



Article

Proteomic and Bioinformatic Investigations of Heat-Treated *Anisakis simplex* Third-Stage Larvae

Maciej Kochanowski *¹, Mirosław Różycki, Joanna Dąbrowska, Aneta Belcik, Jacek Karamon¹, Jacek Sroka and Tomasz Cencek¹

Department of Parasitology and Invasive Diseases, National Veterinary Research Institute, Partyzantów Avenue 57, 24-100 Puławy, Poland; mrozycki@piwet.pulawy.pl (M.R.); joanna.dabrowska@piwet.pulawy.pl (J.D.); aneta.belcik@piwet.pulawy.pl (A.B.); j.karamon@piwet.pulawy.pl (J.K.); jacek.sroka@piwet.pulawy.pl (J.S.); tcencek@piwet.pulawy.pl (T.C.)

* Correspondence: maciej.kochanowski@piwet.pulawy.pl; Tel.: +48-818893025

Received: 24 May 2020; Accepted: 13 July 2020; Published: 16 July 2020



Abstract: *Anisakis simplex* third-stage larvae are the main source of hidden allergens in marine fish products. Some *Anisakis* allergens are thermostable and, even highly processed, could cause hypersensitivity reactions. However, *Anisakis* proteome has not been studied under autoclaving conditions of 121 °C for 60 min, which is an important process in the food industry. The aim of the study was the identification and characterization of allergens, potential allergens, and other proteins of heat-treated *A. simplex* larvae. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was used to identify 470 proteins, including allergens—Ani s 1, Ani s 2, Ani s 3, Ani s 4, Ani s 5—and 13 potential allergens that were mainly homologs of *Anisakis* spp., *Ascaris* spp., and Acari allergens. Ani s 2, Ani s 3, Ani s 5, and three possible allergens were found among the top 25 most abundant proteins. The computational analysis allowed us to detect allergen epitopes, assign protein families, and domains as well as to annotate the localization of proteins. The predicted 3D models of proteins revealed similarities between potential allergens and homologous allergens. Despite the partial degradation of heated *A. simplex* antigens, their immunoreactivity with anti-*A. simplex* IgG antibodies was confirmed using a Western blot. In conclusion, identified epitopes of allergenic peptides highlighted that the occurrence of *Anisakis* proteins in thermally processed fish products could be a potential allergic hazard. Further studies are necessary to confirm the IgE immunoreactivity and thermostability of identified proteins.

Keywords: *Anisakis simplex*; foodborne parasite; allergen; potential allergen; hidden allergen; mass spectrometry; bioinformatics; proteome; autoclaving

1. Introduction

Foodborne parasites are one of the most important causative agents of human infectious diseases, especially in less developed countries [1,2]. The climate changes, new feeding habits, and globalization of food supply chains may increase the worldwide incidence of some foodborne diseases [3,4]. Unfortunately, foodborne parasites remain neglected compared with bacterial and viral pathogens [3]. Therefore, studies in this field, like epidemiological surveys [5,6], development of novel diagnostic tools [7–9], drug discovery [10], or investigation of pathogenicity [11], are particularly valuable.

Anisakis spp. is one of the most important fish-borne parasites [2]. Live third-stage larvae (L3) of *Anisakis simplex* consumed with fish or seafood dishes can cause a human disease called anisakiasis. Over 20,000 cases of anisakiasis had been reported worldwide before 2010 [12]. Bao et al. [13] estimated that the total number of worldwide anisakidosis (almost all anisakiasis) cases up to December 2017 might be over 76,000. According to a report of the Orphanet (the portal of rare diseases and orphan

drugs; January 2020), the worldwide incidence of anisakiasis is 0.32/100,000 [14,15]. However, due to very nonspecific symptoms of the disease, the number of detected cases is considered to be highly underestimated [16]. These assumptions confirm the results of the last study in which the number of anisakiasis cases requiring medical attention in Spain has been estimated at around 8000 per year [17].

A. simplex L3 larvae are also a known source of allergens, and 14 of them (Ani s 1–Ani s 14) are officially recognized by the World Health Organization and International Union of Immunological Societies (WHO/IUIS) Allergen Nomenclature Sub-committee. The list of *Anisakis* allergens seems incomplete as new isoforms [18] and new potential allergens [19–23] are still being identified. *A. simplex* allergens are considered to be the most common hidden allergens in marine fish products [24]. These allergens may cause hypersensitivity reactions in sensitized humans in the form of urticaria, angioedema, asthma, rhinitis, conjunctivitis, gingivostomatitis, contact dermatitis, and even severe anaphylaxis [25,26]. Inhalation and contact allergy to *A. simplex* are mostly linked with occupational diseases of food industry employees, cooks, and restaurant workers [27–30]. Non-occupational airborne-induced anaphylaxis after exposure to *Anisakis* allergens has been also reported [31]. The allergic response caused by *Anisakis* could also be associated with nephrotic syndrome and rheumatic manifestations [32,33].

The prevalence of *Anisakis* hypersensitivity in symptomatic patients varies widely depending on geographical area, characteristics of the population studied, diagnostic criteria, and laboratory assays [30]. Researchers considered that, in endemic countries, the number of highly sensitized humans in the general population could be approximately 7% [34].

As several *A. simplex* allergens are known to be resistant to high and low temperatures or pepsin digestion, allergens of dead larvae may cause hypersensitivity reactions in sensitized humans [35]. Cases of *Anisakis*-related hypersensitivity reactions have been noted even after the ingestion of canned food [26] or meat of chicken fed with fishmeal containing *Anisakis* allergens [36]; however, *Anisakis* allergies after the consumption of processed food have not been well investigated. Similarly, the *A. simplex* allergome and proteome subjected to autoclaving have not been tested so far. This process is one of the most important food preservation methods in the industry and is widely used in the production of canned fish products.

The autoclaving of proteins may result in a reduction of allergenicity by the disruption of IgE epitopes but, overall, it cannot be completely abolished [37]. Furthermore, thermal processing may generate new epitopes in the proteins, leading to neoallergen formation, which could cause an enhancement of allergenicity [38,39]. Therefore, the examination of the heat-treated allergens and possible allergens is important to determine the potential risk of hypersensitivity reactions for sensitized consumers of processed food. Mass spectrometry proteomics approaches seem to be a useful tool for this purpose as they allow for the high-throughput identification of proteins and have been successfully applied to many studies for the analysis of allergens in food [40,41].

In this context, the goal of our study was proteomic and computational investigations of heat-treated *A. simplex* L3 larvae. In addition to the mass spectrometry identification of proteins, we performed an investigation of the influence of high temperature on the antigenic profiles of *A. simplex* and the immunoreactivity of *Anisakis* antigen with IgG antibodies. Subsequently, the identified proteins were processed using a bioinformatics approach to characterize proteins with a particular focus on the allergenic peptides and their epitopes.

2. Materials and Methods

2.1. Ethics Statement

Rabbits purchased from the Center for Experimental Medicine (Katowice, Poland) were housed under standard conditions, and experiments were conducted under the approval of the Local Ethical Commission for Animal Experimentation (license no: 66/2012).

2.2. *A. simplex* L3 Larvae Collection and Identification

Anisakis spp. L3 larvae were collected from marine fishes, as described previously [42]. The larvae were purified by washing with sterile 0.01 M phosphate-buffered saline solution (PBS, pH 7.4; Sigma, St. Louis, MO, USA). Then, the identification of species of *Anisakis* spp. nematodes was performed using PCR-restriction fragment length polymorphism (PCR-RFLP) [43].

2.3. Protein Extraction from *A. simplex* L3 Larvae

Three different antigens of *A. simplex* were obtained:

- the procedure for native crude (CR) antigens of *A. simplex* was performed, as previously reported [42];
- the procedure for heat-treated CR antigens of *A. simplex* was performed by the heating of native CR antigen in a thermomixer at 100 °C for 60 min;
- the procedure for heat-sterilized CR antigens of *A. simplex* was performed by the autoclaving of CR antigens at 121 °C for 60 min.

Protein concentration was determined by measuring the absorbance at 280 nm using an ultraviolet-visible (UV-vis) spectrophotometer (Implen, München, Germany) and adjusted to 1 mg/mL. Protein extracts were kept at −80 °C for further analysis.

2.4. Generation of Rabbit Anti-*A. simplex* Antiserum

Rabbits were immunized by intramuscular injection of 2 mg of *A. simplex* native CR antigen mixed with Freund's complete adjuvant (Sigma, St. Louis, MO, USA). Immunization was performed according to our previously described protocol [42]. Preimmune serum was taken before immunization and used as a negative control. Sera were stored at −80 °C until use.

2.5. Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) and IgG Western Blot (WB)

SDS-PAGE analysis of *A. simplex* antigens and WB reaction of antigens with rabbit anti-*A. simplex* serum was generally performed, as described previously [23], using 4-chloro-1-naphthol (Sigma, St. Louis, MO, USA) as a substrate for horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG antibodies (Sigma, St. Louis, MO, USA).

The molecular weight of the SDS-PAGE and WB bands were estimated using Bio-1D software (ver. 15.07; Vilber Lourmat, Marne-la-Vallée, France). ImageJ software (ver. 1.53a; National Institute of Health, Bethesda, MD, USA, <http://rsb.info.nih.gov/ij/>) was used for semi-quantitative density measurements of the SDS-PAGE and WB profile based on the integrated density (IntDen) calculation (the product of area and mean gray value) [44].

2.6. Sample Processing and Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analysis

The autoclaved *A. simplex* antigens (three independent biological replicates) were subjected to the standard procedure of tryptic digestion during which proteins were reduced with 0.5 M (5 mM f.c.) tris (2-carboxyethyl) phosphine (TCEP; Sigma, St. Louis, MO, USA) for 1 h at 60 °C, and the cysteine residues were subsequently alkylated with 200 mM (10 mM f.c.) methyl methanethiosulfonate (MMTS; Sigma, St. Louis, MO, USA) for 10 min at room temperature and finally cleaved overnight with 10 µL of 0.1 µg/µL trypsin (Promega, Madison, WI, USA) at 37 °C. The resulting peptide mixtures were loaded in equal volumes of 20 µL to a reversed-phase RP-18 pre-column (Waters, Milford, MA, USA) using 0.1% formic acid (FA; Sigma, St. Louis, MO, USA) in water as a mobile phase and then transferred to nano-high-performance liquid chromatography (nano-HPLC) RP-18 column (internal diameter 75 µm; Waters, Milford, MA, USA) using linear acetonitrile (ACN; Sigma, St. Louis, MO, USA) gradient 0–35% over 160 min in the presence of 0.1% FA at a flow rate of 250 nL/min. The nano-HPLC column outlet was coupled directly to the ion source of the Q Exactive mass spectrometer

(Thermo Electron Corp., San Jose, CA, USA) working in the regime of data-dependent MS to MS/MS switch with higher-energy collisional dissociation (HCD) type peptide fragmentation. A blank run ensuring the absence of cross-contamination from previous samples preceded each analysis.

Mass spectrometric data were preprocessed with Mascot Distiller software (ver. 2.6; Matrix Science, London, UK; <http://www.matrixscience.com/distiller.html>) and analyzed with the Mascot search engine server (ver. 2.5; Matrix Science, London, UK; <http://www.matrixscience.com/server.html>) against the A. simplex reference proteome (20,786 sequences; proteome ID: UP000036680) obtained from the Universal Protein Resource (UniProt, <http://www.uniprot.org/>). To reduce mass errors, the peptide and fragment mass tolerance settings were established separately for individual LC-MS/MS runs after a measured mass recalibration, resulting in values of 5 ppm for the parent and 0.01 Da for the fragment ions in higher-energy collisional dissociation (HCD) MS/MS mode. Peptide sequences were searched using trypsin specificity, allowing one missed cleavage; ion type was set as monoisotopic, and protein mass as unrestricted. Beta-methylthiolation of cysteine was used as a fixed modification, whereas oxidation of methionine was set as a variable modification. A score threshold for all samples was set for 50 to match the highest score threshold computed by the Mascot software. Only proteins identified in all three biological replicates were accepted. Quantification of protein abundance was performed using the exponentially modified protein abundance index (emPAI) provided by Mascot.

The mass spectrometry proteomics data were deposited to the ProteomeXchange Consortium via the PRIDE [45] partner repository with the dataset identifier PXD018059 and 10.6019/PXD018059.

2.7. Bioinformatic Analysis

The functional annotation of the identified proteins, including gene ontology (GO) and InterPro analyses, was performed using OmicsBox software (ver. 1.2.4; BioBam Bioinformatics SL, Valencia, Spain, <https://www.biobam.com/omicsbox/>) based on the Blast2GO annotation methodology [46]. Annotations were run with the default settings, as we previously described [23], and a list of annotations was filtered using nematode taxonomy to improve the prediction accuracy.

Experimentally verified epitopes in the peptides of Anisakis allergens were searched in the Immune Epitope Database (IEDB; last updated on June 14, 2020; <https://www.iedb.org/>). Allergens in which epitopes were not found in IEDB were subjected to in silico prediction for the detection of potential epitopes in the peptides of these allergens. Bioinformatic detection of potential epitopes was performed using DNASTAR Protean 3D software (ver. 17.0.2.1; DNASTAR, Madison, WI, USA). B-cell epitopes were predicted by applying a confidence threshold of 0.7. Possible epitopes of major histocompatibility complex class II (MHC II) molecules were detected using the default settings. The mapping of potential T-cell epitopes was performed by combining AMPHI and Rothbard-Taylor methods using the default settings.

