



Proteomics for Development of Food Allergy Vaccines

Mónica Carrera and Susana Magadán

Abstract

Food allergy is a hypersensitivity reaction to food products initiated by immunologic mechanisms, which represents one of the major concerns in food safety. New therapies for food allergies including oral and epicutaneous allergen-specific immunotherapy are required, and B cell epitope-based allergy vaccines are a good promise to improve this field. In this chapter, we describe a workflow for the design of food allergy vaccines using proteomic tools. The strategy is defined based on the characterization of B cell epitopes for a particular food allergen. For that, the workflow comprises five consecutive steps: (1) shotgun proteomics analysis of different protein isoforms for a particular food allergen, (2) downloading all protein sequences for the specific allergen included in UniProtKB database, (3) analysis by protein-based bioinformatics of B cell epitopes, (4) synthesizing of the selected B cell peptide epitopes, and (5) performing of immunoassays using sera from healthy and allergic patients. The results from this method provide a rationale repository of B cell epitopes for the design of new specific immunotherapies for a particular food allergen. The strategy was optimized for all the beta-parvalbumins (β -PRVBs), which are considered as the main fish allergens. Using this workflow, a total of 35 peptides were identified as B cell epitopes, among them the top 4 B cell peptide epitopes that may induce protective immune response were selected as potential peptide vaccine candidates for fish allergy.

Key words Food, Allergy, B cell epitopes, Vaccine, Proteomics

1 Introduction

Food allergy is considered one of the major concerns in food safety. Unlike food intolerance, which is a nonimmune reaction that includes metabolic, toxic, pharmacologic, and undefined mechanisms [1]; food allergy is a hypersensitivity reaction to food products initiated by immunologic mechanisms, which can include immunoglobulin E (IgE) mediated and non-IgE-mediated immune pathways [2]. Whereas the symptoms of non-IgE-mediated food hypersensitivity are typically delayed from hours to weeks after ingestion of the food(s), in IgE-mediated food allergic disease, the clinical symptoms usually develop within minutes to a few hours of ingestion [2–4]. They typically involve dermatologic,

respiratory, gastrointestinal, cardiovascular, and/or neurologic symptoms, and can be associated with fatal anaphylaxis, especially with peanut, tree nuts, and seafood. The prevalence of IgE-mediated food allergy has been increasing during the last two decades. It is estimated to affect 6–8% of young children and about 2–4% of adults [4–6].

IgE-mediated food allergy involves two main phases: sensitization followed by the effector phase [2]. The sensitization of allergic individuals results in inappropriate inflammatory immune responses to common food, instead of inducing immune tolerance [2, 7, 8]. The sensitization to foods arises not only after their oral intake but also by topical or airway routes. The default response to food antigens is typically one of immune tolerance, which is mediated by antigen presenting cells (APCs), such as CD103⁺ dendritic cells (DCs) in the gastrointestinal tract and CD11b⁺ dermal DCs and Langerhans cells in the skin [2, 7]. These APCs capture the antigen and go to lymph nodes, where they present allergen-derived peptides through the major histocompatibility class II molecule (MHC class II) to T cells. In healthy conditions they induce the generation of regulatory T (Treg) cells, which perform an active suppression of immune response (Fig. 1). In patients with food allergy, induction of Treg cells is believed to be compromised and, this results in the generation of antigen-specific Th2 cells that drive IgE class-switching and expansion of allergic effector cells [2, 7, 8]. Murine model systems have provided valuable information about the role of cytokines in this process, such as IL-33 and IL-9, and play an important role in driving the Th2 skewing [9, 10]. IL-33 promotes food allergy through different mechanisms, such as to induce OX40 ligand expression on APCs and to activate the group 2 innate lymphoid cells, both suppress the generation of Treg cells. In addition, IL-33 contributes to acute reactions to food by acting directly on mast cells and enhancing IgE-mediated activation. IL-9 is a growth factor for mast cells and has been shown in mouse models to play an essential role in the pathogenesis of food allergy. The allergen-specific IgE binds to the high-affinity IgE receptor FcεRI on the surface of basophils and mast cells. When an individual is exposed to the same allergen again, the allergen induces the cross-linking of two IgE molecules and mast cells and basophils become activated to release inflammatory mediators like cytokines, histamine, and leukotrienes, which lead to immediate allergic inflammation (Fig. 1). Furthermore, the activation of mast cells through IgE leads to suppression of Treg cell generation and amplification of the Th2 response. Recently, data obtained from animal models and humans support the hypothesis that microbiota also plays an important role in the development of food allergy. The functional epithelial barrier engaged in intimate interplay with innate immune cells and the

