. Agric. Food Chem.* **2020**, *68* (49), 14632–14642.
876 (45) Green, M. R.; Sambrook, J. Rapid Amplification of Sequences from
877 the 3′ Ends of mRNAs: 3′-RACE. *Cold Spring Harb. Protoc.* **2019**,
878 *2019* (5), 426-430.
879 (46) Moraes, A. H.; Ackerbauer, D.; Kostadinova, M.; Bublin, M.; De

Oliveira, G. A.; Ferreira, F.; Almeida, F. C. L.; Breiteneder, H.; Valente, A. P. Solution and High-Pressure NMR Studies of the Structure, Dynamics, and Stability of the Cross-Reactive Allergenic Cod Parvalbumin Gad m 1: Rational Design of Hypoallergenic Molecules. *Proteins Struct. Funct. Bioinforma.* **2014**, *82* (11), 3032–3042.

(47) Vologzhannikova, A. A.; Shevelyova, M. P.; Kazakov, A. S.; Sokolov, A. S.; Borisova, N. I.; Permyakov, E. A.; Kircheva, N.; Nikolova, V.; Dudev, T.; Permyakov, S. E. Strontium Binding to α-Parvalbumin, a Canonical Calcium-Binding Protein of the "EF-Hand" Family. *Biomolecules* **2021**, *11* (8), 1158.

(48) Saptarshi, S. R.; Sharp, M. F.; Kamath, S. D.; Lopata, A. L. Antibody Reactivity to the Major Fish Allergen Parvalbumin Is Determined by Isoforms and Impact of Thermal Processing. *Food Chem.* **2014**, *148*, 321–328.

(49) Griesmeier, U.; Vázquez-Cortés, S.; Bublin, M.; Radauer, C.; Ma, Y.; Briza, P.; Fernández-Rivas, M.; Breiteneder, H. Expression Levels of Parvalbumins Determine Allergenicity of Fish Species. *Allergy* **2010**, *65* (2), 191–198.

(50) Kubota, H.; Kobayashi, A.; Kobayashi, Y.; Shiomi, K.; Hamada-Sato, N. Reduction in IgE Reactivity of Pacific Mackerel Parvalbumin by Heat Treatment. *Food Chem.* **2016**, *206*, 78–84.

(51) Pérez-Tavarez, R.; Moreno, H. M.; Borderias, J.; Loli-Ausejo, D.;

902 Pedrosa, M.; Hurtado, J. L.; Rodriguez-Pérez, R.; Gasset, M. Fish
903 Muscle Processing into Seafood Products Reduces β-Parvalbumin
904 Allergenicity. *Food Chem.* **2021**, *364*, 130308.
905 (52) Taki, A. C.; Ruethers, T.; Nugraha, R.; Karnaneedi, S.; Williamson, N.
906 A.; Nie, S.; Leeming, M. G.; Mehr, S. S.; Campbell, D. E.; Lopata, A.
907 L. Thermostable Allergens in Canned Fish: Evaluating Risks for Fish
908 Allergy. *Allergy* **2023**, *78* (12), 3221–3234.
909 (53) Li, X.; Deng, Y.; Qiu, W.; Feng, Y.; Jin, Y.; Chen, L.; Li, L.; Wang, A.
910 L.; Tao, N.; Jin, Y. Effects of Different Ohmic Heating Treatments on
911 Parvalbumin Structure and Reduction of Allergenicity in Japanese Eel
912 (Anguilla Japonica). *Food Chem.* **2024**, *432*, 137257.
913 (54) Jiang, X.; Rao, Q. Effect of Processing on Fish Protein Antigenicity
914 and Allergenicity. *Foods* **2021**, *10* (5), 969.
915 (55) Zhao, J.; Zeng, J.; Liu, Y.; Lin, H.; Gao, X.; Wang, H.; Zhang, Z.; Lin,
916 H.; Li, Z. Understanding the Mechanism of Increased IgG/IgE
917 Reactivity but Decreased Immunodetection Recovery in Thermally
918 Induced Shrimp (*Litopenaeus Vannamei*) Tropomyosin via
919 Multispectroscopic and Molecular Dynamics Simulation Techniques.
920 *J. Agric. Food Chem.* **2023**, *71* (7), 3444–3458.
921 (56) Kobayashi, A.; Ichimura, A.; Kobayashi, Y.; Shiomi, K. IgE-Binding
922 Epitopes of Various Fish Parvalbumins Exist in a Stereoscopic
923 Conformation Maintained by Ca2+ Binding. *Allergol. Int.* **2016**, *65* (3),

345–348.

(57) Swoboda, I.; Bugajska-Schretter, A.; Linhart, B.; Verdino, P.; Keller, W.; Schulmeister, U.; Sperr, W. R.; Valent, P.; Peltre, G.; Quirce, S.; Douladiris, N.; Papadopoulos, N. G.; Valenta, R.; Spitzauer, S. A Recombinant Hypoallergenic Parvalbumin Mutant for Immunotherapy of IgE-Mediated Fish Allergy. *J. Immunol.* **2007**, *178* (10), 6290–6296.

(58) 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.

(59) Valenta, R.; Campana, R.; Niederberger, V. Recombinant Allergy Vaccines Based on Allergen-Derived B Cell Epitopes. *Immunol. Lett.* **2017**, *189*, 19–26.

(60) Nilvebrant, J.; Rockberg, J. An Introduction to Epitope Mapping. *Methods Mol Biol.* **2018**,*1785*,1-10.