The identified proteins were evaluated for putative allergenicity by searching against The Food Allergy Research and Resource Program (FARRP) AllergenOnline.org database (ver. 20; 2,171 sequences; February 10, 2020; <http://www.allergenonline.com/>) using full-length FASTA alignment (e-value cut-off: 1e-05; 70% identity match).

The ExPASy Compute pI/Mw tool (https://web.expasy.org/compute_pi/) was applied for the calculation of the theoretical isoelectric point (pI) and molecular weight (Mw) of detected proteins.

Unknown 3D structures of allergens were predicted by homology modeling using the Phyre2 server (<http://www.sbg.bio.ic.ac.uk/phyre2/>) in intensive mode (multi-template + ab initio) [47]. Known 3D structures of allergens were derived from RCSB Protein Data Bank (RCSB PDB; <https://www.rcsb.org/>). Visualization of the 3D structures of allergens and structural alignment of allergen models were performed using the PyMOL Molecular Graphics System (ver. 2.0; Schrödinger, LLC, New York, USA).

3. Results

3.1. Comparative SDS-PAGE and IgG-WB Analyses of *A. simplex* Antigens

SDS-PAGE and WB analyses were performed to investigate the influence of high temperature on the *Anisakis* antigen. Figure 1a shows the SDS-PAGE multiband profiles of native and heat-treated CR antigens of *A. simplex*. The bands' profile of antigen heated for 60 min at 100 °C was similar to the native antigen, and just a few high molecular mass bands (132–244 kDa) that were present in the native antigen were not visible in the heated antigen. The intensity of bands was slightly lower compared to the native antigen. The SDS-PAGE profile of antigen autoclaved at 121 °C for 60 min was characterized by diffused band patterns with high background and reduced number of bands compared to the native or heated (at 100 °C) antigens. However, the bands were visible at the following molecular weights: 16–18, 20, 24–26, 34, and about 60 kDa. The IntDen values of all three SDS-PAGE profiles were very similar and ranged from 7.25E+07 to 8.32E+07.

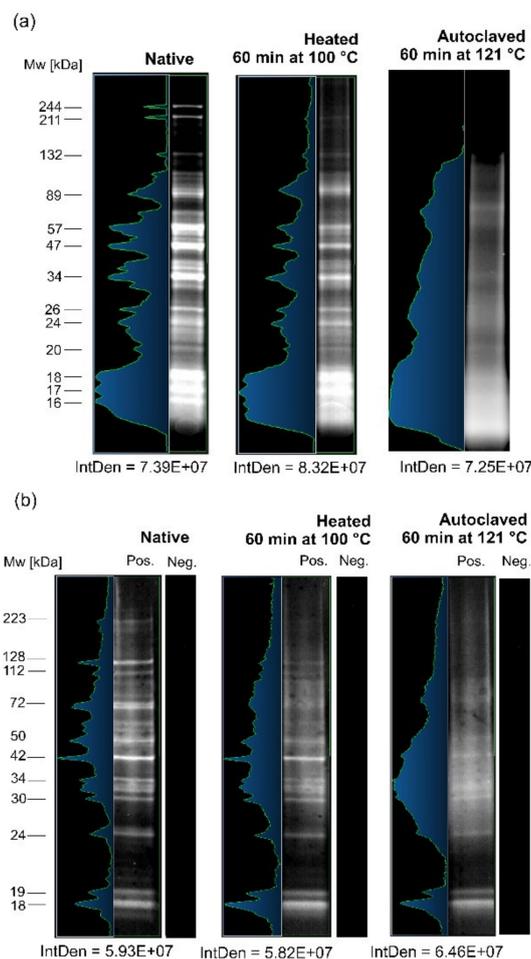


Figure 1. Colloidal Coomassie-stained SDS-PAGE analysis (color inversion mode) of the three following crude (CR) antigens of *A. simplex*: native, heated for 60 min at 100 °C, and autoclaved for 60 min at 121 °C (a). Western blot analysis of anti-*A. simplex* rabbit IgG antibodies reactivity against following CR antigens of *A. simplex*: native, heated 60 min at 100 °C, and autoclaved for 60 min at 121 °C (b). Pos.—membrane incubated with hyperimmune serum from a rabbit immunized with *A. simplex* CR antigen; Neg.—membrane incubated with rabbit preimmune serum. Molecular weight (Mw) estimations are presented in kilodaltons (kDa), as performed by Bio-1D software. The integrated density (IntDen) was calculated by ImageJ software.

The WB profiles of *Anisakis* antigens are presented in Figure 1b. The profiles of both heated antigens were generally consistent with the native antigen. Similarly to the SDS-PAGE profile, the number and intensity of bands were reduced, and a background appeared. The background in the WB profile of autoclaved antigen was slightly higher than in heated at 100 °C. The IntDen values were very similar for the WB profiles, and they were in the range from 5.82E+07 to 6.46E+07.

The background in the SDS and WB profiles of heated antigens was higher than in native antigen, probably due to degradation. However, the band pattern confirmed that degradation was only partial. Furthermore, the epitopes of degraded proteins probably were not damaged.

3.2. Identification and Characterization of *A. simplex* Proteins

A total of 470 proteins were detected in all three biological replicates of shotgun LC-MS/MS analysis. The identification was performed with high confidence as the peptides false discovery rate (FDR) calculated by Mascot in all cases was 0.99%. A detailed list of all identified proteins with UniProt IDs, protein names, gene names, and OmicsBox annotations is presented in Supplemental File S1.

Detected proteins were displayed on the 3D scatter plot (Figure 2) based on theoretical Mw, theoretical pI, and estimated relative protein abundance. The molecular weights of all proteins ranged from 4077 to 807,425 Da. An uncharacterized protein (UniProt ID: A0A0M3JJQ2) had the lowest Mw from all detected proteins, while twitchin (UniProt ID: A0A158PN23) had the highest Mw. The majority of proteins ($n = 396$) were in the range of Mw from about 10 kDa to 94.19 kDa. In the case of the estimated pI values, proteins were in the range of 4.15–10.8. The lowest pI value was calculated for troponin-like protein (UniProt ID: A0A0M3JU57), whereas the highest value was estimated for 60S ribosomal protein L8 (UniProt ID: A0A0M3K2K1). The pI values for the majority of the proteins were in one of the two following ranges: 4.15–7.19 ($n = 320$) and 8–9.59 ($n = 101$).

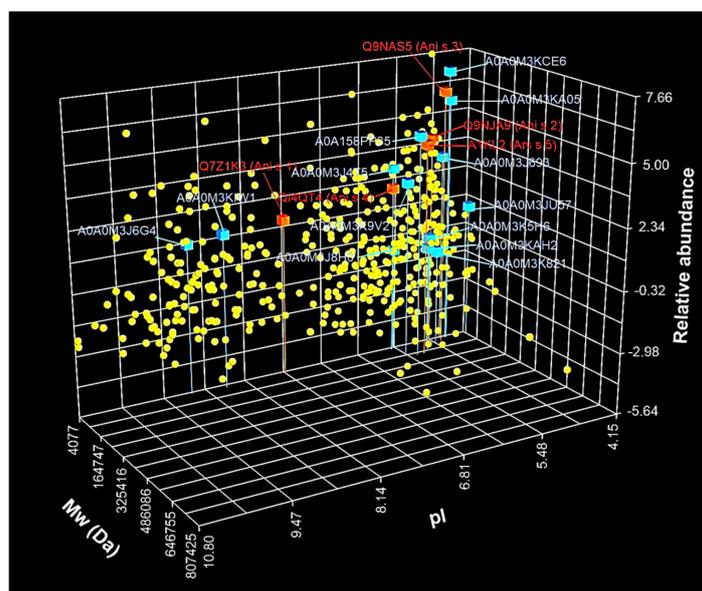


Figure 2. The three-dimensional scatter plot analysis of *A. simplex* proteins ($n = 470$) identified using LC-MS/MS. The proteins were distributed based on molecular weight (Mw), isoelectric point (pI), and the relative abundance of proteins. The Mw and pI values were calculated using the ExPASy Compute pI/Mw tool, while the relative abundance was estimated based on the average exponentially modified protein abundance index (emPAI) values (mean of three biological replicates) calculated by Mascot. The relative abundance is shown in log₂ scale. Red cubes are allergens ($n = 5$), and cyan cubes are potential allergens ($n = 13$). Cubes (allergens and potential allergens) are signed with UniProt ID and have a vertical line for better pI and Mw visualization. Yellow spheres are the other identified proteins. A scatter plot was constructed using Teraplot software (ver. 1.4.06; Kylebank Software Ltd., Ayr, UK).

The protein abundance was estimated using emPAI, which provided an approximate protein relative quantification based on the number of observed peptides divided by the number of observable peptides. The estimated emPAI values of proteins by Mascot were in the range 0.02–201.76. Among all identified proteins, the lowest emPAI values were calculated for collagen alpha-1(IV) chain (emPAI = 0.02; UniProt ID: A0A158PPJ2), uncharacterized protein (emPAI = 0.03; UniProt ID: A0A158PP53), calcium-transporting ATPase (emPAI = 0.03; UniProt ID: A0A0M3JTY4), and uncharacterized protein (emPAI = 0.03; UniProt ID: A0A0M3K3G9). While the most abundant were the following proteins: myosin essential light chain (emPAI = 201.76; UniProt ID: A0A0M3K6N3), tropomyosin (emPAI = 109.71; UniProt ID: A0A0M3KCE6), and DUF4440 domain-containing protein (emPAI = 81.05; UniProt ID: A0A0M3J349).

Table 1 shows the top 25 most abundant proteins, with emPAI values in the range of 13.78–201.76. The Mw and pI values of the majority of highly abundant proteins were in the ranges 10–29.9 kDa (n = 21) and 4.5–6.39 (n = 20), respectively. The distribution of log₂-transformed values of the emPAI of all proteins is illustrated in Figure 2. The average emPAI values were log₂-transformed to normalize the data. As shown in Figure 2, most of the log₂-transformed emPAI values (n = 413) were in the range from −4.2 to 2.99.

InterPro analysis was performed to characterize and classify detected proteins of *A. simplex*. Three hundred twenty different protein families and 203 domains were detected based on sequence analysis of all 470 proteins, and the most highly represented of them are displayed in Figure 3. A complete list of InterPro matches is presented in Supplemental File S1. The most abundant InterPro families (Figure 3a) were immunoglobulin-like fold (19 sequences), followed by the NAD(P)-binding domain superfamily (15 sequences), immunoglobulin-like domain superfamily (12 sequences), P-loop containing nucleoside triphosphate hydrolase (10 sequences), and EF-hand domain pair (9 sequences). Most of the identified protein families (n = 220) were represented only by one sequence. As shown in Figure 3b, the most abundant were the following InterPro domains: immunoglobulin-like domain (12 sequences), EF-hand domain (9 sequences), intermediate filament, rod domain (9 sequences), alpha-crystallin/Hsp20 domain (8 sequences), and intermediate filament rod domain coil 1B (8 sequences). One hundred thirty-six domains had only one assigned sequence.

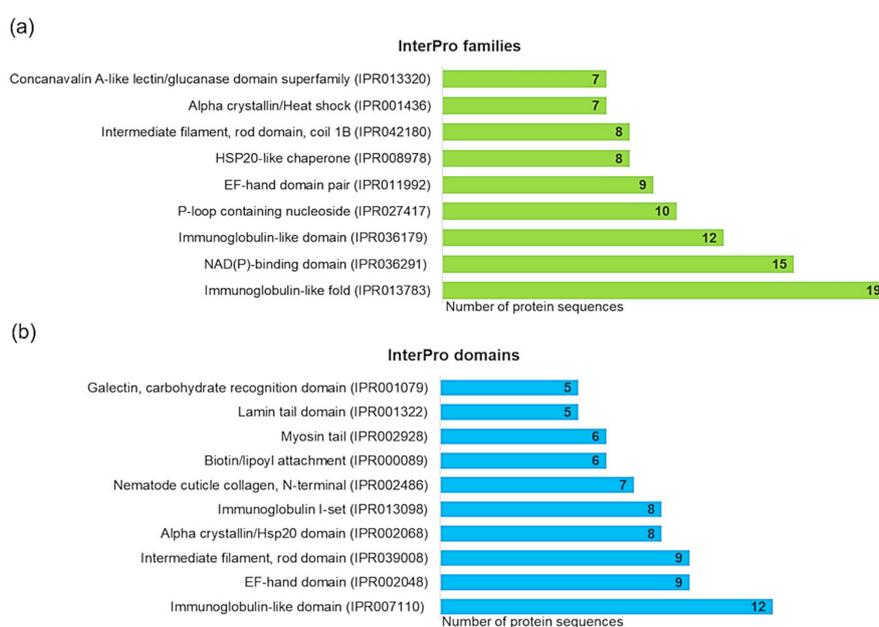


Figure 3. Results of the InterPro analysis of all detected proteins performed using OmicsBox software. Most abundant InterPro families (a); most abundant InterPro domains (b).

Table 1. The top 25 most abundant proteins of *Anisakis simplex* identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS).