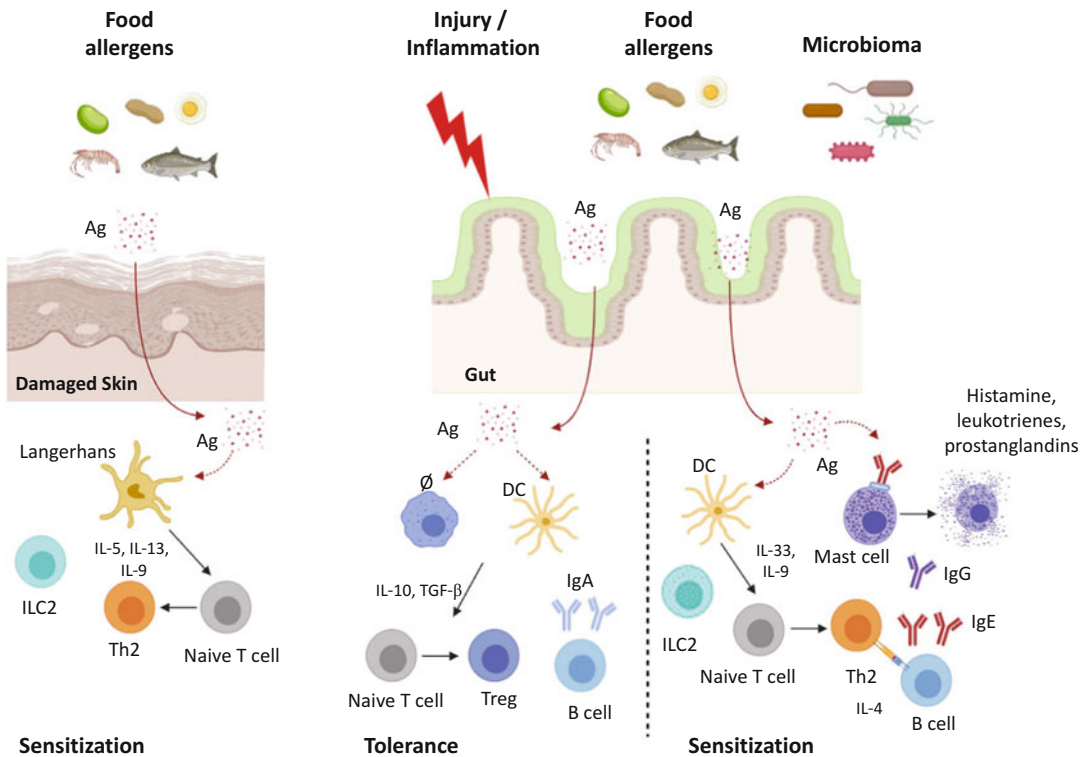


Fig. 1 Overview of sensitization and tolerance in food allergy in skin and gut. Tolerance is an active process: dendritic cells (DC) and macrophages (Ø) capture the antigen which will be presented to naive T cells. The DCs and Ø will induce the differentiation of naive T cells into regulatory T cells (Treg) that migrate to the gut. The production of IL10, TGF-β, and retinoic acid facilitates this process and induces the production of IgA by B cells. The sensitization or breakdown of tolerance to food antigens (Ag) can occur in different tissues, such as skin and gut. A disrupted skin barrier can provide a pathway for the sensitization through the pro-inflammatory action of epithelial cells, innate type 2 lymphoid cells (ILC2), and Langerhans dendritic cells, which secrete cytokines that shift naive T cells differentiation into Th2 phenotype and inhibit Treg. These Th2 can migrate to gut. Th2 cytokines drive the isotypic switching of B lymphocytes to food antigen-specific IgE that anchors to the surface of mast cells. A second contact with the Ag leads to degranulation of mast cells and allergic reaction and symptoms. In gut, there is a similar process driven by DCs and ILC2 cells. Created with BioRender.com

resident microbiota is critical to establishing and maintaining oral tolerance [11, 12].

The majority of food allergens are proteins. Class I food allergens are 10–70 kDa animal or plant proteins that are resistant to food processing and enzymatic digestion and are capable of inducing IgE sensitization via ingestion, such as casein and whey protein in cow's milk, ovalbumin and ovomucoid in egg, and parvalbumin in fish [13]. Class II food allergens are pollen-homologous proteins that are very unstable when subjected to heating or enzymatic digestion, such as in apple, carrot, and celery. Because of their susceptibility to digestion, class II food allergens are considered to induce sensitization via inhalation or topical contact. Rarely,

complex carbohydrates cause food allergy. One example is galactose- α -1,3-galactose (also known as alpha-gal), which can induce delayed anaphylactic reactions after its ingestion via mammalian meat. In this case, patients have usually reported the intake of red meat without problems for many years, and it has been proposed that sensitization is initiated through tick bites, which after feeding on mammals can transmit the antigens through the saliva [11, 14].

Major advances have been achieved in basic, translational, and clinical research, improving our mechanistic understanding of food allergy. New diagnostic tools are being developed, including recombinant allergens for *in vitro* diagnostics and cellular testing [1–4]. Despite such advances, the only proven and effective treatment for this type of hypersensitivity is to conduct a diet free of the allergenic food and their derivatives, which is not always possible and can lead to accidental reactions. However, exciting therapies are in clinical development, including treatment with biological therapies as well as allergen-specific therapies.

The allergen-specific immunotherapy (AIT), based on the administration of repeated doses of allergen (either by an injectable or sublingual administration) to induce immunological tolerance, is a common practice. Allergen-specific immunotherapy (AIT) aims to modulate the immune system against an allergen, thus modifying the natural course of the allergic disease and even having long-lasting benefits. In cases where long-lasting tolerance is developed, it is then possible to discontinue the therapy. The allergen extracts can be obtained from different sources such as cat hair and pelt, mites, different types of pollens, venom protein, food, etc. AIT is currently the only effective therapy to stop the progression of allergy, because treatment with anti-histaminic drugs or corticoids only mitigates the symptoms. Mechanisms of AIT efficacy include the activation of antigen-specific Tregs and IL-10-producing Breg (Br1) subtype cells, together with the anergy of Th2 cells and the production of allergen-specific IgG antibodies that can compete with IgE for binding to the allergens. However, the use of crude extracts in this vaccine has several disadvantages; it could induce severe anaphylactic side reactions or lead to sensitization towards new allergens present in the mixture [15, 16]. Furthermore, AIT generally requires complicated up-dosing schedules and repeated administration which lead to poor patient compliance. Different strategies have been designed to try to overcome these negative effects. They are based on the correct characterization of the allergen at the molecular level, and some of these allergens are now being produced by recombinant technologies, nucleic acid-based strategies, or synthetic peptide chemistry [17–21].