(61) Ruethers, T.; Taki, A. C.; Khangurha, J.; Roberts, J.; Buddhadasa, S.; Clarke, D.; Hedges, C. E.; Campbell, D. E.; Kamath, S. D.; Lopata, A. L.; Koeberl, M. Commercial Fish ELISA Kits Have a Limited Capacity to Detect Different Fish Species and Their Products. *J. Sci.*

946 *Food Agric.* **2020**, *100* (12), 4353–4363.
947 (62) Sharp, M. F.; Stephen, J. N.; Kraft, L.; Weiss, T.; Kamath, S. D.;
948 Lopata, A. L. Immunological Cross-Reactivity between Four Distant
949 Parvalbumins—Impact on Allergen Detection and Diagnostics. *Mol.*
950 *Immunol.* **2015**, *63* (2), 437–448.
951 (63) Ruethers, T.; Nugraha, R.; Taki, A. C.; O'Malley, A.; Karnaneedi, S.;
952 Zhang, S.; Kapingidza, A. B.; Mehr, S.; Kamath, S. D.; Chruszcz, M.;
953 Mackay, G.; Campbell, D. E.; Lopata, A. L. The First Reptilian
954 Allergen and Major Allergen for Fish-allergic Patients: Crocodile B-
955 parvalbumin. *Pediatr. Allergy Immunol.* **2022**, *33* (5), 13781.
956

Figure captions

Figure 1 Purification and identification of *L. maculatus* extracts. (A) SDS-PAGE of extracted proteins. Lane 1, raw extracts; Lane 2, heat extracts; Lane 3, purified parvalbumins. (B) Immunoblotting of purified parvalbumins with polyclonal rabbit anti-cod (*G. morhua*) parvalbumin IgG antibody. (C) Immunoblotting of purified parvalbumins with human sera. QDF1, QDF3, QDF7, QDF9, and QDF13 are from fish-allergic patients; QDS1 and QDS2 are from shrimp-allergic patients; QDN1-3 are from healthy individuals. The patient's clinical information was referred to the Table S 1.

Figure 2 Mass spectroscopic analysis of purified PVII. (A) Cluster analysis

of homologous proteins. (B) Matched peptides in top 3 matching proteins. Higher matched peptides are marked in red. (C) The fragment ion spectrum of fish parvalbumin specific peptide ALTDAETK. (D) The predicted amino acid sequence of PVII by mass spectrometry. Peptides with higher signal intensity was marked in red.

Figure 3 Transcriptome resequencing the nucleic acid sequences of parvalbumins. (A) Heatmap of the top 30 genes based on Fragments Per Kilobase of transcript per Million fragments mapped. (B) The nucleic acid and amino acid sequence of P-PVI. (C) The nucleic acid and amino acid sequence of P-PVII. The suspicious amino acid sequence which is different with the characteristics of fish parvalbumins was marked in yellow shade.

Figure 4 The confirmation of nucleic acid and amino acid sequence of PVII by 3'Race analysis. (A) 1% agarose gel electrophoresis of PCR amplification products. Lane M: marker, indicated in bp. Lane 1, the amplification products of PVII's RNA. (B) The nucleic acid and amino acid sequence of PVII. (C) Comparison of amino acid sequences obtained by MS predicted, RNA-seq, and 3'Race. The same amino acids are marked in yellow shade.

Figure 5 Structural characteristics of PVII. (A) Secondary structure predicted with PSIPRED. (B) ERRAT analysis of AlphaFold-2. (C) Ramachandran plot of AlphaFold-2. (D) The 3D structure model of PVII presented by PyMOL.

Figure 6 Validation of predicted epitopes of β-parvalbumin. (A) Inhibition dot blotting assay analysis of predicted epitopes. QDF1, QDF3, QDF7, QDF9, and QDF15 are human sera from fish-allergic patients. IgG is the polyclonal rabbit anti-*L. maculatus* parvalbumin IgG antibody. (B) The inhibition rate of the predicted epitopes calculated based on the intensity of the dots. (C) The locations of identified epitopes exhibited by PyMOL.

Figure 7 Cross-reactivity assessment of parvalbumins among 26 fishes commonly consumed in China. (A) SDS-PAGE of PBS heated extracts. (B) Immunoblotting of parvalbumins with polyclonal rabbit anti-*L. maculatus* parvalbumin IgG antibody.

Figure 8 Inhibition ELISA analysis of anti-fish parvalbumin IgG antibodies. (A) The Inhibition rate of anti-*L. maculatus* parvalbumin IgG antibody among 26 fish species. (B) The Inhibition rate of anti- *G. morhua* parvalbumin IgG antibody among 26 fish species. The details of 26 fish species are referred in Table S3.

Tables

Table 1. Information of Linear B cell Epitopes Predicted by Bioinformatics Tools.

Bioinformatics tool	Amino acid locations
BepiPred 1.0 Server	5-10, 22-25,36-42,53-62,72-81,91-100
IEDB	6-10, 24-25, 38-42, 52-62, 73-81,91-100

DNAstar Protean system	7-12, 23-29, 39-46, 52-65, 77-85, 89-100
ABCpred	12-23, 26-38,4 5-52, 63-70
Cyprinus carpio (carp)	12-18, 31-50, 72-102, 90-109
Gadus morhua (Atlantic cod)	2-19, 84-109
Salmo salar (Atlantic salmon)	2-8, 7-21, 10-24, 22-36, 55-69, 85-99

Table 2. Information of Synthesized Peptides.

Peptide no.	**Amino acid sequence**	**Position**
P1	LLNDADI	6-12
P2	ADSFKHKEFF	22-31
P3	ADSFKHK	22-28
P4	AKTPDDIK	38-45
P5	SGFIEEDELK	56-65
P6	ALTDAETK	77-84
P7	GDSDGDK	90-97