No	UniProt ID	Protein Name ¹	InterPro Family ² (ID)	InterPro Domain ² (ID)	Mw ³ (Da)	pI ³	Mascot Scor ⁴	Coverage ⁴ (%)	emPAI Value ⁴
1	A0A0M3K6N3	Myosin, essential light chain (inferred by orthology to a <i>Caenorhabditis elegans</i> protein)	EF-hand domain pair (IPR011992); Myosin light chain alkali (IPR029655)	EF-hand domain (IPR002048)	16,418	4.89	4418	90	201.76
2	A0A0M3KCE6	Tropomyosin (inferred by orthology to a <i>C. elegans</i> protein)	Tropomyosin (IPR000533)	..5	13,701	4.54	3023	63	109.71
3	A0A0M3J349	DUF4440 domain-containing protein	Nuclear transport factor 2 (NTF2)-like domain superfamily (IPR032710)	Domain of unknown function DUF4440 (IPR027843)	13,671	6.29	785	72	81.05
4	A0A0M3J4D2	Uncharacterized protein	-	-	23,930	8.72	1897	71	67.66
5	A0A0M3K7H9	Uncharacterized protein	-	-	11,337	9.89	922	70	63.02
6	Q9NAS5	Tropomyosin (allergen Ani s 3)	Tropomyosin (IPR000533)	-	33,203	4.68	3241	72	62.69
7	A0A0M3KA05	SXP/RAL-2 family protein 2 isoform 1	-	Domain of unknown function DUF4440 (IPR027843)	14,517	4.52	5725	71	45.10
8	A0A0M3K8L6	Uncharacterized protein	DVA-1 superfamily (IPR038289)	Polyprotein allergen, nematode (IPR032487)	42,337	7.07	3504	62	42.49
9	A0A0M3KF53	Uncharacterized protein	-	-	26,477	4.98	2630	70	38.24
10	A0A0M3KIS5	IF rod domain-containing protein	Intermediate filament, rod domain, coil 1B (IPR042180)	Intermediate filament, rod domain (IPR039008)	27,871	5.17	1999	73	33.77

Table 1. Cont.

No	UniProt ID	Protein Name ¹	InterPro Family ² (ID)	InterPro Domain ² (ID)	Mw ³ (Da)	pI ³	Mascot Scor ⁴	Coverage ⁴ (%)	emPAI Value ⁴
11	A0A0M3JCV3	KASH_CCD domain-containing protein	-	KASH5-like coiled-coil domain (IPR028168)	17,230	5.4	1885	68	32.84
12	A0A0M3J6C9	Uncharacterized protein	-	-	26,754	7.78	1755	54	30.87
13	A0A0M3JGD2	Uncharacterized protein	-	-	10,889	4.86	771	52	23.42
14	A0A158PP35	Paramyosin (inferred by orthology to a <i>C. elegans</i> protein)	-	Myosin tail (IPR002928)	101,686	5.34	11203	67	22.63
15	A0A0M3JX08	Protein lethal(2)essential for life (inferred by orthology to a <i>Drosophila melanogaster</i> protein)	HSP20-like chaperone (IPR008978); Alpha-crystallin/Heat shock protein (IPR001436)	Alpha-crystallin/Hsp20 domain (IPR002068)	21,063	5.98	1284	80	21.20
16	A0A0M3JAH1	Intermediate filament protein ifb-1 (inferred by orthology to a <i>C. elegans</i> protein)	Lamin tail domain superfamily (IPR036415)	Intermediate filament, rod domain (IPR039008); Lamin tail domain (IPR001322)	24,324	5.33	991	65	20.01
17	A0A0M3JTQ4	Myosin regulatory light chain 2 (inferred by orthology to a <i>D. melanogaster</i> protein)	EF-hand domain pair (IPR011992)	EF-hand domain (IPR002048)	18,715	5.05	2597	81	19.51
18	A0A0M3KKR0	Uncharacterized protein	-	-	5009	4.79	1077	56	18.10
19	A0A0M3J766	Uncharacterized protein	-	-	13,471	5.82	522	48	16.84
20	Q9NJA9	Paramyosin (allergen Ani s 2)	-	Myosin tail (IPR002928)	100,461	5.21	9369	69	16.04

Table 1. Cont.

No	UniProt ID	Protein Name ¹	InterPro Family ² (ID)	InterPro Domain ² (ID)	Mw ³ (Da)	pI ³	Mascot Scor ⁴	Coverage ⁴ (%)	emPAI Value ⁴
21	A1IKL2	SXP/RAL-2 family protein Ani s 5 (allergen Ani s 5)	-	Domain of unknown function DUF4440 (IPR027843)	14,738	4.85	1765	47	14.65
22	A0A0M3JID9	IF rod domain-containing protein	Intermediate filament, rod domain, coil 1B (IPR042180)	Intermediate filament, rod domain (IPR039008)	10,555	6.29	649	43	14.45
23	A0A0M3JGV0	Uncharacterized protein	-	-	13,766	4.73	844	42	14.31
24	A0A0M3IYQ5	Glutamate dehydrogenase, mitochondrial (inferred by orthology to a D. melanogaster protein)	NAD(P)-binding domain superfamily (IPR036291)	-	11,634	8.89	650	61	14.09
25	A0A0M3J159	Small heat shock protein (SHSP) domain-containing protein	HSP20-like chaperone (IPR008978); Alpha-crystallin/Heat shock protein (IPR001436)	Alpha-crystallin/Hsp20 domain (IPR002068)	13,230	6.29	1036	54	13.78

¹ Protein name according to the Universal Protein Resource (UniProt); ² InterPro families and InterPro domains were predicted using OmicsBox software; ³ the theoretical molecular weight (Mw) in daltons (Da) and theoretical isoelectric point (pI) were estimated using the ExPASy Compute pI/Mw tool; ⁴ Mascote score, coverage (in %), and the exponentially modified protein abundance index (emPAI) were calculated as the average of three biological replicates using values derived from Mascot; ⁵ not detected.

InterPro analysis of the most abundant 25 proteins allowed the identification of 10 families and 9 domains (see Table 1). Among them, slightly more often (two matches for each) were reported the following proteins families: alpha-crystallin/heat shock protein, EF-hand domain pair, HSP20-like chaperone, intermediate filament rod domain coil 1B, and tropomyosin. Whereas the most abundant InterPro domain was the intermediate filament rod domain (three matches), followed by the alpha-crystallin/Hsp20 domain, domain of unknown function DUF148, EF-hand domain, and myosin tail (each of the two matches).

Gene ontology cellular component annotation was used to predict the localization of the detected proteins. Figure 4a shows the distribution of the cellular component annotations of all proteins. The most highly abundant cellular component annotations were the following: organelle (12%; GO:0043226), intracellular organelle (12%; GO:0043229), and cytoplasm (12%; GO:0005737). Less abundant were the following GO terms: membrane (7%; GO:0016020), membrane-enclosed lumen (6%; GO:0031974), cytosol (6%; GO:0005829), supramolecular complex (5%; GO:0099080), cell periphery (5%; GO:0071944), extracellular region (4%; GO:0005576), and endomembrane system (4%; GO:0012505). The count of other cellular component annotations was below 4%.

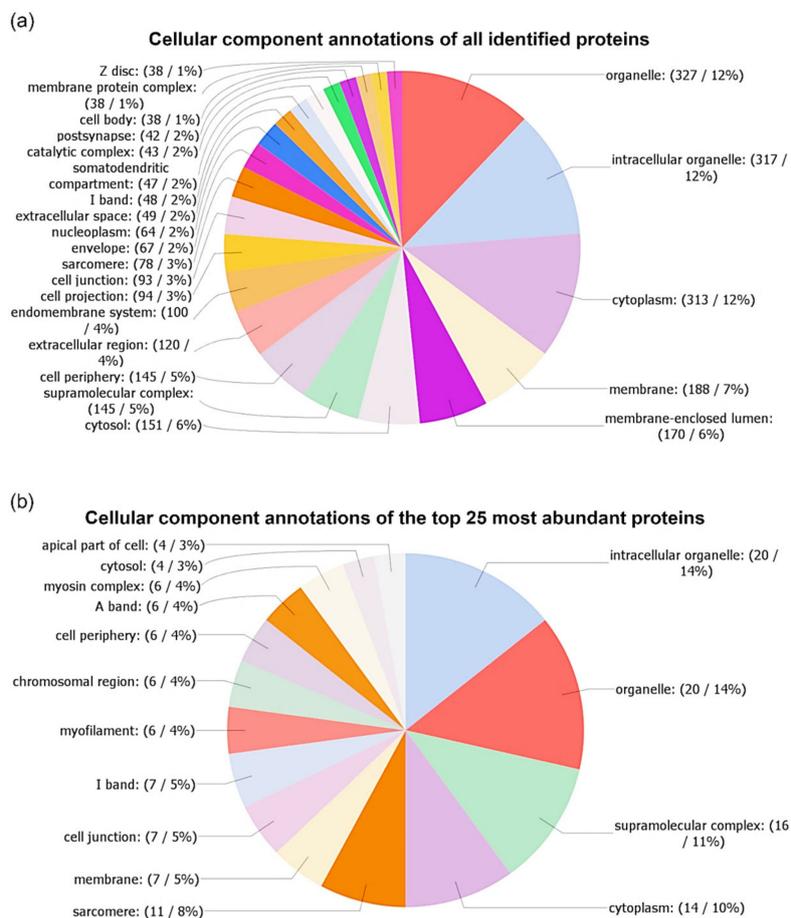


Figure 4. Gene ontology (GO) analysis of cellular components of all identified proteins (a), and the top 25 most abundant proteins (b). The number of annotations was limited for better readability in the following manner: a minimum of 38 matching sequences in GO analysis of all proteins, and a minimum of 4 matching sequences in the case of the top 25 most abundant proteins. Annotations were performed using OmicsBox software.

As shown in Figure 4b, the distribution of the cellular component annotations of the top 25 most abundant proteins was slightly different from the annotations of all proteins. A supramolecular complex (11%; GO:0099080) was the third-largest class of annotations after intracellular organelle (14%;

GO:0043229) and organelle (14%; GO:0043226). Furthermore, sarcomere GO terms (8%; GO:0030017) were of higher abundance, while cytosol (3%; GO:0005829) and membrane-enclosed lumen (2%; GO:0031974) annotations were less represented than in the case of all proteins for cellular component annotations. Other differences in the distribution of cellular component annotations were smaller and are displayed in Figure 4.

3.3. Identification and Characterization of Detected *A. simplex* Allergens and Potential Allergens

Among the identified, the following *A. simplex* allergens were found: Ani s 1, Ani s 2, Ani s 3, Ani s 4, and Ani s 5. We found 13 potential allergens using the AllergenOnline.org server. All detected allergens and potential allergens are listed in Tables 2 and 3, respectively.

Table 4 shows all identified peptides of Ani s 1 and Ani s 5 and matched with these peptides the experimentally verified epitopes from IEDB. B-cell and T-cell epitopes were found in all six peptides of Ani s 1, while B-cell epitopes were detected in 8 of 13 peptides of Ani s 5. Experimentally confirmed epitopes in Ani s 2, Ani s 3, and Ani s 4 were not found in IEDB. In silico predicted epitopes in Ani s 2, Ani s 3, and Ani s 4 peptides are shown in Table 5. Among all 74 Ani s 2 peptides, MHC II, T-cell, and B-cell epitopes were predicted in 15, 31, and 10 peptides, respectively. Ten and nine of all 34 Ani s 3 peptides were matched with T-cell and B-cell epitopes, respectively. T-cell epitopes were found in all five peptides of Ani s 4.

Additionally, the native tertiary structures of the identified allergens were modeled to obtain structural insights into the characteristics of the proteins. Figure 5 shows the native 3D structures of Ani s 1, Ani s 2, Ani s 3, and Ani s 4 modeled with high confidence ($\geq 77\%$ of residues modeled at $> 90\%$ confidence) using the Phyre2 server and the structure of Ani s 5 derived from RCSB PDB (ID: 2MAR). The displayed structures clearly differentiated various protein and allergen classes.

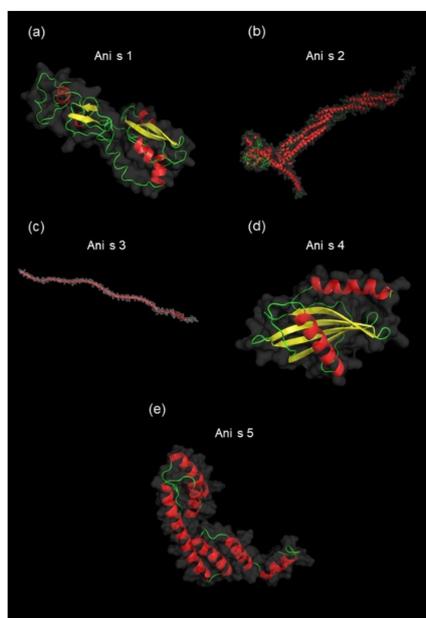


Figure 5. The native tertiary structures of identified allergens. Prediction of unknown 3D structures of allergens: Ani s 1 (77% of residues modeled at $> 90\%$ confidence, model dimensions in angstroms (Å): X:64.335 Y:49.060 Z:49.310) (a), Ani s 2 (99% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:160.179 Y:213.187 Z:160.224) (b), Ani s 3 (99% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:368.887 Y:109.731 Z:146.232) (c), and Ani s 4 (86% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:31.951 Y:46.659 Z:32.835) (d) were performed using the Phyre2 server. The allergen Ani 5 structure was visualized based on the structure from the Protein Data Bank PDB ID: 2MAR (e). PyMOL software was used to visualize all structures. Colors of the allergen models are according to secondary structure: helices highlighted in red, sheets in yellow, and loops in green.

Table 2. Allergens of *A. simplex* identified using LC-MS/MS.