Thus, the emerging use of allergen-derived T cell and B cell epitope databases has greatly facilitated the development of epitope-based immunotherapies that aims to modulate the immune response of the patients towards specific allergens. Different

computer programs that predict specific epitopes have also been developed [22, 23]. Currently, it is feasible to predict T cell epitopes in association with appropriate MHC molecules using specific programs, such as ProPred [24], NetMHC [25], and EpDis [26]. Moreover, it is possible to predict B cell epitopes using different *in silico* algorithms based on various physicochemical properties of amino acids such as hydrophilicity, flexibility, beta-turns, and surface accessibility [27, 28].

In this chapter, we describe a workflow for the design of food allergy vaccines using proteomic tools. The strategy is defined based on the characterization of B cell epitopes for a particular food allergen, parvalbumin, a fish-food allergen.

Fish-sensitive patients are commonly allergic to numerous fish species [29]. This allergy is typically life-long affecting up to 5% of all children and 2% of adults, but can reach up to 8% among fish processing workers. The symptoms of this type of allergy appear within 60 min of exposure and include acute and generalized urticaria, nausea, vomiting, abdominal cramps, diarrhea, wheezing, and asthma. In the most severe cases, anaphylaxis shocks can be potentially life threatening. It is worthy to note that allergens responsible for shellfish and fish allergy are different, but are often discussed together due to nutritional habits [13, 29].

Different fish allergens have been characterized such as the hormone vitellogenin, collagen, and enzymes such as enolase and aldolase. In Table 1, those fish allergens recognized by the International Union of Immunological Societies Allergen Nomenclature Subcommittee database (www.allergen.org) are shown. Studies have supported that allergic properties of these proteins are related with their resistance to heat and to certain gastrointestinal enzymes [30, 31]; however, more research is required to demonstrate the correlation between *in vivo* human digestion of proteins and their allergenic status [32]. Parvalbumin (PRVB), a small calcium binding protein (between 10 and 15 kDa), is the major allergen identified in most fish species, which is abundant in fish muscle [29, 33]. Most fish PRVBs belong to the beta-subtype (β -PRVB) [33]. Furthermore, more than two isoforms of β -PRVBs are expressed by one fish species, which can differ in amino acid sequence as demonstrated for Atlantic salmon (*Salmo salar*) where their β 1 and β 2 isoforms have only 64% identity [29, 34] or in their polymeric structure. These differences can result in an allergic patient reacting to one isoform more than another, which adds to the complexity of diagnosing fish allergy and detecting allergenic parvalbumin.

Table 1

Fish protein allergens recognized by the International Union of Immunological Societies Allergen Nomenclature Subcommittee database (www.allergen.org)

Protein	Fish	Allergen name	Molecular weight (kDa)	Isoallergen and variants	GenBank protein
Beta-parvalbumin	<i>Clupea harengus</i> (Atlantic herring)	Clu h 1	12	Clu h 1.0101	CAQ72970
				Clu h 1.0201	CAQ72971
				Clu h 1.0301	CAQ72972
	<i>Ctenopharyngodon idella</i> (grass carp)	Cten i 1	9	Cten i 1.0101	
	<i>Cyprinus carpio</i> (common carp)	Cyp c 1	12	Cyp c 1.0101	CAC83658
				Cyp c 1.0201	CAC83659
	<i>Gadus callarias</i> (Baltic cod)	Gad c 1	12	Gad c 1.0101	P02622
	<i>Gadus morhua</i> (Atlantic cod)	Gad m 1	12	Gad m 1.0101	AAK63086
				Gad m 1.0102	CAM56785
				Gad m 1.0201	AAK63087
				Gad m 1.0202	CAM56786
	<i>Lates calcarifer</i> (barramundi)	Lat c 1	11,5	Lat c 1.0101	AAV97933
				Lat c 1.0201	AHW83198
					AAT45383
	<i>Lepidorhombus whiffiagonis</i> (megrim, whiff, turbot fish)	Lep w 1	11,5	Lep w 1.0101	CAP17694
	<i>Oncorhynchus mykiss</i> (rainbow trout)	Onc m 1	12	Onc m 1.0101	P86431
				Onc m 1.0201	P86432
	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 1	11	Pan h 1.0101	XP_026772003.1
				Pan h 1.0201	XP_026803769.1
	<i>Rastrelliger kanagurta</i> (Indian mackerel) ^a	Ras k 1	11,3	Ras k 1.0101	ANW10058
	<i>Salmo salar</i> (Atlantic salmon)	Sal s 1	12	Sal s 1.0101	CAA66403
	<i>Sardinops sagax</i> (Pacific pilchard)	Sar sa 1	12	Sar sa 1.0101	CAQ68366

(continued)

Table 1
(continued)