Figure1

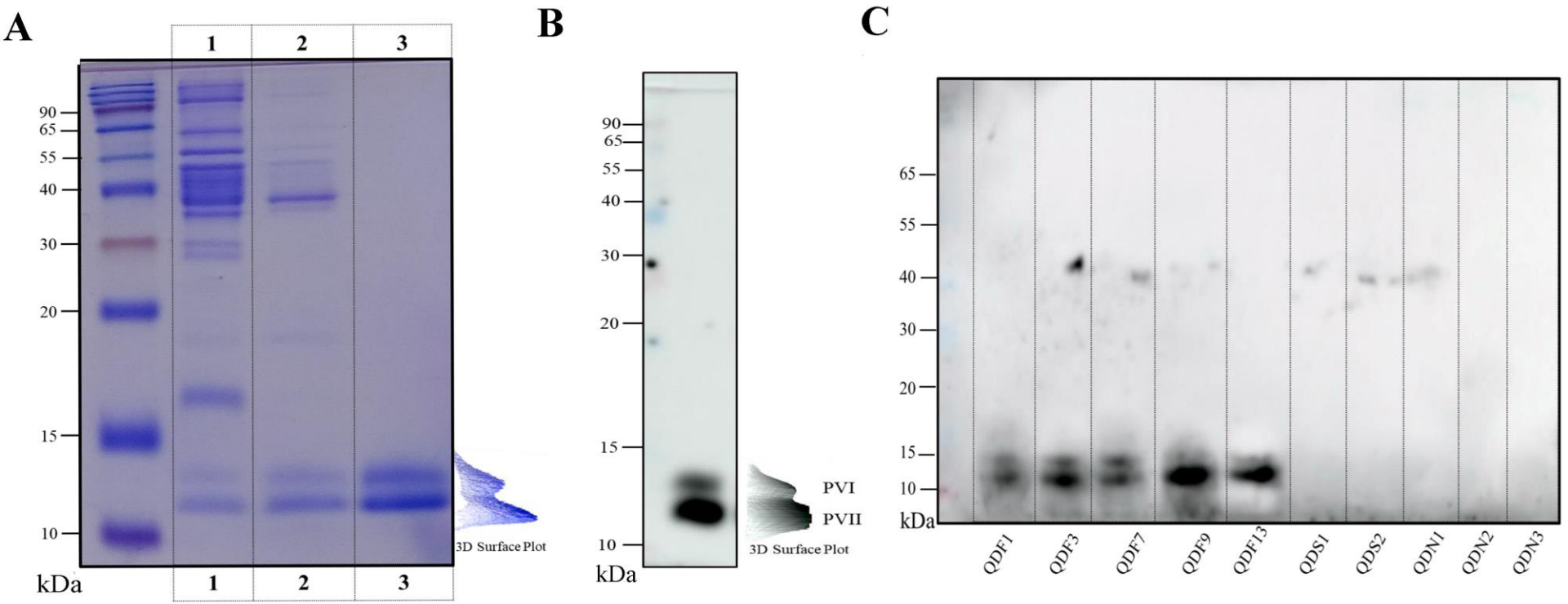
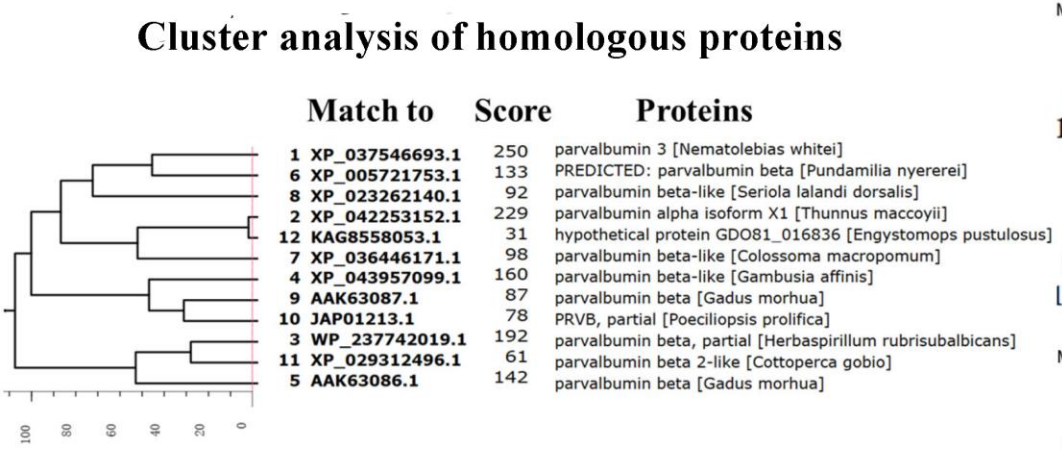

A
1
2
3
90
65
55
40
30
20
15
10
kDa
3D Surface Plot
B
PVI
PVII
3D Surface Plot
C
65
55
40
30
20
15
10
kDa
QDF1
QDF3
QDF7
QDF9
QDF13
QDS1
QDS2
QDN1
QDN2
QDN3

Figure2

A



B

Match to: **XP_037546693.1** Score: **250 parvalbumin 3 [Nematolebias whitei]**

1 MAFAGILSEA DITAALAACQ AADSFKHKEF FAKVGLAAKS ADDVKKAFAV
51 IDQDKSGFIE EEELKLFLQN FSAGARALTD AETKTFLQAG DSDGDGKIGV
101 DEFAAMIKG

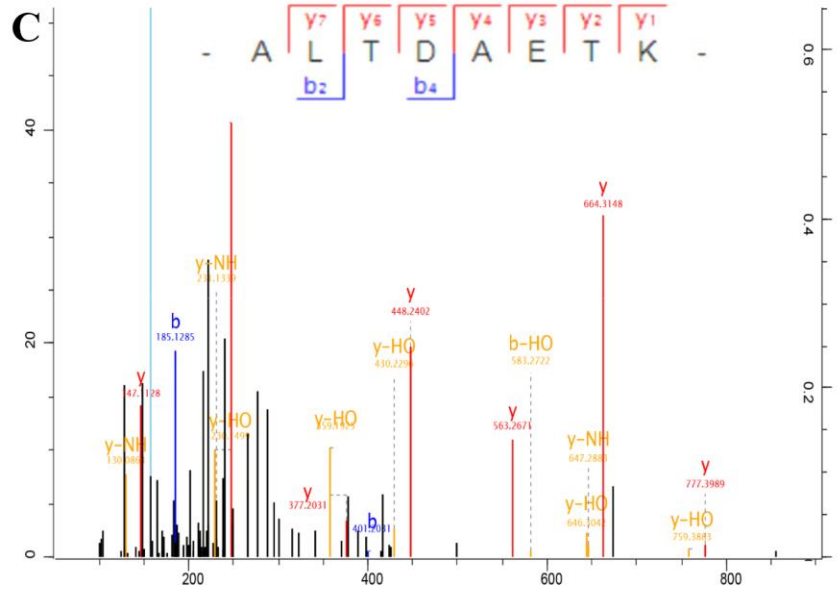
Match to: **XP_042253152.1** Score: **229 parvalbumin alpha isoform X1 [Thunnus maccoyii]**

1 MMMLKMAFAG ILTEADITAA LAACQAADSF KYKDFFAKVG LAAKTPDDIK
51 KAFAVIDQDK SGFIEEDELK LFLQNFSAGA RALTDAETKA FLMAGDSDGD
101 GKIGIDEFAA LVKA

Match to: **WP_237742019.1** Score: **192 parvalbumin beta, partial [Herbaspirillum rubrisubalbicans]**