No	UniProt ID	Allergen Name ¹	Protein Name	InterPro Family (ID)	InterPro Domain (ID)	AllFam ² (ID)	Mw (Da)	pI	Mascot Score	Coverage (%)	emPAI Value
1	Q7Z1K3	Ani s 1	Major allergen Ani s 1 (21 kDa allergen)	Pancreatic trypsin inhibitor Kunitz domain superfamily (IPR036880)	Pancreatic trypsin inhibitor Kunitz domain (IPR002223)	Animal Kunitz serine protease inhibitor (AF003)	18,869	7.48	481	21	2.53
2	Q9NJA9	Ani s 2	Paramyosin	-	Myosin tail (IPR002928)	Myosin heavy chain (AF100)	100,461	5.21	9369	69	16.04
3	Q9NAS5	Ani s 3	Tropomyosin	Tropomyosin (IPR000533)	-	Tropomyosin (AF054)	33,203	4.68	3241	72	62.69
4	Q14QT4	Ani s 4	Ani s 4 allergen	-	Cystatin domain (IPR000010)	Cystatin (AF005)	10,291	5.54	196	27	3.67
5	A1IKL2	Ani s 5	SXP/RAL-2 family protein Ani s 5	-	Domain of unknown function DUF148 (IPR003677)	SXP/RAL-2 family (AF137)	14,738	4.85	1765	47	14.65

¹ Allergen name approved by the World Health Organization and International Union of Immunological Societies (WHO/IUIS) Allergen Nomenclature Sub-committee; ² classification of allergen family was conducted according to 'AllFam - The Database of Allergen Families' (ver. 2017.03.07).

Table 3. Potential allergens of *A. simplex* identified by LC-MS/MS and AllergenOnline.org analysis.

No	UniProt ID	Protein Name	Possible Allergens ¹					Homologous Allergens ¹				AllergenOnline.org Result ¹	
			InterPro Family (ID)	InterPro Domain (ID)	Mw (Da)	pI	emPAI Value	Genbank ID	ALLERGEN name	Protein Name (Taxonomy)	AllFam (ID)	Identity ¹	Similarity ¹
1	A0A0M3KCE6	Tropomyosin (inferred by orthology to a <i>C. elegans</i> protein)	Tropomyosin (IPR000533)	-	13,701	4.54	109.71	ACN32322.1	Asc 1 3	Tropomyosin (<i>Ascaris lumbricoides</i>)	Tropomyosin (AF054)	0.946	0.973
2	A0A0M3KA05	SXP/RAL-2 family protein 2 isoform 1	-	Domain of unknown function DUF148 (IPR003677)	14,517	4.52	45.10	BAF75681.1	Ani s 8	SXP/RAL-2 family protein 2 isoform 1 (<i>Anisakis simplex</i>)	SXP/RAL-2 family (AF137)	0.987	1
3	A0A158PP35	Paramyosin (inferred by orthology to a <i>C. elegans</i> protein)	-	Myosin tail (IPR002928)	101,686	5.34	22.63	Q9NJA9.1	Ani s 2	Paramyosin (<i>Anisakis simplex</i>)	Myosin heavy chain (AF100)	0.854	0.958
4	A0A0M3J4T5	DUF148 domain-containing protein	-	Domain of unknown function DUF148 (IPR003677)	16,646	5.56	7.10	BAF43534.1	Ani s 5	SXP/RAL-2 family protein (<i>Anisakis simplex</i>)	SXP/RAL-2 family (AF137)	0.971	0.993
5	A0A0M3J693	Tropomyosin (inferred by orthology to a <i>C. elegans</i> protein)	Tropomyosin (IPR000533)	-	11,304	4.64	7.79	ACN32322.1	Asc 1 3	Tropomyosin (<i>Ascaris lumbricoides</i>)	Tropomyosin (AF054)	0.824	0.953
6	A0A0M3K9V2	Heat shock 70 kDa protein cognate 1 (inferred by orthology to a <i>D. melanogaster</i> protein)	Heat shock protein 70 kD, peptide-binding domain superfamily (IPR029047); Heat shock protein 70 family (IPR013126); Heat shock protein 70 kD, C-terminal domain superfamily (IPR029048)	-	67,585	5.45	4.84	AOD75395.1	Tyr p 28	Heat shock-like protein (<i>Tyrophagus putrescentiae</i>)	Heat shock protein (Hsp70) (AF002)	0.737	0.86

Table 3. Cont.

No	UniProt ID	Protein Name	Possible Allergens ¹					Homologous Allergens ¹				AllergenOnline.org Result ¹	
			InterPro Family (ID)	InterPro Domain (ID)	Mw (Da)	pI	emPAI Value	Genbank ID	ALLERGEN name	Protein Name (Taxonomy)	AllFam (ID)	Identity ¹	Similarity ¹
7	A0A0M3KIW1	Uncharacterized protein	Heat shock protein 70 kD, peptide-binding domain superfamily (IPR029047); Heat shock protein 70 family (IPR013126)	-	24,599	8.49	2.31	AOD75395.1	Tyr p 28	Heat shock-like protein (Tyrophagus putrescentiae)	Heat shock protein (Hsp70) (AF002)	0.851	0.95
8	A0A0M3J6G4	Peptidyl-prolyl cis-trans isomerase	Cyclophilin-like domain superfamily (IPR029000)	Cyclophilin-type peptidyl-prolyl cis-trans isomerase domain (IPR002130)	12,614	9.04	1.99	AVV30163.1	Ole e 15	Cyclophilin 0101 (Olea europaea)	-	0.789	0.916
9	A0A0M3JU57	Troponin-like protein	EF-hand domain pair (IPR011992)	EF-hand domain (IPR002048)	19,444	4.15	1.39	CAB58171.1	Ani s troponin C	Troponin-like protein (Anisakis simplex)	EF hand family (AF007)	0.994	1
10	A0A0M3K5H6	78 kDa glucose-regulated protein	Heat shock protein 70 kD, peptide-binding domain superfamily (IPR029047); Heat shock protein 70 family (IPR013126); Heat shock protein 70 kD, C-terminal domain superfamily (IPR029048)	Endoplasmic reticulum chaperone BIP, nucleotide-binding domain (IPR042050)	75,409	5.07	0.77	ABF18258.1	Aed a 8	Heat shock cognate 70 (Aedes aegypti)	Heat shock protein (Hsp70) (AF002)	0.777	0.93
11	A0A0M3J8H0	Uncharacterized protein	-	-	9756	5.52	0.51	ABV55106.1	Ani s 9	Ani s 9 allergen precursor (Anisakis simplex)	SXP/RAL-2 family (AF137)	1	1
12	A0A0M3K821	Tubulin alpha chain	-	-	47,784	4.83	0.42	AIO08861.1	Der f 33	Der f 33 allergen (Dermatophagoides farinae)	Tubulin/FtsZ family (AF025)	0.735	0.866
13	A0A0M3KAH2	Tubulin alpha chain	-	-	48,222	5	0.51	AUX14773.1	Der f 33-like	Der f 33-like protein (Dermatophagoides pteronyssinus)	-	0.743	0.891

¹ Potential allergen and homologous allergens with identity and similarity values (1 is equal to 100%) were calculated using the full-length FASTA searching algorithm by AllergenOnline.org.

Table 4. List of known B-cell/T-cell epitopes that were bioinformatically identified in Ani s 1 and Ani s 5 peptides derived from the autoclaved antigen of *A. simplex*.

Peptide ¹		Epitope ³				
Sequence ¹	Position ¹	IEDB Epitope ID ³	Sequence ³	Position ³	B-Cell/T-Cell ³	Reference ³
Q7Z1K3 (Ani s 1) ⁴						
LFANCKK ²	188-194	137143	EDAKCER <u>G</u> KL <u>F</u> ANCKK ⁵	179-194	B-cell	[48]
		421049	GK <u>L</u> FANCKK	186-194	T-cell	[49]
EEELFAR	137-143	421391	MGLCCPTKEEEL <u>F</u> A	129-142	B-cell	[49]
		421053	GLCCPTKEEEL <u>F</u> A	130-142	B-cell	[49]
		420891	EELFAREYEG <u>V</u> C	138-149	T-cell	[49]
		420774	CCPTKEEEL <u>F</u> AR	132-143	T-cell	[49]
GSGWMMTILGK	160-170	137278	R <u>G</u> S <u>G</u> W <u>M</u> M <u>T</u> I <u>L</u> G <u>K</u> S <u>C</u> D	159-173	B-cell	[48]
		421086	G <u>W</u> M <u>M</u> T <u>I</u> L <u>G</u> K <u>S</u> C <u>D</u>	162-173	T-cell	[49]
		421087	G <u>W</u> M <u>M</u> T <u>I</u> L <u>G</u> K <u>S</u> C <u>D</u> D <u>Q</u> F <u>C</u> P <u>E</u> D <u>A</u>	162-181	B-cell	[49]
		421233	K <u>M</u> D <u>R</u> G <u>S</u> G <u>W</u> M <u>M</u> T <u>I</u>	156-167	T-cell	[49]
		421406	<u>M</u> T <u>I</u> L <u>G</u> K <u>S</u> C <u>D</u> D <u>Q</u> F <u>C</u> P <u>E</u> D <u>A</u>	165-181	B-cell	[49]
		421613	R <u>G</u> S <u>G</u> W <u>M</u> M <u>T</u> I <u>L</u> G	159-169	B-cell	[49]
MMAFMGLCCPTK	125-136	420792	CP <u>N</u> G <u>Y</u> Q <u>C</u> K <u>M</u> MA <u>F</u> M <u>G</u>	117-130	B-cell	[49]
		421053	GLCCPTKEEEL <u>F</u> A	130-142	B-cell	[49]
		421091	GYQCK <u>M</u> MA <u>F</u> M <u>G</u> L	120-131	T-cell	[49]
		421092	GYQCK <u>M</u> MA <u>F</u> M <u>G</u> LCCPTK	120-136	B-cell	[49]
		421384	MA <u>F</u> M <u>G</u> LCCPT <u>K</u> E	126-137	T-cell	[49]
		421391	MGLCCPTKEEEL <u>F</u> A	129-142	B-cell	[49]
MDRSGWMMTILGK	157-170	137278	R <u>G</u> S <u>G</u> W <u>M</u> M <u>T</u> I <u>L</u> G <u>K</u> S <u>C</u> D	159-173	B-cell	[48]
		421086	G <u>W</u> M <u>M</u> T <u>I</u> L <u>G</u> K <u>S</u> C <u>D</u>	162-173	T-cell	[49]
		421087	G <u>W</u> M <u>M</u> T <u>I</u> L <u>G</u> K <u>S</u> C <u>D</u> D <u>Q</u> F <u>C</u> P <u>E</u> D <u>A</u>	162-181	B-cell	[49]
		421233	K <u>M</u> D <u>R</u> G <u>S</u> G <u>W</u> M <u>M</u> T <u>I</u>	156-167	T-cell	[49]
		421406	<u>M</u> T <u>I</u> L <u>G</u> K <u>S</u> C <u>D</u> D <u>Q</u> F <u>C</u> P <u>E</u> D <u>A</u>	165-181	B-cell	[49]
		421613	R <u>G</u> S <u>G</u> W <u>M</u> M <u>T</u> I <u>L</u> G	159-169	B-cell	[49]

Table 4. Cont.

Peptide ¹		Epitope ³				
Sequence ¹	Position ¹	IEDB Epitope ID ³	Sequence ³	Position ³	B-Cell/T-Cell ³	Reference ³
SGICLSFK	56-63	420822	DKSGICLSFKYT	54-65	T-cell	[49]
		421875	WHDDKSGICLSFKY	51-64	B-cell	[49]
A1IKL2 (Ani s 5)						
ADAELSK	105-111	230144	<u>KKADAELSKIA</u>	103-113	B-cell	[50]
FETFKK	73-78	-	-	-	-	-
TDPEIEK	50-56	-	-	-	-	-
AKEAELAK	82-89	-	-	-	-	-
KADAELSK	104-111	230144	<u>KKADAELSKIA</u>	103-113	B-cell	[50]
AFFELLK	38-44	-	-	-	-	-
DETKTDPEIEK	46-56	230147	<u>KTDPEI</u>	49-54	B-cell	[50]
DELEKGIGPAVPQ	140-152	230136	<u>IQAIYKTLPQSVKDELEKGI</u>	127-146	B-cell	[50]
TLPQSVKDELEK	133-144	230136	<u>IQAIYKTLPQSVKDELEKGI</u>	127-146	B-cell	[50]
		230137	<u>IQKAQKIQAIYKTLPQSVKD</u>	121-140	B-cell	[50]
EAELAKAHEEAVAK	84-97	-	-	-	-	-
DLDAWVDTLGGDYK	57-70	230149	<u>LDAWVDTLG</u>	58-66	B-cell	[50]
		230150	<u>LDAWVDTLGGDYKAKFETFK</u>	58-77	B-cell	[50]
DLDAWVDTLGGDYKAK	57-72	230149	<u>LDAWVDTLG</u>	58-66	B-cell	[50]
		230150	<u>LDAWVDTLGGDYKAKFETFK</u>	58-77	B-cell	[50]
TDPEIEKDLDAWVDTLGGDYK	50-70	230149	<u>LDAWVDTLG</u>	58-66	B-cell	[50]
		230150	<u>LDAWVDTLGGDYKAKFETFK</u>	58-77	B-cell	[50]

¹Peptides were identified using LC-MS/MS and retrieved from the Mascot search results; ²sequences of unique peptides are bolded; ³epitope mapping data were retrieved from the Immune Epitope Database and Analysis Resource (IEDB); ⁴UniProt ID and allergen name; ⁵the epitope sequences found in the peptides are underlined.

Table 5. List of in silico predicted epitopes (MHC II, T-cell, and B-cell) in Ani s 2, Ani s 3, and Ani s 4 peptides derived from the autoclaved antigen of *A. simplex*.