Protein	Fish	Allergen name	Molecular weight (kDa)	Isoallergen and variants	GenBank protein
	<i>Scomber scombrus</i> (Atlantic mackerel) ^a	Sco s 1	12	Sco s 1	
	<i>Sebastes marinus</i> (ocean perch, redfish, snapper)	Seb m 1	11	Seb m 1.0101	CAQ72968 CAQ72969
	<i>Thunnus albacares</i> (yellowfin tuna)	Thu a 1	11	Seb m 1.0201 Thu a 1.0101	CAQ72967
	<i>Xiphias gladius</i> (swordfish)	Xip g 1	11,5	Xip g 1.0101	CAR48256
Beta-enolase	<i>Gadus morhua</i> (Atlantic cod)	Gad m 2	47,3	Gad m 2.0101	B3A0L6
	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 2	50	Pan h 2.0101	XP_026777250.1
	<i>Salmo salar</i> (Atlantic salmon)	Sal s 2	47,3	Sal s 2.0101	ACH70931 ACH70932
	<i>Thunnus albacares</i> (yellowfin tuna)	Thu a 2	50	Thu a 2.0101	P86978
Aldolase A	<i>Gadus morhua</i> (Atlantic cod)	Gad m 3	40	Gad m 3.0101	P86980
	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 3	40	Pan h 3.0101	XP_026771637.1
	<i>Salmo salar</i> (Atlantic salmon)	Sal s 3	40	Sal s 3.0101	ACH70901
	<i>Thunnus albacares</i> (yellowfin tuna)	Thu a 3	40	Thu a 3.0101	P86979
Collagen alpha	<i>Lates calcarifer</i> (barramundi)	Lat c 6	130–140	Lat c 6.0101 Lat c 6.0201 Lat c 6.0301	XP_018521723.1 XP_018553992.1 XP_018522130.1
	<i>Salmo salar</i> (Atlantic salmon)	Sal s 6	130–140	Sal s 6.0101 Sal s 6.0102 Sal s 6.0201 Sal s 6.0202	XP_014059932.1 XP_014048044.1 XP_013998297.1 XP_014033985.1
Beta-prime-component of vitellogenin	<i>Oncorhynchus keta</i> (chum salmon)	Onc k 5	18	Onc k 5.0101	BAJ07603

(continued)

Table 1
(continued)

Protein	Fish	Allergen name	Molecular weight (kDa)	Isoallergen and variants	GenBank protein
Tropomyosin	<i>Oreochromis mossambicus</i> (Mozambique tilapia)	Ore m 4	33	Ore m 4.0101	AFV53352
	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 4	35	Pan h 4.0101	XP_026781482 XP_026775428.1
	<i>Salmo salar</i> (Atlantic salmon)	Sal s 4	37	Pan h 4.0201 Sal s 4.0101	NP_001117128.1
Creatine kinase	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 7	43	Pan h 7.0101	XP_026780620.1
	<i>Salmo salar</i> (Atlantic salmon)	Sal s 7	43	Sal s 7.01	ACH70914.1
Triosephosphate isomerase	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 8	25	Pan h 8.0101	XP_026795867.1
	<i>Salmo salar</i> (Atlantic salmon)	Sal s 8	25	Sal s 8.01	ACM09737.1
Pyruvate kinase PKM-like	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 9	65	Pan h 9.0101	XP_026775867.1
L-lactate dehydrogenase	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 10	34	Pan h 10.0101	P_026774991.1
Glucose 6-phosphate isomerase	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 11	60	Pan h 11.0101	XP_026782721.1
Glyceraldehyde-3-phosphate dehydrogenase	<i>Pangasianodon hypophthalmus</i> (striped catfish)	Pan h 13	36	Pan h 13.0101	XP_026782131.1

^aThe allergen is registered as parvalbumin

The workflow introduced in this chapter comprises five consecutive steps: (1) shotgun proteomics analysis of different protein β -PRVB isoforms, (2) downloading all protein sequences for β -PRVBs included in UniProtKB database, (3) analysis by protein-based bioinformatics of B cell epitopes, (4) synthesizing the selected B cell peptide epitopes, and (5) performing of immunoassays using sera from healthy and allergic patients. The results from this method provide a rationale repository of B cell epitopes for the design of new specific immunotherapies for a particular food allergen. Using this workflow, a total of 35 peptides were identified as B

cell epitopes, among them the top 4 B cell peptide epitopes that may induce protective immune response were selected as potential peptide vaccine candidates for fish allergy.

In vitro and in vivo studies have shown that B cell epitope-based allergy vaccines present a reduced potential to activate allergen-specific T cells and to induce pro-inflammatory cytokines [35, 36]. The B cell peptide epitopes can be synthesized and administered as a mixture of peptides combined with specific adjuvants in order to induce oral tolerance or anergy [37]. B cell epitopes identified and validated by ELISA could be chemically coupled to a carrier protein or be produced as recombinant fusion proteins and administered to achieve these objectives. One classical protein carrier is the keyhole limpet hemocyanin, which induces allergen-specific IgG against allergens [38]. 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 [39, 40]. The results of different works, including clinical trials, confirmed a good safety profile for this kind of vaccines and its capacity of inducing a highly selective allergen-specific IgG response that does not boost allergen-specific IgE responses [35, 36]. Advances in nanotechnology and polymer design have also enabled the design of new delivery systems for immunization therapy that are also suitable for mucosal and oral administration routes. Therefore, shotgun proteomics, protein-based bioinformatics, and IgE reactive approaches provide a good strategy for the global characterization of B cell epitopes for food allergen. The generation of B cell epitopes database will be very useful for further development of new therapeutic treatments based on peptide vaccines for food allergy.