1 SFILANPNPN PKDNQRLIMA FAGILNDADV TAALQACQAV DSFNYKGFFA
51 KVGLSAKTPD VIKKAFAVID QDKSGFIEED ELKLFLQNFS AGARALTDAE
101 TKAFLKAGDS DGDGKIGVDE FAAMVKA

C



D

MS predicted:

1 * * * * * * * * * * * AALDACE AAESFSYKEF

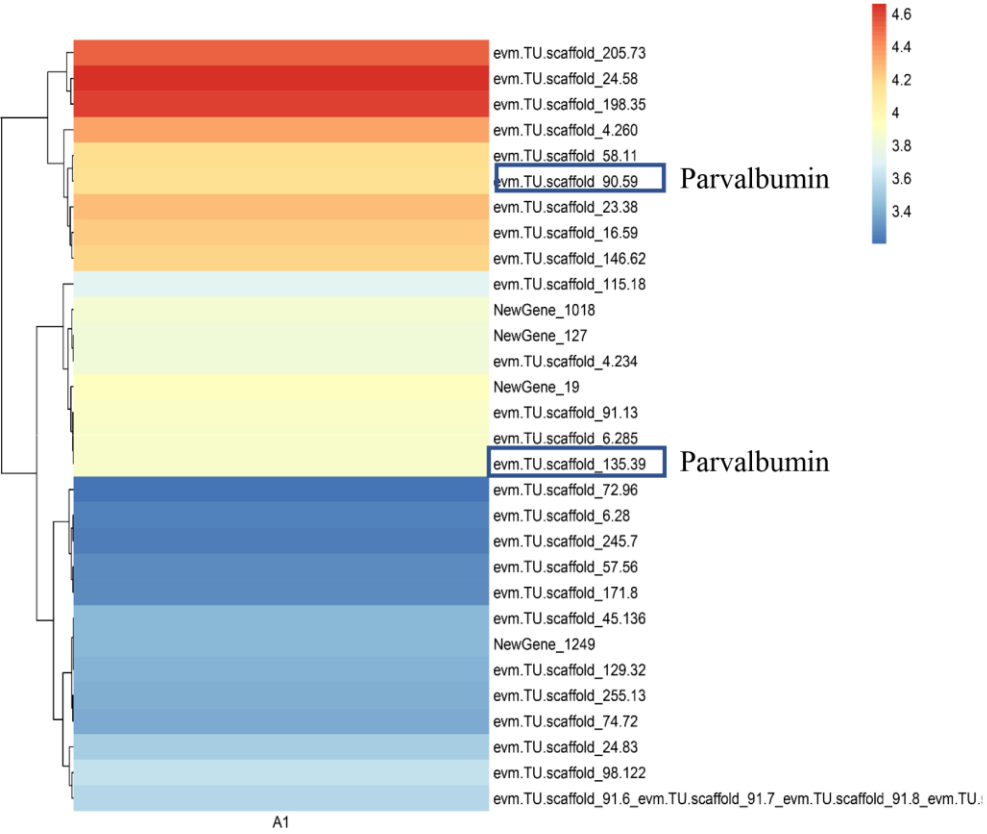
31 FAKVGLAAKT PDDIKKAFAV LDQDK**SGFLE**

61 **EDELK**LFLQN FQAGAR**ALTD AETK**AFLKAG

91 DSDGDGKIGV DEFAAMIKG

Figure3

A



B

>evm.TU.scaffold_90.59
P-PVI

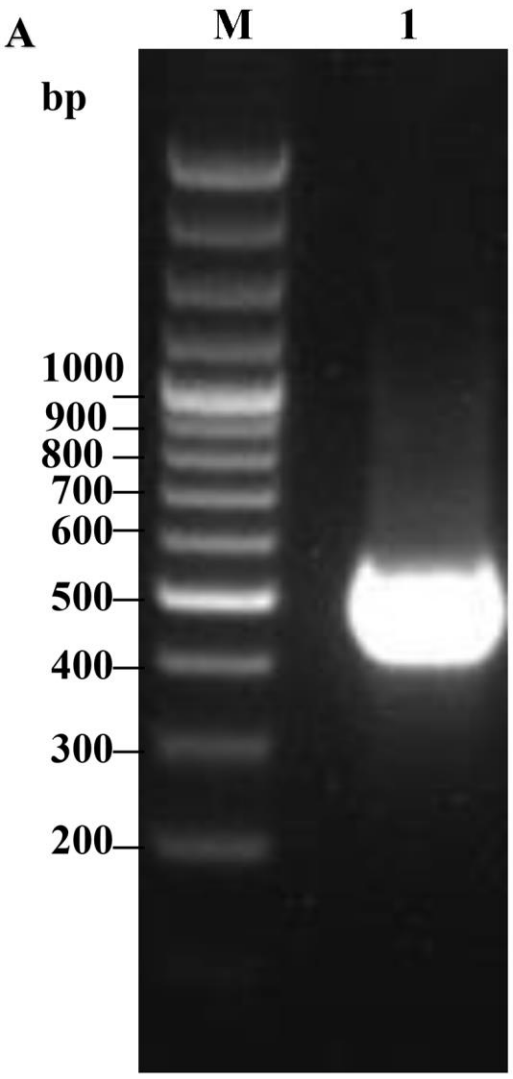
```
ATG GCC TTC GCA GGT GTA CTG AGC GAT GCG GAC ATC ACT GCA GCC CTG GAA GGG TGC AAA
  1 M   A   F   A   G   V   L   S   D   A   D   I   T   A   A   L   E   G   C   K
GGC GCC GGC ACA TTC GAC CAC GTG AAG TTC TTT AAG GCA TGC GGT ATG GCC GGC AAG TCT
 21 G   A   G   T   F   D   H   V   K   F   F   K   A   C   G   M   A   G   K   S
CCC GAG GAC GTC AAG AAG GCC TTC TTC ATC ATC GAC CAG GAC AAG AGC GGC TTC ATT GAG
 41 P   E   D   V   K   K   A   F   F   I   I   D   Q   D   K   S   G   F   I   E
GAG GAC GAG CTC AAG CTG TTC CTG CAG AAC TTC CAG GCA GGT GCA CGT GAA CTG ACC GCC
 61 E   D   E   L   K   L   F   L   Q   N   F   Q   A   G   A   R   E   L   T   A
GAC GAG ACC AAG ACT TTT CTC AAG GCC GGT GAC TCC GAC GGT GAC GGC AAG ATT GGC GCT
 81 D   E   T   K   T   F   L   K   A   G   D   S   D   G   D   G   K   I   G   A
GAC GAG TTC GCT GCC ATG GTT AAG TCA TAA
101 D   E   F   A   A   M   V   K   S   -
```