Peptide ¹		Epitope MHC II ³		Epitope T-Cell ³		Epitope B-Cell ³	
Sequence ¹	Position ¹	Sequence ³	Position ³	Sequence ³	Position ³	Sequence ³	Position ³
Q9NJA9 (Ani s 2) ⁴							
ANLEAQK ²	562-568	-	-	-	-	-	-
DLQVALK	465-471	-	-	-	-	-	-
SLEEQVK	712-718	-	-	-	-	-	-
AALAEQK	402-409	-	-	<u>LAELQKMKQLYEKAVE</u> ⁵	404-419	-	-
TALDNAIR	617-624	-	-	<u>CKTALDN</u>	615-621	-	-
ELEDAEGR	829-836	-	-	-	-	<u>RELEDAEGRA</u>	828-837
ELHAADER	674-681	-	-	<u>LDEVTKELHAAD</u>	668-679	-	-
ALAEQQVR	487-495	-	-	<u>RAQRALAE</u>	483-491	-	-
VALDEESAAR	271-280	-	-	-	-	-	-
HKETQSALR	760-768	-	-	-	-	<u>RIRDLEVALDEETRRHKETQSALRKKDRRI</u>	745-774
LTAALADAEAR	523-533	<u>LTAALA</u>	523-528	-	-	-	-
AQNTAILER	361-370	-	-	-	-	-	-
EEEMEALRK	506-514	-	-	-	-	-	-
VQLDNLQHVK	222-231	-	-	-	-	-	-
LHELDLENAR	449-458	-	-	-	-	-	-
LLQDDFESER	43-52	-	-	-	-	<u>QDDFESERELRNRIERERA</u>	45-63
MFVMAQDTADR	787-797	-	-	<u>DTADRMLE</u>	793-800	-	-
ADQAESSLNLR	837-848	-	-	-	-	-	-
FEQQTIELSNK	178-188	-	-	-	-	-	-
KLHELDLENAR	448-458	-	-	-	-	-	-
ISALSAELEECK	605-616	<u>ISALSA</u>	605-610	-	-	-	-
YQLAQQLEESR	232-242	-	-	-	-	-	-
QAEADLEEAHVR	628-639	<u>AHVRISDLTS</u>	635-645	-	-	-	-
EVQMQIDEEHK	776-786	-	-	-	-	<u>EVQMQI</u>	776-781
LSLANTEITQWK	287-298	-	-	-	-	-	-

Table 5. Cont.

Peptide ¹		Epitope MHC II ³		Epitope T-Cell ³		Epitope B-Cell ³	
Sequence ¹	Position ¹	Sequence ³	Position ³	Sequence ³	Position ³	Sequence ³	Position ³
ADLSVQLIALTDR	63-75	-	-	-	-	-	-
DLQVALKESEAAR	465-477	-	-	-	-	-	-
KISALSAELEECK	604-616	<u>ISALSA</u>	605-610	-	-	-	-
KQAEADLEEAHVR	627-639	-	-	-	-	-	-
IRDLEVALDEETR	746-758	-	-	-	-	<u>RIRDLEVALDEETRRHKETQSALRKKDRRI</u>	745-774
ISDLTSINSNLTAIK	640-654	<u>AHVRISDLTS</u>	635-645	<u>ISDLTSI</u>	640-646	-	-
LQSEVEVLIVDLEK	347-360	-	-	-	-	-	-
IKEVQMQUIDEEHK	774-786	-	-	-	-	<u>EVQMQUI</u>	776-781
QLQQTLDQYALAQR	590-603	-	-	<u>LEDTQRQLQQT</u>	584-594	-	-
ERADLSVQLIALTDR	61-75	-	-	-	-	-	-
AVQELHEEQEHSMK	692-705	-	-	<u>ANRALADAARAVQELHE</u>	682-698	-	-
QLGEAEAMTMQNLQR	808-822	-	-	<u>AEAMTMQNLQRVRRYQRELEDA</u>	812-833	-	-
SRIDELLVELEAAQR	384-398	-	-	<u>IDELLVE</u>	386-392	-	-
QAEYEEQIEIMLQK	323-336	-	-	-	-	-	-
VEDLNKHVNDLAQQR	189-203	-	-	<u>SNKVEDLNKHVNDLAQ</u>	186-201	<u>LNKHVN</u>	192-197
MAQKFEQQTIELSNK	174-188	-	-	<u>HVAEKMAQKFE</u>	169-179	-	-
QLQQTLDQYALAQRK	590-604	-	-	<u>LEDTQRQLQQT</u>	584-594	-	-
LQAENSLLAEIHDQK	206-221	-	-	<u>DLLAEIH</u>	212-218	-	-
ISDLTSINSNLTAIKNK	640-656	<u>AHVRISDLTS</u>	635-645	<u>ISDLTSI</u>	640-646	-	-
ARLQSEVEVLIVDLEK	345-360	-	-	-	-	-	-
LETTELSTAQADLDEVTK	657-673	<u>LSTAQA</u>	661-666	<u>LDEVTKELHAAD</u>	668-679	-	-
LQHEVIELTATIDQLQK	150-166	<u>IELTAT</u>	155-160	<u>EVIELTATIDQLQK</u>	153-166	-	-
LLEESQLENEDAMNVLR	103-119	-	-	<u>EDAMN</u>	112-116	-	-
YQAEIAELEMVTDNLNR	545-561	-	-	<u>EMTVDNLN</u>	553-560	-	-
QLQVQIQEAEAAALLGGKR	719-737	-	-	-	-	-	-
FEQQTIELSNKVEDLNK	178-194	-	-	<u>SNKVEDLNKHVNDLAQ</u>	186-201	-	-
VEAEHKLSLANTEITQWK	281-298	-	-	-	-	-	-

Table 5. Cont.

Peptide ¹		Epitope MHC II ³		Epitope T-Cell ³		Epitope B-Cell ³	
Sequence ¹	Position ¹	Sequence ³	Position ³	Sequence ³	Position ³	Sequence ³	Position ³
NKLETELSTAQADLDEVTK	655-673	<u>LSTAQA</u>	661-666	<u>LDEVTKELHAAD</u>	668-679	-	-
SMQFEIDRLTAALADAEAR	515-533	<u>LTAALA</u>	523-528	-	-	-	-
QRLQAENSDLLAEIHDQK	204-221	-	-	<u>DLLAEIH</u>	212-218	-	-
SQMQAQLHQVQLELDSVR	253-270	<u>LHQVQL</u>	259-264	-	-	-	-
SKFDAEVALHHEEVEDLR	299-316	-	-	<u>EVEDLRKKMMQ</u>	311-321	-	-
SQMQAQLHQVQLELDSVR	253-270	<u>LHQVQL</u>	259-264	-	-	-	-
LLEESQLENEDAMNVLRLK	103-120	-	-	<u>EDAMN</u>	112-116	-	-
KLLEESQLENEDAMNVLRL	102-119	-	-	<u>EDAMN</u>	112-116	-	-
QRLQHEVIELTATIDQLQK	148-166	-	-	<u>EVIELTATIDQLQK</u>	153-166	-	-
QAEYEEQIEIMLQKVSQLEK	323-342	-	-	-	-	-	-
MKAEIAR	534-540	-	-	-	-	-	-
MKQLYEK	410-416	-	-	-	-	-	-
EVELSKLR	94-101	-	-	-	-	<u>ESNRKREVELS</u>	88-98
DLEVALDEETRR	748-759	-	-	-	-	<u>RIRDLEVALDEETRRHKETQSALRKKDRRI</u>	745-774
LQDAECATDSQIESNR	76-91	-	-	-	-	<u>ESNRKREVELS</u>	88-98
KQSEQIHLQANLEDTQR	572-589	<u>IIQLQA</u>	577-582	<u>LEDTQRQLQQT</u>	584-594	-	-
SSTADMGALTSMSVADLGSLTR	15-36	<u>LTSMSVA</u>	23-29	<u>MGALTSMSVADLGSLTR</u>	20-37	-	-
ISALSAELEECKTALDNAIR	605-624	<u>ISALSA</u>	605-610	<u>CKTALDN</u>	615-621	-	-
VSQLEK	337-342	-	-	-	-	-	-
LNIQKR	802-807	-	-	-	-	-	-
KSLEEQVK	711-718	-	-	<u>DALRKSLEEQV</u>	707-717	-	-
KYQAEIAELEMTVDNLNR	544-561	-	-	<u>EMTVDNLN</u>	553-560	-	-
Q9NAS5 (Ani s 3)							
AEFAER	239-244	-	-	-	-	-	-
ADAAEEK	22-28	-	-	<u>DAAEEKVRQMTDKLERIEEELRDT</u> <u>QKKMMQTENDLDKAQEDL</u>	23-64	-	-
KVMENR	128-133	-	-	-	-	<u>RKVMENRSFQDEE</u>	127-139
IEEELR	39-44	-	-	-	-	-	-

Table 5. Cont.

Peptide ¹		Epitope MHC II ³		Epitope T-Cell ³		Epitope B-Cell ³	
Sequence ¹	Position ¹	Sequence ³	Position ³	Sequence ³	Position ³	Sequence ³	Position ³
LKEAETR	232-238	-	-	-	-	-	-
SFQDEER	134-140	-	-	-	-	RKVMENRSFQDEE	127-139
SLEVSEEK	206-213	-	-	-	-	-	-
ANTVESQLK	141-149	-	-	-	-	-	-
QMTDKLER	31-38	-	-	-	-	<u>TDKLERIEEELRDT</u>	33-46
IEKDNALDR	13-21	-	-	-	-	-	-
LEDELVHEK	256-264	-	-	<u>EFAERSVQKLQKEVDRLEDELVH</u>	240-262	<u>DRLEDELV</u>	254-261
IVELEEELR	190-198	-	-	<u>KIVELEE</u>	189-195	<u>LEEEL</u>	193-197
AEFAERSVQK	239-248	-	-	<u>EFAERSVQKLQKEVDRLEDELVH</u>	240-262	-	-
EDSYEEQIR	218-226	-	-	<u>EQIRTVS</u>	223-229	<u>LQREDSYEEQIR</u>	215-226
MMQTENDLDK	50-59	-	-	-	-	-	-
IEEELRDTQK	39-48	-	-	-	-	<u>TDKLERIEEELRDT</u>	33-46
EAQMLAEEADR	150-160	-	-	-	-	-	-
MTLLEEELER	92-101	-	-	-	-	-	-
SLEVSEEKALQR	206-217	-	-	-	-	-	-
EAQMLAEEADRK	150-161	-	-	-	-	-	-
VQEAEAEVAALNR	78-90	-	-	-	-	-	-
DNALDRADAAEEK	16-28	-	-	<u>DAAEEKVRQMTDKLERIEEELRDTQ</u> <u>KKMMQTENDLLDKAQEDL</u>	23-64	-	-
RMTLLEEELER	91-101	-	-	-	-	-	-
LEEATHADESER	113-125	-	-	<u>LEEATHADESERVRKVME</u>	113-131	-	-
KVQEAEAEVAALNR	77-90	-	-	-	-	-	-
VQEAEAEVAALNRR	78-91	-	-	-	-	-	-
EVDREDELVHEK	252-264	-	-	<u>EFAERSVQKLQKEVDRLEDELVH</u>	240-262	<u>DRLEDELV</u>	254-261
ALQREDSYEEQIR	214-226	-	-	<u>EQIRTVS</u>	223-229	<u>LQREDSYEEQIR</u>	215-226
AQEDLSTANSNLEEK	60-74	-	-	-	-	-	-
SISEELDQTFQELSGY	269-284	-	-	-	-	-	-
YKSISEELDQTFQELSGY	267-284	-	-	-	-	-	-

Table 5. Cont.

Peptide ¹		Epitope MHC II ³		Epitope T-Cell ³		Epitope B-Cell ³	
Sequence ¹	Position ¹	Sequence ³	Position ³	Sequence ³	Position ³	Sequence ³	Position ³
ANTVESQLKEAQMLAEEADR	141-160	-	-	-	-	-	-
KMMQTENDLDK	49-59	-	-	-	-	-	-
MMQTENDLDKAQE DLSTANSNLEEK	50-74	-	-	DAAEEKVRQMTDKLERIEEELRDT QKKMMQTENDLDKAQEDL	23-64	-	-
Q14QT4 (Ani s 4)							
KQVVAGDK	69-76	-	-	-	-	-	-
KWENFEEVK	100-108	-	-	WQKKWENFEEVKVLKCDH	97-114	-	-
ISAMINDGKPHLVK	49-63	-	-	<u>ELAGKSIAKISAMI</u>	40-53	-	-
		-	-	<u>GKPHLVKVVVS</u>	56-66	-	-
ELAGKSIK	40-48	-	-	<u>ELAGKSIKISAMI</u>	40-53	-	-
WENFEEVK	101-108	-	-	WQKKWENFEEVKVLKCDH	97-114	-	-

¹ Peptides were identified using LC-MS/MS and retrieved from the Mascot search results; ² sequences of unique peptides are bolded; ³ major histocompatibility complex class II (MHC II) molecules, T-cell, and B-cell epitopes were predicted using DNA STAR Protean 3D software; ⁴ UniProt ID and allergen name; ⁵ the epitope sequences found in the peptides are underlined.