2 Materials

2.1 Reference Food/ Fish Samples

1. Specimens of the most consumed fish species that cause fish allergy.
2. FishID Kit (Bionostra, Madrid, Spain).

2.2 Protein Extraction

1. Lysis buffer: 10 mM Tris-HCl, pH 7.2 with 5 mM phenyl-methylsulfonyl fluoride (PMFS). Prepared at the moment and stored at 4 °C.
2. Ultra-Turrax device (IKA-Werke, Staufen, Germany).
3. Centrifuge.

2.3 Allergen Purification

1. Incubator set at 100 °C.
2. Centrifuge.

2.4 Protein Concentration

1. Bicinchoninic acid (BCA) method.
2. Centrifuge.
3. Microplate reader.

2.5 Protein Digestion with Trypsin

1. 8 M Urea.
2. 100 mM TCEP (Tris(2-carboxyethyl)phosphine hydrochloride).
3. 500 mM IAA (Iodoacetamide).
4. Sequencing grade modified trypsin.
5. 25 mM ammonium bicarbonate pH 8.0.
6. Incubator set at 37 °C.
7. 5% Formic acid (FA) grade for liquid chromatography.
8. SpeedVac concentrator.
9. C18 MicroSpin™ column.

2.6 Shotgun Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS) Analysis

1. Proxeon EASY-nLC II LC system (Thermo Fisher Scientific, San Jose, CA, USA) coupled to a LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific).
2. Reverse phase (RP) column (75 µm × 10 cm) packed in-house with C₁₈ resin (Magic C₁₈ AQ 3 µm; Michrom BioResources, Auburn, CA, USA).
3. Mobile phase A: 0.1% FA in Milli-Q water.
4. Mobile phase B: 0.1% FA in 98% Acetonitrile (ACN).

2.7 MS Data Processing

1. Proteome Discoverer 2.4 package (Thermo Fisher Scientific) implemented with SEQUEST-HT license.
2. UniProt/TrEMBL database downloaded from Uniprot.org.
3. PeptideProphet algorithm, https://en.wikipedia.org/wiki/Trans-Proteomic_Pipeline.

2.8 Bioinformatics Analysis of Allergen Sequences and B Cell Epitopes

1. BLASTp program against protein sequences registered in the NCBI database.
2. Clustal W (BioEdit, <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>).
3. Kolaskar and Tongaonkar on the Immune Epitope Database and Analysis Resource (IEDB) (<http://tools.immuneepitope.org/bcell/>).

2.9 Synthetic Peptide Epitopes

1. Track Peptide Libraries (JPT Peptide Technologies GmbH, Berlin, Germany).
2. Phosphate-Buffered Saline (PBS).

2.10 Sera from Different Fish Allergic Patients and Healthy Donors

1. Sera from fish allergic patients and healthy controls obtained from hospitals.
2. Fish-specific IgE antibodies test.
3. Informed consent from all subjects and all protocols approved by the board of the local institutional authority (CEIC).

2.11 Enzyme-Linked Immunosorbent Assay (ELISA)

1. 96-well high-binding plates.
2. Phosphate-Buffered Saline (PBS).
3. Blocking buffer: PBS/1% BSA (Bovine Serum Albumin) (w/v).
4. Washing buffer: PBS/0.05% Tween 20, pH 7.4.
5. Sera from different fish allergic patients and healthy donors.
6. Goat anti-human IgE-HRP (horseradish peroxidase).
7. Incubator at 37 °C.
8. *o*-Phenylenediamine (OPD).
9. 3 M sulfuric acid.
10. ELISA Microplate Reader.

2.12 Statistical Analysis

1. Statistical package R version 3.4.1 (<http://www.r-project.org>).

2.13 3D Structural Modeling

1. Swiss-Model server (<http://swissmodel.expasy.org>).
2. Swiss-PdbViewer 4.1.

3 Methods

The method was optimized for beta-parvalbumins (β -PRVBs), which are considered as the major fish allergens [41] (*see Note 1*).

3.1 Shotgun Proteomics Analysis of β -PRVBs**3.1.1 Reference Fish Species**

1. Acquire from marketplaces specimens of different fish species (i.e., 15 different fish species) to comprise the most consumed fish species that cause fish allergy.
2. Identify genetically the fish species using the FishID Kit (Bio-nostra, Madrid, Spain) (*see Note 2*).

3.1.2 Allergen Protein Extraction and Purification

1. Homogenize 0.5 g of white muscle of the selected fish species in 4 mL of lysis buffer for 2 min in an Ultra-Turrax device.
2. Purify the β -PRVBs by taking advantage of their thermostability [42].

3. Centrifuge protein extracts at $40,000 \times g$ for 20 min at 4 °C and purify the supernatants into new tubes.
4. Measure protein concentration in the extracts using BCA, Bradford, or Lowry method. The concentration is expected to be approximately 1–5 mg/mL. If sample lysates need to be stored at this point, store them at –80 °C until use.

3.1.3 Trypsin Protein Digestion

1. Dissolve a total of 100 µg of protein in 8 M urea and 25 mM of ammonium bicarbonate pH 8.0. Add TCEP to a final concentration of 5 mM for 30 min at 37 °C. Add IAA to a final concentration of 50 mM incubate 60 min in dark. Dilute 4× the sample with 25 mM of ammonium bicarbonate pH 8.25 (*see Note 2*).
2. Digest 100 µg of protein with 1 µg of trypsin (1:100 protease to protein) in 25 mM ammonium bicarbonate pH 8.0 overnight at 37 °C (*see Note 3*).
3. Acidify the peptide digest by addition of 5% FA to reduce the pH of the solution to ~2.7.
4. Clean the samples on C18 MicroSpin™ columns and dry in a SpeedVac.