C

>evm.TU.scaffold_135.39
P-PVII

```
ATG GCA TTC GCA GGA CTT CTC AAC GAT GCT GAC ATC ACT GCA GCC CTT GCA GCT TGC CAA
  1 M   A   F   A   G   L   L   N   D   A   D   I   T   A   A   L   A   A   C   Q
GCT GCC GAC TCT TTC AAG CAC AAG GAG TTC TTT GCT AAG GTC GGC CTG GCT GCC AAG ACC
 21 A   A   D   S   F   K   H   K   E   F   F   A   K   V   G   L   A   A   K   T
CCC GAT GAC ATC AAG AAG GCC TTC GCC GTC ATT GAC CAG GAC AAG AGT GGC TTC ATT GAG
 41 P   D   D   I   K   K   A   F   A   V   I   D   Q   D   K   S   G   F   I   E
GAG GAT GAG CTG AAG CTG TTC CTG CAG AAC TTC TCT GCC GGT GCC AGA GCT CTG ACC GAT
 61 E   D   E   L   K   L   F   L   Q   N   F   S   A   G   A   R   A   L   T   D
GCT GAG ACC AAG ACG TTT CTC CAG GCC GGT GAC TCT GAT GGT GAC GGC AAG ATC GGA GTT
 81 A   E   T   K   T   F   L   Q   A   G   D   S   D   G   D   G   K   I   G   V
GAT GCT CCG GAC TCA TTC TGC TAC AAA AAG TTC TTC CAG TTG TGT GGC CTG TCC TCA AAG
101 D   A   P   D   S   F   C   Y   K   K   F   F   Q   L   C   G   L   S   S   K
ACG CCC AAG GAA GTC AAA GAC GTC TTC CAG ATC CTC GAC GAG GAC AAC AGC GGC TTT ATC
121 T   P   K   E   V   K   D   V   F   Q   I   L   D   E   D   N   S   G   F   I
GAG GAG TCT GAG CTC AAA ATT CCA GAC TAT GGT CCT GTC CTG AGT TCT ACA GAT GAA ATG
141 E   E   S   E   L   K   I   P   D   Y   G   P   V   L   S   S   T   D   E   M
TCT TCG GGT CTT CAG TCC TCT CCC TTC CTC TGG AGT CTC ATG TCA TCA AAA TCC CCA GTG
161 S   S   G   L   Q   S   S   P   F   L   W   S   L   M   S   S   K   S   P   V
TCA AAC GAG AAC TAA
181 S   N   E   N   -
```

Figure4



B

```
ATG GCA TTC GCA GGA CTT CTC AAC GAT GCT GAC ATC ACT GCA GCC CTT GCA GCT TGC CAA
  1 M   A   F   A   G   L   L   N   D   A   D   I   T   A   A   L   A   A   C   Q
GCT GCC GAC TCT TTC AAG CAC AAG GAG TTC TTT GCT AAG GTC GGC CTG GCT GCC AAG ACC
 21 A   A   D   S   F   K   H   K   E   F   F   A   K   V   G   L   A   A   K   T
CCC GAT GAC ATC AAG AAG GCC TTC GCC GTC ATT GAC CAG GAC AAG AGT GGC TTC ATT GAG
 41 P   D   D   I   K   K   A   F   A   V   I   D   Q   D   K   S   G   F   I   E
GAG GAT GAG CTG AAG CTG TTC CTG CAG AAC TTC TCT GCC GGT GCC AGA GCT CTG ACC GAT
 61 E   D   E   L   K   L   F   L   Q   N   F   S   A   G   A   R   A   L   T   D
GCT GAG ACC AAG ACG TTT CTC CAG GCC GGT GAC TCT GAT GGT GAC GGC AAG ATC GGA GTT
 81 A   E   T   K   T   F   L   Q   A   G   D   S   D   G   D   G   K   I   G   V
GAT GAG TTT GCT GCC ATG ATC AAG GGT TAA
101 D   E   F   A   A   M   I   K   G   -
```

C

```
MS predicted ----------AALDACEAAESFSYKEFFAKVGLAAKTPDDIKKAFAVLDIKKSGFLE  60
RNA-seq      MAFAGLLNDADITAALAACQAADSFKHKEFFAKVGLAAKTPDDIKKAFAVIDQDKSGFIE  60
3'RACE       MAFAGLLNDADITAALAACQAADSFKHKEFFAKVGLAAKTPDDIKKAFAVIDQDKSGFIE  60

MS predicted EDELKLFLQNFQAGARALTDAETKAFLKAGDSDGDGKIGVDEFAAMIKG---------- 109
RNA-seq      EDELKLFLQNFSAGARALTDAETKTFLQAGDSDGDGKIGVDAPDSFCYKKFFQLCGLSSK 120
3'RACE       EDELKLFLQNFSAGARALTDAETKTFLQAGDSDGDGKIGVDEFAAMIKG---------- 109

MS predicted ------------------------------------------------------------ 109
RNA-seq      TPKEVKDVFQILDEDNSGFIEESELKIPDYGPVLSSTDEMSSGLQSSPFLWSLMSSKSPV 180
3'RACE       ------------------------------------------------------------ 109

MS predicted ----  109
RNA-seq      SNEN  184
3'RACE       ----  109
```

Percent Identity Matrix

MS	100.00%	79.17%	87.50%
RNA-seq	79.17%	100.00%	92.66%
3'RACE	87.50%	92.66%	100.00%