The largest group of potential allergens (five proteins) were homologous to *A. simplex* allergens (Ani s 2, Ani s 5, Ani s 8, Ani s 9, and Ani s troponin C). Four potential allergens were homologous to the following Acari (*Tyrophagus putrescentiae*, *Dermatophagoides farinae*, *Dermatophagoides pteronyssinus*) allergens: Tyr p 28, Der f 33, and Der f 33-like. Another two proteins were homologous to *Ascaris lumbricoides* allergen Asc l 3. The last two potential allergens of *A. simplex* were homologous to *Olea europaea* allergen Ole e 15 and *Aedes aegypti* allergen Aed a 8. Three-dimensional native structures of homologous allergens for potential allergens of *A. simplex* are shown in Figure 6 ($\geq 85\%$ of residues modeled at $> 90\%$ confidence), except for the Ani s 2 and Ani s 5 models, which are presented in Figure 5. Additionally, the alignment of 3D models of potential allergens and their homologous allergens was performed to visualize the similarities in protein structures (see Supplemental File S2). The displayed 3D models of allergens and potential allergens showed similarities among the proteins belonging to the same allergen classes.

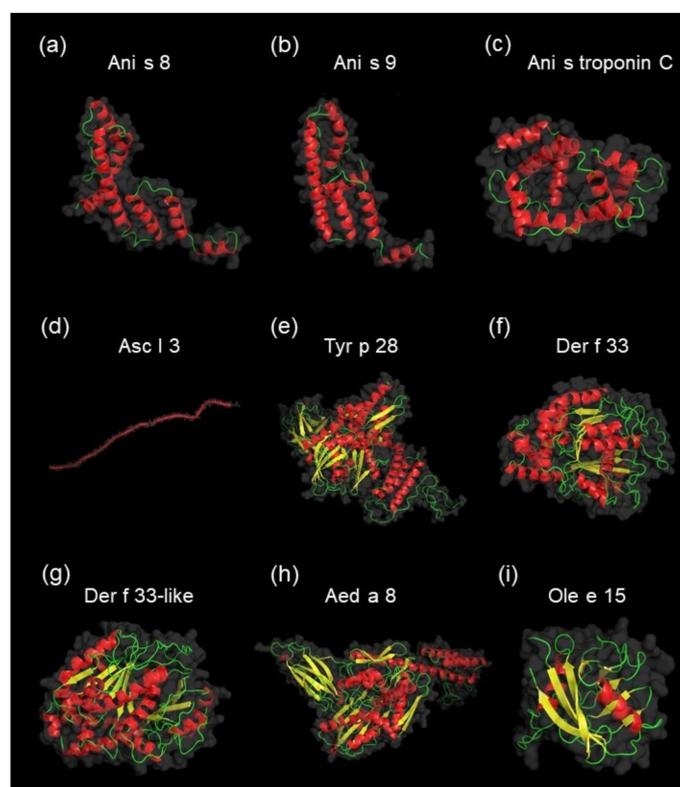


Figure 6. The native tertiary structures of homologous allergens, as identified using AllergenOnline.org for potential allergens of *A. simplex*. Prediction of unknown 3D structures of allergens: Ani s 8 (85% of residues modeled at $> 90\%$ confidence, model dimensions in angstroms (Å): X: 62.875 Y: 60.531 Z: 35.486) (a), Ani s 9 (87% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:54.278 Y: 58.349 Z:28.680) (b), Ani s troponin C (99% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:50.390 Y: 40.831 Z:35.415) (c), Asc l 3 (99% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:361.607 Y:132.976 Z:152.396) (d), Tyr p 28 (97% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:81.072 Y:91.096 Z:63.286) (e), Der f 33 (89% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:63.736 Y:59.230 Z:58.888) (f), Der f 33-like (90% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:57.763 Y:55.796 Z:56.582) (g), Aed a 8 (90% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:57.763 Y:55.796 Z:56.582) (h), Ole e 15 (100% of residues modeled at $> 90\%$ confidence, model dimensions in Å: X:38.017 Y:36.365 Z:47.475) (i) were performed using the Phyre2 server. Allergen 3D structures of Ani s 2 and Ani s 5, which are also homologs of potential allergens, are shown in Figure 5. PyMOL software was used to visualize all structures. Colors of the allergen models are according to secondary structure: helices highlighted in red, sheets in yellow, and loops in green.

A. simplex allergens and homologs of potential allergens (see Tables 2 and 5) were classified to allergen families based on the AllFam database. The five following AllFam families were detected among identified allergens: animal Kunitz serine protease inhibitor (Ani s 1), myosin heavy chain (Ani s 2), tropomyosin (Ani s 3), cystatin (Ani s 4), and SXP/RAL-2 family (Ani s 5). In the case of homologs of potential allergens, six families of AllFam were found. The most highly represented AllFam families of homologous allergens were heat shock protein 70 (Hsp70) and SXP/RAL-2 family (three matches for each), followed by tropomyosin (two matches). The lowest abundant allergen families (one match for each) were the following: EF-hand family, myosin heavy chain, and tubulin/FtsZ family. Two homologous allergens (Ole e 15 and Der f 33-like) were not reported in the AllFam database.

The detected allergens were in the Mw/pI range of 10291–100461 Da/4.68–7.48, while potential allergens were in the range of 9756–101686 Da/4.15–9.04. The lowest Mw of the detected allergens and potential allergens was found for the Ani s 4 and uncharacterized protein (UniProt ID: A0A0M3J8H0), respectively. Ani s 2 (paramyosin) was the allergen with the highest calculated Mw, while the highest Mw value of potential allergens was found for the paramyosin (UniProt ID: A0A158PP35). Ani s 3 (tropomyosin) and troponin-like protein (UniProt ID: A0A0M3JU57) had the lowest calculated pI values of allergens and potential allergens, respectively. Ani s 1 was the allergen with the highest pI, whereas the highest pI of potential allergens was calculated for the peptidyl-prolyl cis-trans isomerase (UniProt ID: A0A0M3J6G4).

The identified allergens were in the range of emPAI values from 2.53 (Ani s 1) to 62.69 (Ani s 3); while the potential allergens were in the range from 0.42 (tubulin alpha chain; UniProt ID: A0A0M3K821) to 109.71 (tropomyosin; UniProt ID: A0A0M3KCE6). In the top 25 most abundant proteins (see Table 1), the following allergens were found: Ani s 2, Ani s 3, and Ani s 5, and the following possible allergens: tropomyosin (UniProt ID: A0A0M3KCE6), SXP/RAL-2 family protein 2 isoform 1 (UniProt ID: A0A0M3KA05), and paramyosin (UniProt ID: A0A158PP35). The distribution of the detected allergens and potential allergens based on the Mw, pI, and emPAI values is displayed in Figure 2.

All five detected allergens were classified using InterPro analysis, whereas the three following potential allergens: uncharacterized protein (UniProt ID: A0A0M3J8H0), tubulin alpha chain (UniProt ID: A0A0M3K821), and tubulin alpha chain (UniProt ID: A0A0M3KAH2) had no domains or families matches. Among the identified allergens, two InterPro families (pancreatic trypsin inhibitor Kunitz domain superfamily and tropomyosin) and four InterPro domains (pancreatic trypsin inhibitor Kunitz domain, myosin tail, cystatin domain, and domain of unknown function DUF148) were detected. Six InterPro families of potential allergens were found, from which the most abundant (three matches for each) were the heat shock protein 70 family and heat shock protein 70 kD peptide-binding domain superfamily. Less represented (two matches for each) were the following InterPro families: tropomyosin and heat shock protein 70 kD C-terminal domain superfamily, while the least abundant (one match for each) were EF-hand domain pair and cyclophilin-like domain superfamily. Five domains of all possible allergens were identified. The most abundant was the domain of unknown function DUF148 (two matches), and other domains (one match for each) were the following: myosin tail, cyclophilin-type peptidyl-prolyl cis-trans isomerase domain, EF-hand domain, and endoplasmic reticulum chaperone binding immunoglobulin protein (BIP) nucleotide-binding domain.

The distribution of cellular component annotations of allergens and potential allergens are displayed in Figure 7a,b, respectively. A detailed list of all GO terms with assigned sequences is shown in Supplemental File S1. The most abundant class of cellular component annotations of allergens was cytoplasm (14%; GO:0005737), with three matched sequences: Ani s 2, Ani s 3, and Ani s 5. Lower represented (9%) were the following annotations: intracellular organelle (matches: Ani 2, Ani s 3; GO:0043229), organelle (matches: Ani s 2, Ani s 3; GO:0043226), supramolecular complex (matches: Ani s 2, Ani s 3; GO:0099080), myofilament (matches: Ani s 2, Ani s 3; GO:0036379), extracellular region (matches: Ani s 1, Ani s 5; GO:0005576), and sarcomere (matches: Ani s 2, Ani s 3; GO:0030017). The least abundant (5%) cellular component annotations were muscle thin filament

tropomyosin (match: Ani s 3; GO:0005862), I band (match: Ani s 3; GO:0031674), A band (match: Ani s 2; GO:0031672), myosin complex (match: Ani s 2; GO:0016459), microtubule organizing center (match: Ani s 2; GO:0005815), extracellular space (match: Ani s 5; GO:0005615), and extracellular matrix (match: Ani s 1; GO:0031012).

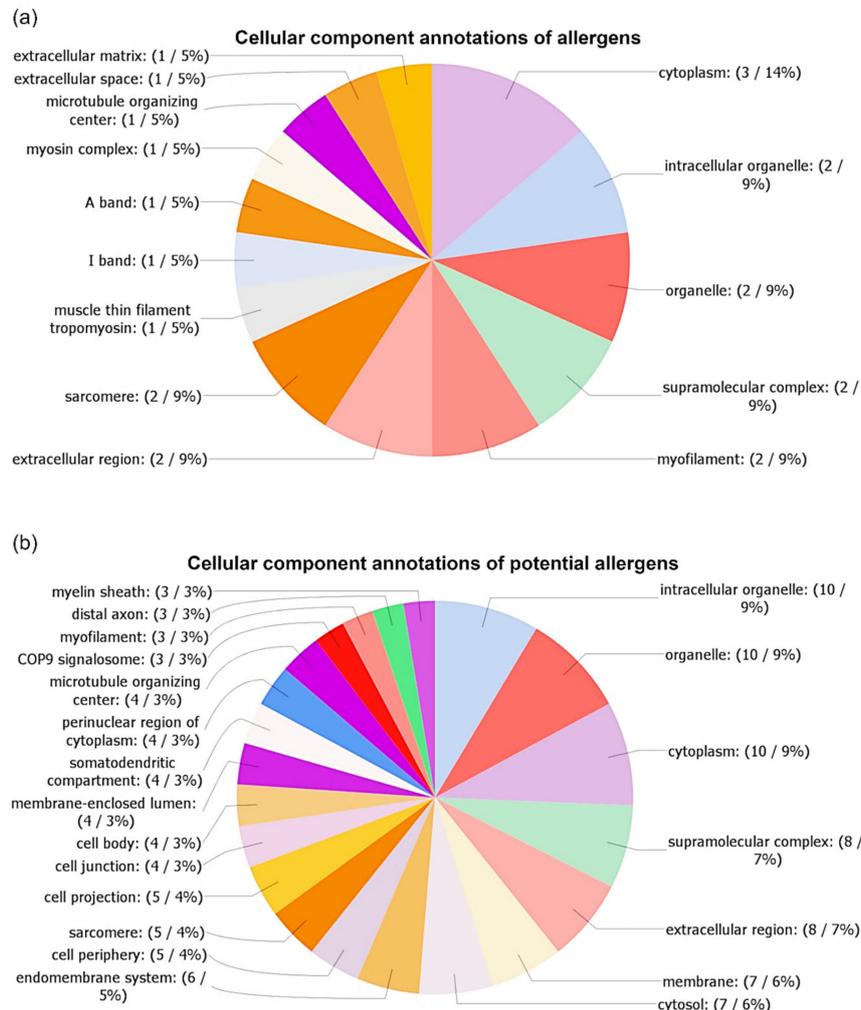


Figure 7. Gene ontology analysis of cellular components of allergens (a) and potential allergens (b). The number of annotations was limited for better readability in the case of potential allergens (minimum of three sequences). Annotations were performed using OmicsBox software.

Twenty-one cellular components GO terms were found based on the analysis of potential allergens. The most abundant of them were the following GO terms: intracellular organelle (9%; GO:0043229), organelle (9%; GO:0043226), cytoplasm (9%; GO:0005737), followed by supramolecular complex (7%; GO:0099080), extracellular region (7%; GO:0005576), membrane (6%; GO:0016020), and cytosol (6%; GO:0005829). Other cellular components GO terms were below 6%.

4. Discussion

The sterilization process is widely used for preserving many types of food, including products containing the meat of sea fish. Studies evaluating the effects of such a process on the *A. simplex* proteins, which could contaminate fish products, are only fragmentary. Therefore, in the present work, we attempted to reduce the knowledge gap in this topic, especially focusing on the proteome of *Anisakis*. A current survey was the first proteome and allergome profiling of *A. simplex* in autoclaving conditions.

4.1. Electrophoretic and Immunological Investigations of the Influence of High Temperature on *A. simplex* Antigen

We performed a comparative analysis of SDS-PAGE and IgG-WB profiles of the following *A. simplex* antigens: (i) native antigen, (ii) antigen heated at 100 °C, and (iii) autoclaved antigen. We showed that antigen heating and autoclaving caused a reduction in the number and intensity of SDS-PAGE and IgG-WB bands. Furthermore, both heating processes caused an appearance of background in IgG-WB profiles. Based on this experiment, we knew that heating caused changes in the antigens, but did not interfere with the ability to bind antibodies. Similar observations regarding the reduction of the number and intensity of bands in electrophoretic and WB profiles of autoclaved *Anisakis* antigen were reported by Carballeda-Sangiao et al. [51] and Klapper et al. [52].

We thought that the background in the IgG-WB profiles of the heated antigens was the result of degradation and/or alteration of some of the *Anisakis* proteins. However, heated antigens were degraded only partially. Similarly to our study, several other investigations concerning the influence of autoclaving on antigens/allergens of tree nuts [53], shrimps [54], lentil, and chickpeas [55] have shown smear appearing on the SDS-PAGE and WB profiles while maintaining their antigenicity.