3.1.4 LC-MS/MS Analysis

1. Analyze peptide digests by LC-MS/MS in a Proxeon EASY-nLC II LC system coupled to a LTQ-Orbitrap XL mass spectrometer (*see Note 4*).
2. Separate a total of 1 µg of peptides on an RP column (75 µm × 10 cm) packed with C₁₈ resin (Magic C₁₈ AQ 3 µm) using 0.1% FA in Milli-Q water and 0.1% FA in 98% ACN as mobile phases A and B, respectively. Use a 60-min linear gradient from 5 to 35% B at a flow rate of 300 nL/min.
3. For ionization follow the parameters: 1.95 kV spray voltage and 230 °C capillary temperature. Peptides are analyzed in positive ion mode, and complete high-resolution full scans are obtained from 400 to 1600 amu (1 µscan) in the Orbitrap analyzer, followed by four data-dependent collision-induced dissociation (CID) MS/MS scans (1 µscans) using an isolation width of 3 amu and a normalized collision energy of 35%.
4. Fragmented masses are assigned during dynamic exclusion for 30 s after the second fragmentation event, and unassigned charged ions are excluded from the MS/MS analysis.

3.1.5 MS Data Processing

1. Analyze MS/MS spectra using SEQUEST-HT™ (Proteome Discoverer 2.4 package) against the Teleostei UniProt/TrEMBL protein database.
2. Use the following constraints for the searches: semi-tryptic cleavage with up to two missed cleavage sites and tolerances

of 1.2 Da for precursor ions and 0.5 Da for MS/MS fragment ions. The variable modifications allowed are carbamidomethylation (+57.021465 Da) on cysteine, methionine oxidation (+15.99492 Da) and acetylation (+42.01057 Da) on the N-terminus of the protein.

3. Finally, the results are subjected to statistical analysis with the PeptideProphet algorithm to ensure that the false discovery rate (FDR) was less than 1%.

3.2 Bioinformatics Analysis of β -PRVB Sequences

1. Analyze all β -PRVB sequences identified by shotgun proteomics using BLASTp algorithm to determine homologies with proteins registered in the NCBI database.
2. Download in FASTA format all β -PRVBs for all Teleostei species ($n = 98$ protein sequences).
3. Align all sequences with Clustal W.

3.3 Bioinformatics Analysis of B Cell Epitopes

1. Compute linear B cell epitopes for each β -PRVB protein sequence using the Kolaskar and Tongaonkar method included in the IEDB.
2. Select a window size of five to seven amino acids that is appropriate for finding regions that may potentially be antigenic [43]. The features of the B cell epitopes include surface accessibility, flexibility, and hydrophilicity [44]. A total of 35 different B cell epitopes with seven amino acids are identified for all β -PRVBs [41]. Among them the 12 B cell epitopes with the highest percentage score determined using the Kolaskar and Tongaonkar method are synthesized to validate the predicted B cell peptide epitopes.

3.4 Synthesis of Selected B Cell Peptide Epitopes

1. Synthesize the selected B cell epitopes (seven amino acid residues) and negative control peptides (seven amino acid residues with a shuffled amino acid sequence) with more than 70% of purity using Track Peptide Libraries, JPT Peptide Technologies GmbH, Berlin, Germany.
2. Reconstitute the lyophilized peptides individually with $1 \times$ PBS and stored at -80°C .

3.5 Sera from Different Fish Allergic Patients and Healthy Donors

1. Collect sera from allergic patients with a clinical history of allergic reactions to fish and sera from healthy controls with a negative clinical history for any Type-I allergy from the hospitals.
2. Confirm the diagnosis of IgE-mediated fish allergy by measuring fish-specific IgE antibodies.
3. Provide the informed consent from all subjects and all protocols approved by the board of the local institutional authority.

3.6 Immunoassay Using Sera from Healthy and Allergic Patients by ELISA

1. Dilute each B cell synthetic peptide or negative control peptide in PBS at a final concentration of 20 µg/mL.
2. Coat the wells of a 96-well ELISA plate with 50 µL of the corresponding peptide and incubate overnight at 4 °C.
3. Remove the coating solution and block the remaining binding sites by adding 200 µL of blocking buffer per well and incubate for 2 h at 37 °C.
4. Wash the plate three times with 50 µL of washing buffer.
5. Incubate 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.
6. Wash the plate three times with 50 µL of washing buffer.
7. Detect bound IgE antibodies using 50 µL per well of goat anti-human IgE-HRP (1:1000 in PBS/1% BSA) for 1.5 h at 37 °C.
8. Wash the plate three times with 50 µL of washing buffer.
9. Add 50 µL/well of freshly prepared OPD, the horseradish peroxidase substrate, incubate at room temperature for 5–20 min, depending on the rate of color development.
10. Stop the reaction by adding 50 µL/well of 3 M sulfuric acid.
11. Read the optical density (OD) at 492 nm using an ELISA microplate reader. Perform the analyses in triplicate.
12. Finally, correct the OD value obtained for each B cell synthetic peptide by subtracting the OD value determined in negative control peptide samples. Report the mean intensities for each peptide epitope.