Figure5

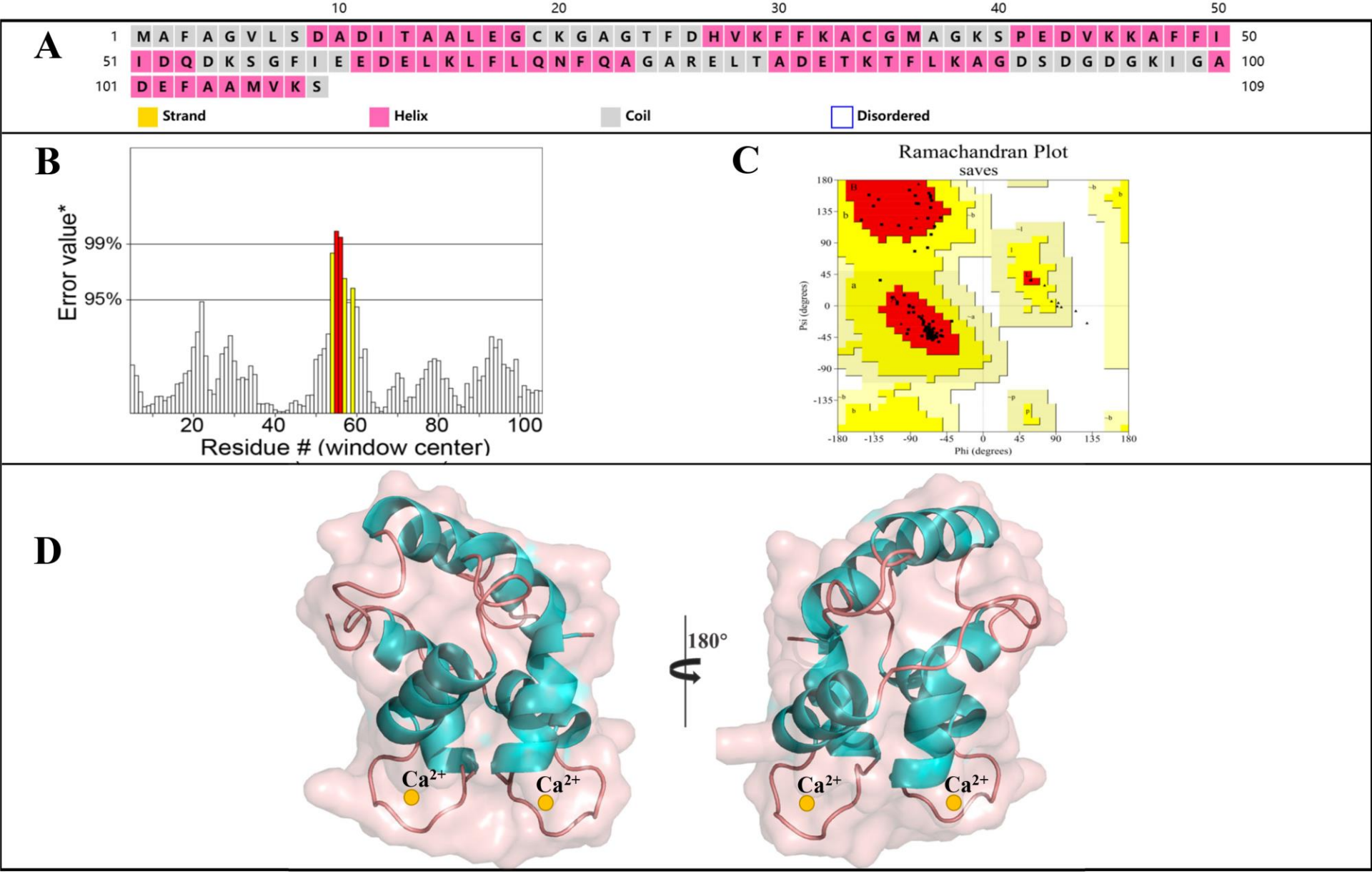

A
10
20
30
40
50
1 M A F A G V L S D A D I T A A L E G C K G A G T F D H V K F F K A C G M A G K S P E D V K K A F F I 50
51 I D Q D K S G F I E E D E L K L F L Q N F Q A G A R E L T A D E T K T F L K A G D S D G D G K I G A 100
101 D E F A A M V K S 109
Strand
Helix
Coil
Disordered
B
Error value*
99%
95%
20
40
60
80
100
Residue # (window center)
C
Ramachandran Plot
saves
Psi (degrees)
Phi (degrees)
D
Ca2+
Ca2+
180°
Ca2+
Ca2+