Based on the calculation of the IntDen values of the SDS-PAGE and IgG-WB profiles, we measured the protein amounts in the antigens and the signal intensity of the immunoreactivity of antigens, respectively. The IntDen values of SDS-PAGE, as well as the IgG-WB profiles, were similar for the native and both heated antigens; therefore, we supposed that the temperature conditions we used did not cause a drastic reduction of antigenicity.

The present study confirmed the immunoreactivity of the heated/autoclaved *Anisakis* antigens with anti-*A. simplex* IgG antibodies that were linked to delayed IgG-based allergy. This type of hypersensitivity is associated with chronic urticaria that is a frequent symptom of anisakiasis [56,57]. Detection of IgG antibodies specific to *Anisakis* is also useful in the serodiagnosis of anisakiasis [42]. The IgE immunoreactivity of heated *Anisakis* antigens was not investigated in this study, and further studies on this issue are needed. However, the IgE immunoreactivity and thermostability of autoclaved Ani s 1 and Ani s 4 have been shown by Carballeda-Sangiao et al. [51]. Results of other studies that investigated the influence of high temperature on the antigen of *A. simplex* [51,52,58–61] were consistent with our findings regarding the high thermal resistance of *Anisakis* antigenic profile.

4.2. Identification and Label-Free Quantification of Proteins, Allergens, and Potential Allergens of *A. simplex*

Mass spectrometry allowed us to detect 470 proteins of *A. simplex*. Another published dataset of proteins derived from heat-treated *Anisakis* larvae consists of 146 proteins detected in nematode extract heated 5 min at 110 °C [62]. There are no other thermal proteome profiles of *A. simplex* reported in the scientific literature. Comparing the total number of detected proteins ($n = 470$) from the present study with the total number of proteins of native *Anisakis* antigen from our previous investigation [23], we could see a 27% decrease in protein number after autoclaving. In the proteomic profiling of native antigens of *A. simplex* [23], we used a slightly different bioinformatic approach; nevertheless, this comparison clearly indicated that autoclaving did not drastically reduce the number of proteins.

Among all the detected proteins, we identified peptides derived from the five following allergens of *A. simplex*: Ani s 1, Ani s 2, Ani s 3, Ani s 4, and Ani s 5. However, it was impossible to confirm the thermostability of the detected allergens since the identification was conducted by the gel-free LC-MS/MS approach, and, using this technology, only peptides could be detected. However, epitopes, as well as possible epitopes, were found in identified allergenic peptides. As was mentioned, among the *A. simplex* allergens from the WHO/IUIS list, until now, only the autoclaving resistance of Ani s 1 and Ani s 4 is known [51], and further thermostability studies of other *Anisakis* allergens are necessary. The total number of allergens detected in the autoclaved antigens ($n = 5$) was 64% lower compared to the number of allergens found in the native *Anisakis* antigen ($n = 14$) [23]. However, it should be emphasized that, as identified in our present study, Ani s 1 and Ani s 2 are particularly harmful as they are major allergens, which cause hypersensitivity reactions in more than 50% of the allergic

population [63]. Ani s 2 and Ani s 3 are also panallergens that are ubiquitously distributed with highly conserved sequences and structures, and, therefore, they are responsible for cross-reactions, even between phylogenetically distant and unrelated organisms [64–66].

In recent studies, novel allergens and novel potential allergens of *Anisakis* have been described. As the list of *A. simplex* allergens seems to be still incomplete, we performed bioinformatic screening of the detected proteins for possible allergenicity. We identified 13 potential allergens of *Anisakis* (see Table 5) using the AllergenOnline.org server, which is commonly used for such computational evaluations [67]. To increase the confidence of the allergenicity predictions, we applied a high level of identity between potential allergens and homologous allergens (70%) as the cut-off, as the level of identity already just above 50% indicates the possibility of cross-reactions [68]. Most of the homologous identified potential allergens were allergens of *Anisakis* ($n = 5$) and *Ascaris* ($n = 2$), which are relative phylogenetic closely related as both nematodes belong to the same order of Ascaridida. Among the potential allergens, we also frequently detected homologs of Acari allergens ($n = 4$). The occurrence of cross-reactions of *A. simplex* antigen with Acari [69] and *Ascaris* [70,71] antigens was experimentally proven.

Three following potential allergens identified in this work: paramyosin (UniProt ID: A0A158PP35), peptidyl-prolyl cis-trans isomerase (UniProt ID: A0A0M3J6G4), and 78 kDa glucose-regulated protein (UniProt ID: A0A0M3K5H6) were also detected by us in the native antigens of *A. simplex* [23]. We identified potential *Anisakis* allergen SXP/RAL-2 family protein 2 isoform 1 (UniProt ID: A0A0M3KA05) in the native antigen of *Contraecaecum osculatum* [23]. Except for these four potential allergens identified in our previous study, the other nine sequences were identified as potential allergens of *Anisakis* for the first time.

To know the content of allergens and potential allergens in heat-treated *A. simplex* larvae, we measured the relative abundance of proteins using mass spectrometry label-free quantification. We analyzed the protein abundance calculated by Mascot software, and this data revealed that among the 25 most abundant proteins were present the following allergens: Ani s 2, Ani s 3, and Ani s 5, and the following potential allergens: tropomyosin (UniProt ID: A0A0M3KCE6), SXP / RAL-2 family protein 2 isoform 1 (UniProt ID: A0A0M3KA05), and paramyosin (UniProt ID: A0A158PP35). Comparing these results with the relative quantification of allergens found among the 25 most abundant proteins of the native antigen of *Anisakis* [23], we could conclude that autoclaving caused a slight reduction in the number of identified allergens. The following five allergens were found in the native antigens: Ani s 8, Ani s 2, Ani s 13, and Ani s 3 (two isoforms).

4.3. Computational Investigations of Detected Proteins, Allergens, and Potential Allergens of *A. simplex*

Especially important bioinformatics analysis was the detection of epitopes in allergenic peptides originated from the autoclaved antigen of *A. simplex*. This investigation confirmed the results of WB and showed that the epitopes of autoclaved peptides were not destroyed. Among all *Anisakis* allergens, T-cell/B-cell epitopes in Ani s 1 [48,49] and Ani s 5 [50] were experimentally verified, and we used these datasets via IEDB. Epitopes were identified in the majority (74%) of Ani s 1 and Ani s 5 peptides (see Table 3). MHC II, T-cell, and B-cell epitopes in Ani s 2, Ani s 3, and Ani s 4 were predicted for the first time. Probable epitopes were found in 61% of autoclaved allergenic peptides (Table 4) using algorithms of DNASTAR Protean 3D software. This software has been successfully used to predict epitopes in many other studies [72–74].

Previously, among all *Anisakis* allergens, the only 3D structure of Ani s 5 has been determined experimentally by nuclear magnetic resonance [50]. Therefore, to acquire more insights into the nature of novel potential allergens, as well as all identified allergens, we predicted 3D models, representing the native conformation of these proteins. The models were predicted with high confidence using an algorithm combining ab initio and homology-based prediction implemented in the Phyre2 server [47]. Pairs of protein models representing the potential allergen and its homologs were subjected to structural alignment within the PyMOL software. Structural alignment analysis confirmed a potential allergen's structural similarities with its homologs.

Computational analyses showed that identified proteins, allergens, and potential allergens were very diverse in properties, such as molecular weights and isoelectric points. High diversity also occurred at the level of protein classification into the InterPro family, the identification of InterPro domains, and allergen assignment according to the AllFam database. For example, the most highly represented family and domain among all detected proteins, i.e., immunoglobulin-like fold (IPR013783) and immunoglobulin-like domain (IPR007110) were assigned only to 4% and 3% of sequences, respectively. This superfamily represents domains with an immunoglobulin-like (Ig-like) fold, and Ig-like domains are one of the most common protein modules found in different organism proteins [75]. Proteins with this fold vary in their cellular localization, amino acid sequence, and biological role [76]. EF-hand domains were relatively abundant among identified proteins (nine sequences). Troponin-like protein (UniProt ID: A0A0M3JU57), a potential allergen detected in this study, and its homolog (Ani s troponin C) also contain EF-hand domains. EF-hand allergen family is the second-largest group of allergens [77] that are detected in different organisms like parvalbumins in a specific species of fish and fungus (*Trichophyton violaceum*) [78] or polcalcins in the pollen of trees, grasses, and weeds [79]. It has been found that antigenic sites of parvalbumins in *Trichophyton violaceum* are located on both sides of the Ca^{2+} -binding site of the first EF-hand domain and parvalbumin proteins possessing conserved amino acid motifs (cysteine, lysine, and arginine) [78]. Noteworthy proteins among the highly abundant proteins families detected by us were heat shock proteins, like HSP20-like chaperones (eight sequences), which were highly represented among all thermostable proteins, while hsp70s were relatively higher abundant (three sequences) in the group of potential allergens. In our previous study, we found that heat shock proteins were also one of the most abundant proteins in the native antigen of *Anisakis* [23]. This is not surprising because HSPs are extremely heterogeneous in nature and function mainly as molecular chaperones that help other proteins maintain their native structure and, especially under stresses, are highly expressed [80–82]. HSPs are major immune dominant antigens in many parasite infections, and they play a key role in host–parasite interactions [83,84]. Allergens belonging to the Hsp70 family are found in a heterogeneous range of sources. Among others, HSP70s are inhalant allergens of house dust mites, storage mites, biting midges, black flies, and cockroaches [20,85,86].

AllFam classification is an effective way to characterize allergens and potential allergens. This was also the case in our study in which, despite a large variety of detected allergen, the following classes were more common: hsp70 (AF002), which is described above, SXP/RAL-2 family (AF137), tropomyosin (AF054), and myosin heavy chain (AF100).

The allergens of the SXP/RAL-2 family include only three allergens: Ani s 5, Ani s 7, and Ani s 8. In addition to the Ani s 5 allergen, in the present study, we identified two potential allergens related to this family. Members of the SXP/RAL-2 family are characterized by the presence of the domain of unknown function (DUF)148 protein [87]. This family of secreted proteins seems to be specific for nematodes, and several members have been reported in animal parasitic nematodes and in *Caenorhabditis elegans* [87–89]. The role of these proteins is unrecognized; however, it is known that the structure of Ani s 5 resembles that of calmodulin but binds Mg^{2+} instead of Ca^{2+} [50].

Among allergens from the tropomyosin class, we found Ani s 3 and two potential allergens that were homologs of Asc l 3. Tropomyosin has been identified as a minor inhalation allergen in arthropods (mites, cockroaches) and as a major food allergen in crustaceans and mollusks [90,91], while vertebrate tropomyosin seems to be non-allergenic [92]. Due to repetitive coiled-coil structures, which were visualized on our 3D models, tropomyosin retains IgE antibodies binding ability even after heat treatment or partial digestion [92,93]. Tropomyosin sequences are highly conserved, which causes frequent cases of hypersensitivity cross-reactions with phylogenetically distant allergens [94].

During the analysis of allergen class of myosin heavy chain (AF100), we assigned Ani s 2 and one potential homolog allergen of Ani s 2 to this group. Both of these proteins were paramyosins. Paramyosin is a filamentous protein that is found in many invertebrates, including parasites. This protein may regulate the host's immune responses by inhibiting the classical pathway of complement cascade through inhibition of the complement C1 function [95]. Paramyosin is engaged in the

immunological protection mechanism of parasites by acting as Fc receptors and has been shown to induce hypersensitivity reactions in humans [96–98].

We performed cellular component GO annotation to predict the localization of detected proteins in relation to cellular compartments and structures. Obtaining this data allowed for a deeper characterization of proteins as it provided context enabling, understanding of their function. The results of the cellular components annotations (see Figures 4 and 7) showed a large variation in the distribution of GO terms. Among all annotations, the most abundant (about one-third of annotations) were the following GO terms: organelle (GO:0043226), intracellular organelle (GO:0043229), and cytoplasm (GO:0005737). The GO term organelle means the organized structure of distinctive morphology and function, includes the nucleus, mitochondria, vesicles, ribosomes, and the cytoskeleton, but excludes the plasma membrane. The GO term intracellular organelle is an organized structure of distinctive morphology and function, occurring within the cell. The GO term cytoplasm is all of the contents of a cell, excluding the plasma membrane and nucleus, but including other subcellular structures.

In our comparison of the cellular component annotations among all identified proteins, the most abundant 25 proteins, allergens, and potential allergens showed some interesting differences. The GO term supramolecular complex (GO:0099080) was the most represented in the 25 most abundant proteins (11% of sequences), followed by allergens (9% of sequences), potential allergens (7% of sequences), and the lowest represented in case of all proteins (5%). The GO term supramolecular complex is a cellular component that consists of an indeterminate number of proteins or macromolecular complexes, organized into regular, higher-order structures, such as a polymer, sheets, networks, or fibers. Among the proteins belonging to this class of GO term, we also assigned the following allergens and potential allergens, which were filament proteins: tropomyosin (including Ani s 3), paramyosin (including Ani s 2), tubulin alpha chain, and troponin-like protein sequences. Slightly higher abundance of sequences of allergens, as well as potential allergens, were assigned to the extracellular region (GO:0005576). The GO term extracellular region is the space external to the outermost structure of a cell; for cells without external protective or external encapsulating structures, this refers to space outside of the plasma membrane. The following proteins were assigned to this GO term: like tubulin alpha chain, SXP/RAL-2 family protein, 78 kDa glucose-regulated protein, DUF148 domain-containing protein, and heat shock 70 kDa protein cognate 1. These results corresponded with the fact that Ani s 1 and Ani s 5 are excretory-secretory allergens.