3.7 Statistical Analysis

1. Evaluate the ELISA data using statistical package R.
2. Make heat map and hierarchical clustering using the function `heatmap.2()` on R, using the Euclidean distance and the complete linkage method.
3. Perform Box-plot graphics using the `boxplot()` command showing the outliers on R package.
4. Select the top B cell peptide epitopes based on the highest percentage score determined using the Kolaskar and Tongaonkar algorithm and the IgE immunoassay results to be considered as potential B cell peptide vaccine candidates. A total of 4 top B cell peptide epitopes for all β-PRVBs are selected as potential B cell peptide vaccine candidates [41].

3.8 3D Structural Modeling

1. Use the Swiss-Model server (<http://swissmodel.expasy.org>) to create the 3D models and the Swiss-PdbViewer 4.1 to visualize the 3D structure.

2. Use as a protein template the Accession Number of the corresponding β -PRVBs. As a result the 4 top B cell peptide epitopes for all β -PRVBs are considered good peptide vaccine candidates because they are localized in the surface of the β -PRVBs. These peptide vaccine candidates further need to be validated using new experiments with cell lines and animal models to obtain an efficient and safe immunotherapy for fish allergy [41] (*see Note 5*).

4 Notes

1. Different food allergens and different lysis buffer need to be optimized.
2. TCEP (Tris (2-carboxyethyl) phosphine hydrochloride) is used to stabilize proteins, to reduce disulfide bonds, and to maintain monothiols in a reduced state. This can be substituted by DTT (DL-Dithiothreitol). IAA (Iodoacetamide) is a sulfhydryl-reactive alkylating reagent used to bind covalently with the thiol group of cysteines in order to avoid disulfide bond formation for protein characterization and peptide mapping.
3. Maximal trypsin activity occurs at pH 7–9; the enzyme is reversibly inactivated at pH < 4. If the pH of the sample is too acidic/alkaline is too acidic add Tris-HCl buffer pH 8 to the sample and check the pH again.
4. Different LC-MS instruments from different commercial brands may require specific parameters for peptide ionization and fragmentation.
5. All these potential peptide vaccine candidates need to be validated by further bioassays using cell lines and animal models. However, compared with the classical approaches, the proteomics and bioinformatics methods are faster and low-cost alternatives that predict and reduce the number of potential targets to be investigated.

Acknowledgments

This work was supported by the GAIN-Xunta de Galicia Project (IN607D 2017/01) and the Spanish AEI/EU-FEDER PID2019-103845RB-C21 project. Dr. Mónica Carrera is supported by the Ramón y Cajal contract (Ministry of Science and Innovation of Spain). Dr. Susana Magadan is supported by the Talent Retention program of Universidade de Vigo.

References

1. Turnbull JL, Adams HN, Gorard DA (2015) Review article: the diagnosis and management of food allergy and food intolerances. *Aliment Pharmacol Ther* 41:3–25
2. Yu W, Freeland DMH, Nadeau KC (2016) Food allergy: immune mechanisms, diagnosis and immunotherapy. *Nat Rev Immunol* 16:751–765
3. Sicherer SH, Sampson HA (2014) Food allergy: epidemiology, pathogenesis, diagnosis, and treatment. *J Allergy Clin Immunol* 133:291–308
4. Sicherer SH, Sampson HA (2018) Food allergy: a review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. *J Allergy Clin Immunol* 141:41–58
5. Ben-Shoshan M, Turnbull E, Clarke A (2012) Food allergy: temporal trends and determinants. *Curr Allergy Asthma Rep* 12:346–372
6. Savage J, Johns CB (2015) Food allergy: epidemiology and natural history. *Immunol Allergy Clin North Am* 35:45–59
7. Satitsuksanoa P, Jansen K, Głobińska A, van de Veen W, Akdis M (2018) Regulatory immune mechanisms in tolerance to food allergy. *Front Immunol* 9:2939
8. Chinthrajah RS, Hernandez JD, Boyd SD, Galli SJ, Nadeau KC (2016) Molecular and cellular mechanisms of food allergy and food tolerance. *J Allergy Clin Immunol* 137:984–997
9. Schülke S, Albrecht M (2019) Mouse models for food allergies: where do we stand? *Cells* 8:546
10. Kanagaratham C, Sallis BF, Fiebiger E (2018) Experimental models for studying food allergy. *CMGH* 6:356–369.e1
11. De Martinis M, Sirufo MM, Suppa M, Ginaldi L (2020) New perspectives in food allergy. *Int J Mol Sci* 21(4):1474
12. Iweala OI, Nagler CR (2019) The microbiome and food allergy. *Annu Rev Immunol* 37:377–403
13. Leung NYH, Wai CYY, Shu S, Wang J, Kenny TP, Chu KH, Leung PSC (2014) Current immunological and molecular biological perspectives on seafood allergy: a comprehensive review. *Clin Rev Allergy Immunol* 46:180–197
14. Commins SP, Jerath MR, Cox K, Erickson LD, Platts-Mills T (2016) Delayed anaphylaxis to alpha-gal, an oligosaccharide in mammalian meat. *Allergol Int* 65:16–20
15. Dreskin SC, Halsey NA, Kelso JM et al (2016) International Consensus (ICON): allergic reactions to vaccines. *World Allergy Organ J* 9(1):32
16. Valenta R, Karaulov A, Niederberger V et al (2018) Allergen extracts for in vivo diagnosis and treatment of allergy: is there a future? *J Allergy Clin Immunol Pract* 6:1845–1855.e2
17. Zhernov Y, Curin M, Khaitov M, Karaulov A, Valenta R (2019) Recombinant allergens for immunotherapy. *Curr Opin Allergy Clin Immunol* 19:402–414
18. Valenta R, Campana R, Focke-Tejkl M, Niederberger V (2016) Vaccine development for allergen-specific immunotherapy based on recombinant allergens and synthetic allergen peptides: lessons from the past and novel mechanisms of action for the future. *J Allergy Clin Immunol* 137:351–357
19. Valenta R, Campana R, Niederberger V (2017) Recombinant allergy vaccines based on allergen-derived B cell epitopes. *Immunol Lett* 189:19–26
20. O’Hehir RE, Prickett SR, Rolland JM (2016) T cell epitope peptide therapy for allergic diseases. *Curr Allergy Asthma Rep* 16:14
21. Woodfolk JA (2006) Epitope-specific T-cell responses and allergic phenotypes: implications for T-cell peptide therapy. *Expert Rev Clin Immunol* 2:309–318
22. Salimi N, Fleri W, Peters B, Sette A (2010) Design and utilization of epitope-based databases and predictive tools. *Immunogenetics* 62:185–196
23. Martini S, Nielsen M, Peters B, Sette A (2020) The immune epitope database and analysis resource program 2003–2018: reflections and outlook. *Immunogenetics* 72:57–76
24. Singh H, Raghava GP (2001) ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 17(12):1236–1237
25. Nielsen M, Lundegaard C, Lund O (2007) Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics* 8:238
26. Jandrić DR, Lazić GM, Mitić NS, Pavlović MD (2016) Software tools for simultaneous data visualization and T cell epitopes and disorder prediction in proteins. *J Biomed Inform* 60:120–131
27. Yao B, Zheng D, Liang S, Zhang C (2013) Conformational B-cell epitope prediction on antigen protein structures: a review of current algorithms and comparison with common binding site prediction methods. *PLoS One* 8(4):e62249