Figure6

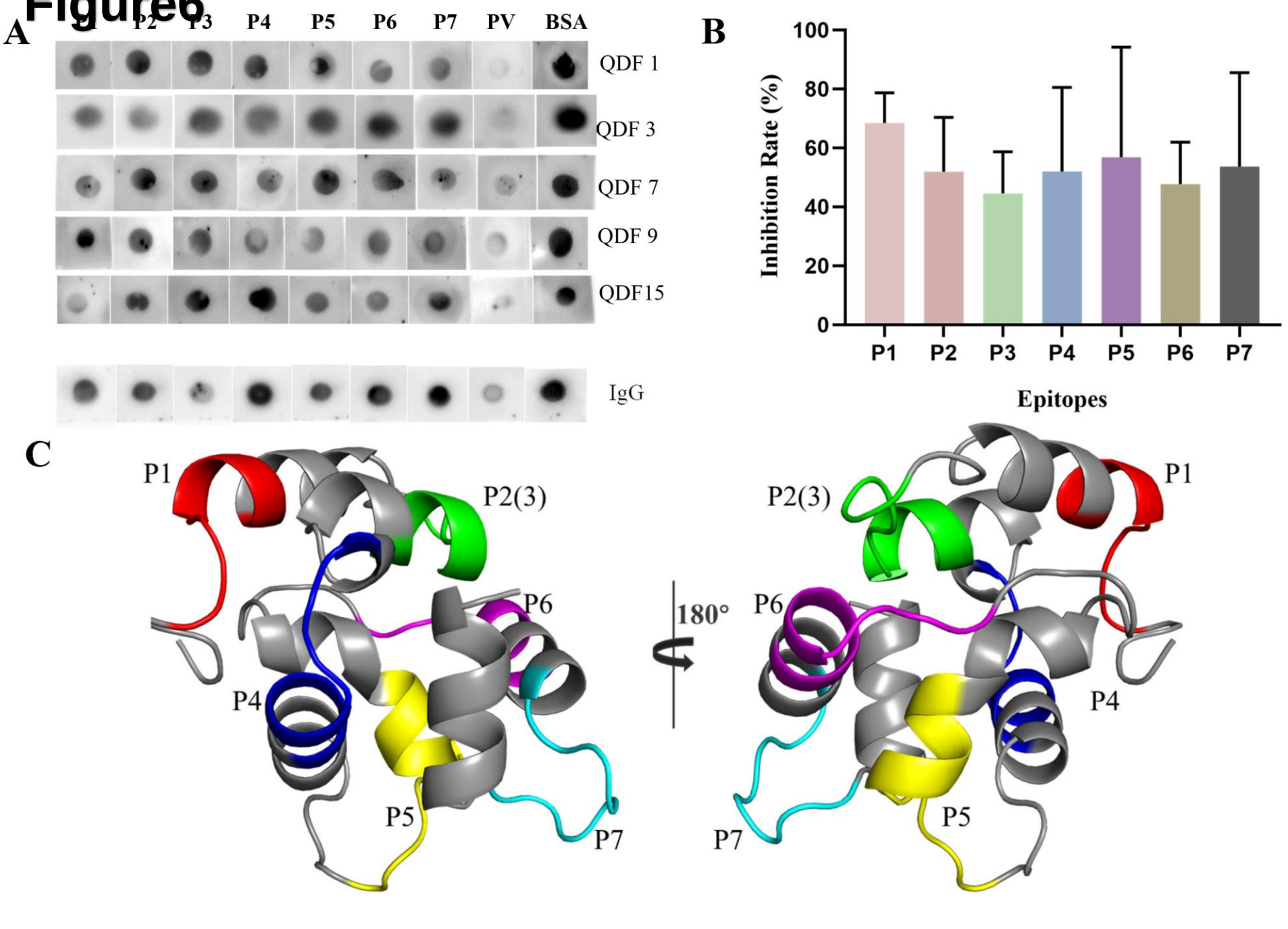
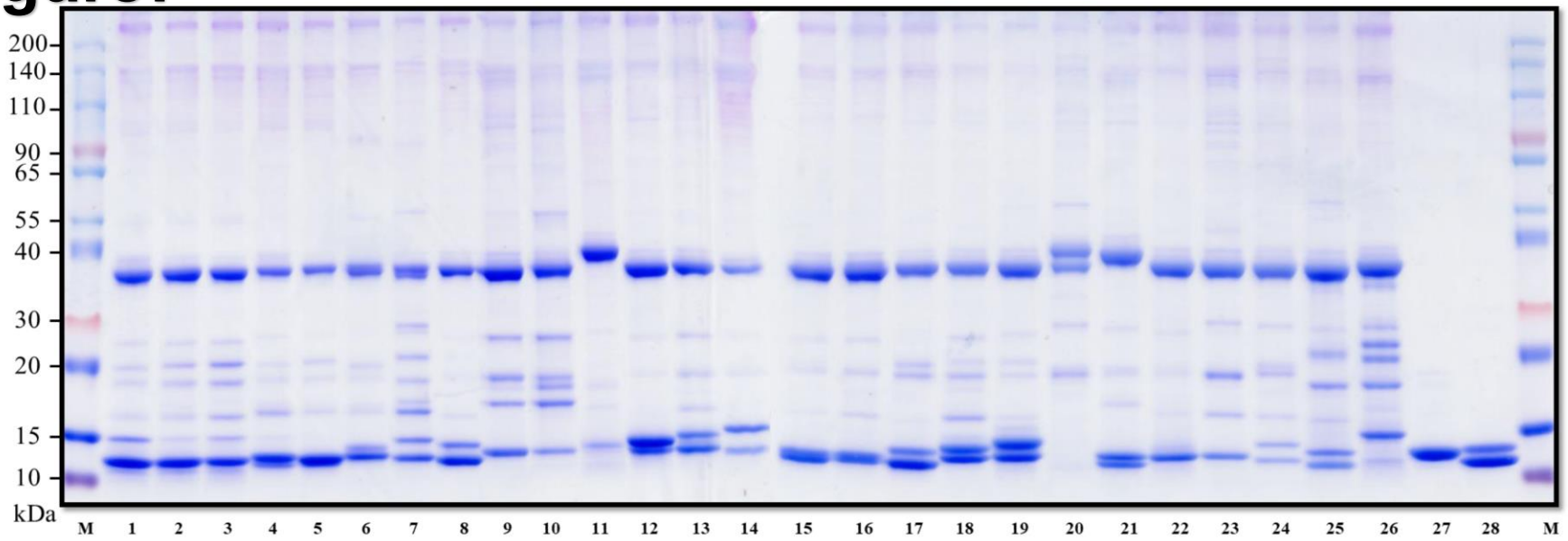

A
P1
P2
P3
P4
P5
P6
P7
PV
BSA
QDF 1
QDF 3
QDF 7
QDF 9
QDF15
IgG
B
Inhibition Rate (%)
100
80
60
40
20
0
P1
P2
P3
P4
P5
P6
P7
Epitopes
C
P1
P2(3)
P6
P4
P5
P7
180°
P2(3)
P1
P6
P4
P5
P7

Figure7

A


200
140
110
90
65
55
40
30
20
15
10
kDa
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 M

B

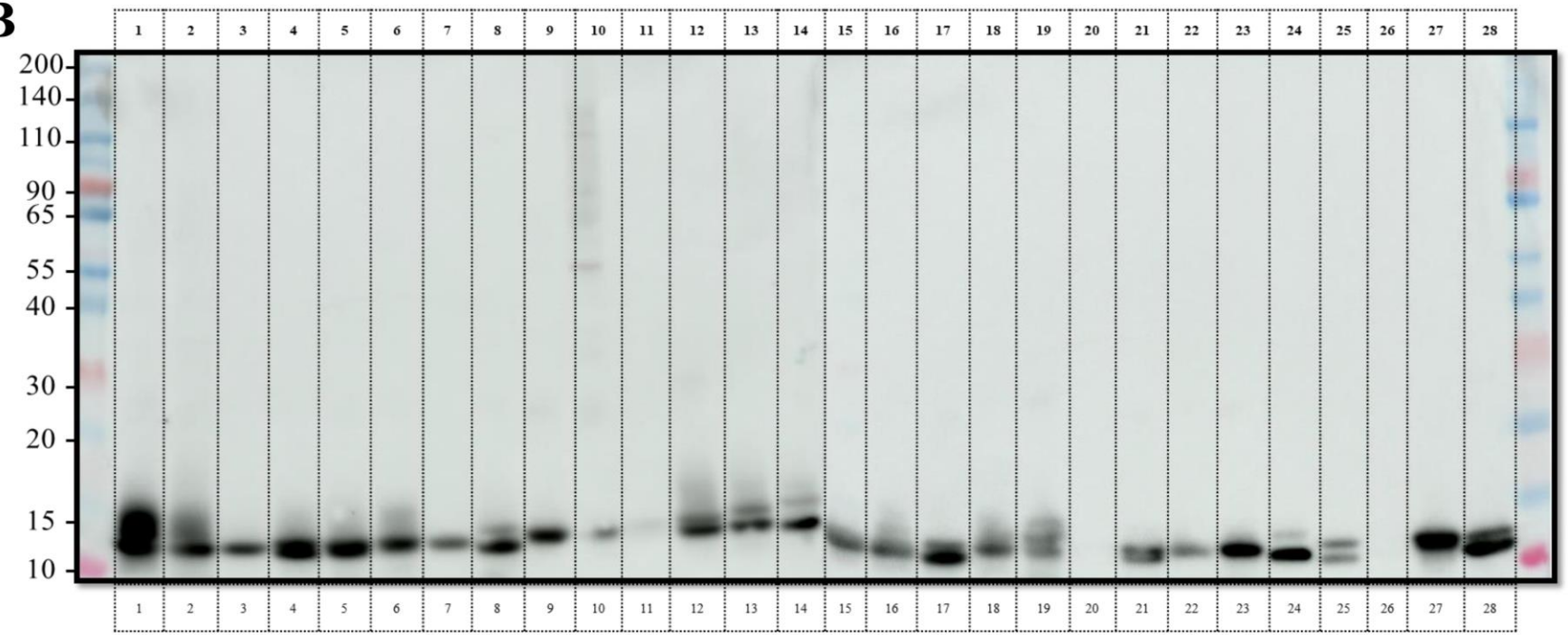
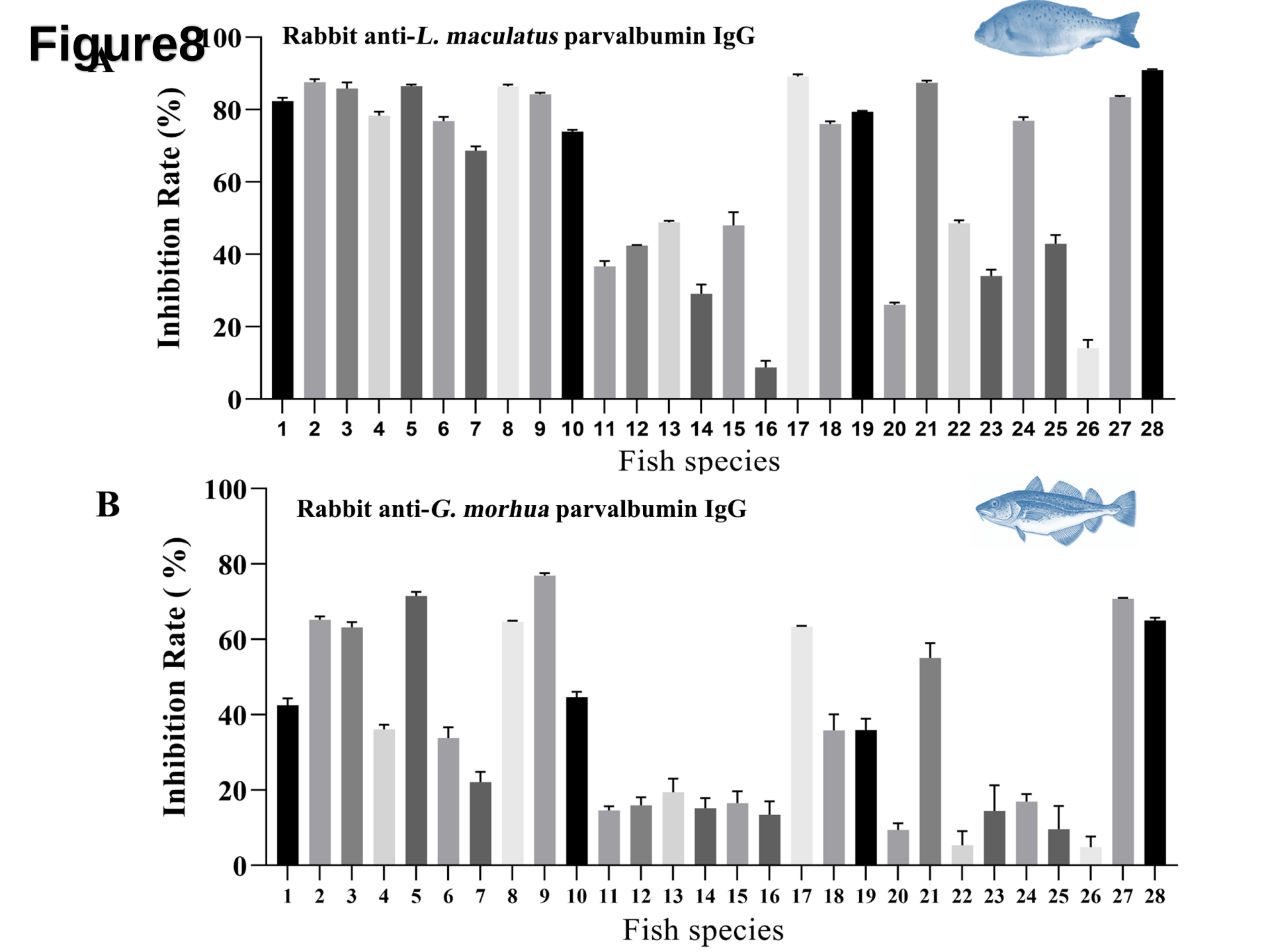

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
200
140
110
90
65
55
40
30
20
15
10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

Figure8


A
Rabbit anti-L. maculatus parvalbumin IgG
Inhibition Rate (%)
0
20
40
60
80
100
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
Fish species
B
Rabbit anti-G. morhua parvalbumin IgG
Inhibition Rate (%)
0
20
40
60
80
100
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
Fish species