28. Potocnakova L, Bhide M, Pulzova LB (2016) An introduction to B-cell epitope mapping and in silico epitope prediction. *J Immunol Res* 2016:6760830
29. Sharp MF, Lopata AL (2014) Fish allergy: in review. *Clin Rev Allergy Immunol* 46:258–271
30. Astwood JD, Leach JN, Fuchs RL (1996) Stability of food allergens to digestion in vitro. *Nat Biotechnol* 14:1269–1273
31. Arif SH, Jabeen M, Hasnain A-U (2007) biochemical characterization and thermostable capacity of parvalbumins: the major fish-food allergens. *J Food Biochem* 31:121–137
32. Herman RA, Roper JM, Zhang JXQ (2020) Evidence runs contrary to digestive stability predicting protein allergenicity. *Transgenic Res* 29:105–107
33. Kuehn A, Swoboda I, Arumugam K, Hilger C, Hentges F (2014) Fish allergens at a glance: variable allergenicity of parvalbumins, the major fish allergens. *Front Immunol* 5:179
34. Saptarshi SR, Sharp MF, Kamath SD, Lopata AL (2014) Antibody reactivity to the major fish allergen parvalbumin is determined by isoforms and impact of thermal processing. *Food Chem* 148:321–328
35. Eckl-Dorna J, Weber M, Stanek V et al (2019) Two years of treatment with the recombinant grass pollen allergy vaccine BM32 induces a continuously increasing allergen-specific IgG4 response. *EBioMedicine* 50:421–432
36. Focke-Tejkl M, Weber M, Niespodziana K et al (2015) Development and characterization of a recombinant, hypoallergenic, peptide-based vaccine for grass pollen allergy. *J Allergy Clin Immunol* 135:1207–1217.e11
37. Cook QS, Burks AW (2018) Peptide and recombinant allergen vaccines for food allergy. *Clin Rev Allergy Immunol* 55:162–171
38. Tabesh S, Fanuel S, Fazlollahi MR, Yekaninejad MS, Kardar GA, Razavi SA (2019) Design and evaluation of a hypoallergenic peptide-based vaccine for Salsola kali allergy. *Int Immunopharmacol* 66:62–68
39. Niespodziana K, Focke-Tejkl M, Linhart B, Civaj V, Blatt K, Valent P, Van Hage M, Grönlund H, Valenta R (2011) A hypoallergenic cat vaccine based on Fel d 1-derived peptides fused to hepatitis B PreS. *J Allergy Clin Immunol* 127:1562–1570.e6
40. Edlmayr J, Niespodziana K, Linhart B et al (2009) A combination vaccine for allergy and rhinovirus infections based on rhinovirus-derived surface protein VP1 and a nonallergenic peptide of the major Timothy Grass Pollen Allergen Phl p 1. *J Immunol* 182:6298–6306
41. Carrera M, González-Fernández Á, Magadán S, Mateos J, Pedrós L, Medina I, Gallardo JM (2019) Molecular characterization of B-cell epitopes for the major fish allergen, parvalbumin, by shotgun proteomics, protein-based bioinformatics and IgE-reactive approaches. *J Proteomics* 200:123–133
42. Carrera M, Cañas B, Vázquez J, Gallardo JM (2010) Extensive de Novo sequencing of new parvalbumin isoforms using a novel combination of bottom-up proteomics, accurate molecular mass measurement by FTICR-MS, and selected MS/MS ion monitoring. *J Proteome Res* 9:4393–4406
43. Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* 276:172–174
44. Karplus PA, Schulz GE (1985) Prediction of chain flexibility in proteins - a tool for the selection of peptide antigens. *Naturwissenschaften* 72:212–213