5. Conclusions

This study provided novel data on the *A. simplex* proteome. Based on mass spectrometry analysis, it could be concluded that 470 proteins were detected in heat-treated *A. simplex* larvae. Among identified proteins, peptides of the following allergens were found: Ani s 1, Ani s 2, Ani s 3, Ani s 4, and Ani s 5. In silico predicted and known epitopes in peptides originated from these allergens were detected using bioinformatics tools. Furthermore, thirteen potential allergens were detected, nine of which were identified for the first time. The identified proteins, allergens, and potential allergens were very diverse in terms of properties, such as their molecular weight, isoelectric point, tertiary structure, domain and family classifications, and cellular component annotations. The reactivity of the autoclaved *A. simplex* antigen with anti-*A. simplex* IgG antibodies that were relevant to delayed IgG-based allergy was confirmed by WB. The IgE-binding capacity and thermostability of identified allergens and potential allergens were not tested in this study, and, therefore, further studies are needed to investigate these aspects.

Due to the presence of epitopes in allergenic peptides derived from the autoclaved antigen, thermally processed fish products that might contain *A. simplex* proteins could be a potential threat to sensitized consumers. These findings have implications for the fish processing industry and food safety authorities. It is necessary to search for more effective methods to reduce the allergenicity of food products contaminated by *Anisakis*. A practical solution to this issue can provide removal of *Anisakis* allergens during the washing of fish muscle, as described by Olivares et al. [99]. Furthermore,

an extensive examination of fish products for *A. simplex* allergens can improve the protection of *Anisakis*-allergic consumers. Hence, the implementation of diagnostic tools for the detection of *A. simplex* allergens is essential for food safety laboratories. Publicly deposited mass spectrometry data could be useful for future studies, such as the development of new diagnostic assays.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2218-273X/10/7/1066/s1>, **File S1:** The list of all proteins identified using LC-MS/MS, and results of the functional analysis performed by OmicsBox software; **File S2:** The structural alignments of the following potential allergens and their homologous allergens (native structures): A0A0M3KCE6 (97% of residues modeled at > 90% confidence, model dimensions in angstroms (Å): X:27.759 Y:173.490 Z:45.247) vs. Asc l 3 (a), A0A0M3KA05 (81% of residues modeled at > 90% confidence, model dimensions in Å: X:75.957 Y:58.757 Z:35.306) vs. Ani s 8 (b), A0A158PP35 (89% of residues modeled at > 90% confidence, model dimensions in Å: X:125.866 Y:222.697 Z:221.902) vs. Ani s 2 (c), A0A0M3J4T5 (67% of residues modeled at > 90% confidence, model dimensions in Å: X:47.481 Y:73.842 Z:41.442) vs. Ani s 2 (d), A0A0M3J693 (97% of residues modeled at > 90% confidence, model dimensions in Å: X:46.182 Y:84.389 Z:117.031) vs. Asc l 3 (e), A0A0M3K9V2 (95% of residues modeled at > 90% confidence, model dimensions in Å: X:65.744 Y:74.745 Z:102.212) vs. Tyr p 28 (f), A0A0M3KIW1 (97% of residues modeled at > 90% confidence, model dimensions in Å: X:52.953 Y:70.661 Z:62.596) vs. Tyr p 28 (g), A0A0M3J6G4 (92% of residues modeled at > 90% confidence, model dimensions in Å: X:39.808 Y:40.355 Z:35.353) vs. Ole e 15 (h), A0A0M3JU57 (91% of residues modeled at > 90% confidence, model dimensions in Å: X:43.290 Y:44.288 Z:50.719) vs. Ani s troponin C (i), A0A0M3K5H6 (89% of residues modeled at > 90% confidence, model dimensions in Å: X:120.529 Y:61.462 Z:88.699) vs. Aed a 8 (j), A0A0M3J8H0 (53% of residues modeled at > 90% confidence, model dimensions in Å: X:49.537 Y:38.719 Z:27.443) vs. Ani s 9 (k), A0A0M3K821 (90% of residues modeled at > 90% confidence, model dimensions in Å: X:62.766 Y:54.503 Z:55.905) vs. Der f 33 (l), A0A0M3KAH2 (91% of residues modeled at > 90% confidence, model dimensions in Å: X:71.490 Y:51.790 Z:60.479) vs. Der f 33-like (m). Models of unknown 3D structures were performed using the Phyre2 server. Model dimensions, coverage, and confidence of structure prediction of homologous allergens for potential allergens are reported in Figures 5 and 6. PyMOL software was used to perform alignment and visualize all structures. Potential allergen structure is presented in cartoon view mode and highlighted in white, while the homologous allergen structure is presented in the ribbon mode view and highlighted in red.

Author Contributions: Conceptualization, M.K., M.R., J.D., and T.C.; methodology, investigation, M.K., M.R., J.D., J.K., J.S., and A.B.; software, M.K. and J.S.; visualization, M.K. and A.B.; supervision, T.C.; writing—original draft, M.K.; writing—review and editing, J.K., J.S., M.R., and T.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Center for Research and Development under the Strategic Program Biostrateg (grant no. 350 BIOSTRATEG2/296211/4/NCBR/2016) and the National Veterinary Research Institute in Puławy, Poland (projects nos.: S/261, and F/109).

Acknowledgments: This paper is part of Maciej Kochanowski's Ph.D. dissertation (National Veterinary Research Institute, Puławy, Poland).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Jenkins, E.J.; Castrodale, L.J.; de Rosemond, S.J.; Dixon, B.R.; Elmore, S.A.; Gesy, K.M.; Hoberg, E.P.; Polley, L.; Schurer, J.M.; Simard, M. Tradition and transition: Parasitic zoonoses of people and animals in Alaska, northern Canada, and Greenland. *Adv. Parasitol.* **2013**, *82*, 33–204. [PubMed]
- Zolfaghari Enameh, R.; Purmonen, S.; Sukura, A.; Parkkila, S. Surveillance and diagnosis of zoonotic foodborne parasites. *Food Sci. Nutr.* **2018**, *6*, 3–17. [CrossRef] [PubMed]
- Trevisan, C.; Torgerson, P.R.; Robertson, L.J. Foodborne Parasites in Europe: Present Status and Future Trends. *Trends Parasitol.* **2019**, *35*, 695–703. [CrossRef] [PubMed]
- Pozio, E. How globalization and climate change could affect foodborne parasites. *Exp. Parasitol.* **2020**, *208*, 107807. [CrossRef]
- Różycki, M.; Biliska-Zajac, E.; Kochanowski, M.; Grądziel-Krukowska, K.; Zdybel, J.; Karamon, J.; Wiśniewski, J.; Cencek, T. First case of *Trichinella spiralis* infection in beavers (*Castor fiber*) in Poland and Europe. *Int. J. Parasitol. Parasites Wildl.* **2020**, *11*, 46–49. [CrossRef]
- Sroka, J.; Karamon, J.; Wójcik-Fatla, A.; Piotrowska, W.; Dutkiewicz, J.; Biliska-Zajac, E.; Zajac, V.; Kochanowski, M.; Dąbrowska, J.; Cencek, T. *Toxoplasma gondii* infection in slaughtered pigs and cattle in Poland: Seroprevalence, molecular detection and characterization of parasites in meat. *Parasites Vectors* **2020**, *13*, 1–11. [CrossRef]

7. Zolfaghari Emameh, R.; Kuuslahti, M.; Näreaho, A.; Sukura, A.; Parkkila, S. Innovative molecular diagnosis of *Trichinella* species based on β -carbonic anhydrase genomic sequence. *Microb. Biotechnol.* **2016**, *9*, 172–179. [[CrossRef](#)]
8. Armentia, A.; Santos, J.; Serrano, Z.; Martín, B.; Martín, S.; Barrio, J.; Fernández, S.; González-Sagrado, M.; Pineda, F.; Palacios, R. Molecular diagnosis of allergy to *Anisakis simplex* and *Gymnorhynchus gigas* fish parasites. *Allergol Immunopathol* **2017**, *45*, 463–472. [[CrossRef](#)]
9. Lim, H.; Jung, B.K.; Cho, J.; Yooyen, T.; Shin, E.H.; Chai, J.Y. Molecular diagnosis of cause of anisakiasis in humans, South Korea. *Emerging Infect. Dis.* **2015**, *21*, 342–344. [[CrossRef](#)]
10. Kunz, S.; Balmer, V.; Sterk, G.J.; Pollastri, M.P.; Leurs, R.; Müller, N.; Hemphill, A.; Spycher, C. The single cyclic nucleotide-specific phosphodiesterase of the intestinal parasite *Giardia lamblia* represents a potential drug target. *PLoS Negl. Trop. Dis.* **2017**, *11*, e0005891. [[CrossRef](#)]
11. Sánchez-Alonso, I.; Carballeda-Sangiao, N.; González-Muñoz, M.; Navas, A.; Arcos, S.C.; Mendizábal, A.; Tejada, M.; Careche, M. Pathogenic potential of *Anisakis* L3 after freezing in domestic freezers. *Food Control* **2018**, *84*, 61–69. [[CrossRef](#)]
12. Panel on Biological Hazards. Scientific opinion on risk assessment of parasites in fishery products. *EFSA J.* **2010**, *8*, 1543. [[CrossRef](#)]
13. Bao, M.; Pierce, G.J.; Strachan, N.J.; Pascual, S.; González-Muñoz, M.; Levsen, A. Human health, legislative and socioeconomic issues caused by the fish-borne zoonotic parasite *Anisakis*: Challenges in risk assessment. *Trends Food Sci. Technol.* **2019**, *86*, 298–310. [[CrossRef](#)]
14. Guardone, L.; Armani, A.; Nucera, D.; Costanzo, F.; Mattiucci, S.; Bruschi, F. Human anisakiasis in Italy: A retrospective epidemiological study over two decades. *Parasite* **2018**, *25*, 41. [[CrossRef](#)] [[PubMed](#)]
15. Orphanet. Prevalence and Incidence of Rare Diseases. Available online: http://www.orpha.net/orphacom/cahiers/docs/GB/Prevalence_of_rare_diseases_by_alphabetical_list.pdf (accessed on 12 June 2020).
16. Bao, M.; Pierce, G.J.; Strachan, N.J.; Martínez, C.; Fernández, R.; Theodossiou, I. Consumers' attitudes and willingness to pay for *Anisakis*-free fish in Spain. *Fish. Res.* **2018**, *202*, 149–160. [[CrossRef](#)]
17. Bao, M.; Pierce, G.J.; Pascual, S.; González-Muñoz, M.; Mattiucci, S.; Mladineo, I.; Cipriani, P.; Bušelić, I.; Strachan, N.J. Assessing the risk of an emerging zoonosis of worldwide concern: Anisakiasis. *Sci. Rep.* **2017**, *7*, 43699. [[CrossRef](#)]
18. Carballeda-Sangiao, N.; Rodríguez-Mahillo, A.I.; Careche, M.; Navas, A.; Caballero, T.; Dominguez-Ortega, J.; Jurado-Palomo, J.; González-Muñoz, M. Ani s 11-like protein is a pepsin-and heat-resistant major allergen of *Anisakis* spp. and a valuable tool for *Anisakis* allergy component-resolved diagnosis. *Int. Arch. Allergy Immunol.* **2016**, *169*, 108–112. [[CrossRef](#)]
19. Baird, F.J.; Su, X.; Aibinu, I.; Nolan, M.J.; Sugiyama, H.; Otranto, D.; Lopata, A.L.; Cantacessi, C. The *Anisakis* transcriptome provides a resource for fundamental and applied studies on allergy-causing parasites. *PLoS Negl. Trop. Dis.* **2016**, *10*, e0004845. [[CrossRef](#)]
20. Fæste, C.K.; Jonscher, K.R.; Dooper, M.M.; Egge-Jacobsen, W.; Moen, A.; Daschner, A.; Egaas, E.; Christians, U. Characterisation of potential novel allergens in the fish parasite *Anisakis simplex*. *EuPA Open Proteom.* **2014**, *4*, 140–155. [[CrossRef](#)] [[PubMed](#)]
21. Llorens, C.; Arcos, S.C.; Robertson, L.; Ramos, R.; Futami, R.; Soriano, B.; Ciordia, S.; Careche, M.; González-Muñoz, M.; Jiménez-Ruiz, Y.; et al. Functional insights into the infective larval stage of *Anisakis simplex* s.s., *Anisakis pegreffii* and their hybrids based on gene expression patterns. *BMC Genom.* **2018**, *19*, 592. [[CrossRef](#)]
22. Arcos, S.C.; Ciordia, S.; Robertson, L.; Zapico, I.; Jiménez-Ruiz, Y.; Gonzalez-Muñoz, M.; Moneo, I.; Carballeda-Sangiao, N.; Rodriguez-Mahillo, A.; Albar, J.P.; et al. Proteomic profiling and characterization of differential allergens in the nematodes *Anisakis simplex sensu stricto* and *A. pegreffii*. *Proteomics* **2014**, *14*, 1547–1568. [[CrossRef](#)] [[PubMed](#)]
23. Kochanowski, M.; Dąbrowska, J.; Różycki, M.; Karamon, J.; Sroka, J.; Cencek, T. Proteomic profiling reveals new insights into the allergomes of *Anisakis simplex*, *Pseudoterranova decipiens* and *Contracaecum osculatatum*. *J. Parasitol.* **2020**, in press.
24. Anibarro, B.; Seoane, F.; Mugica, M. Involvement of hidden allergens in food allergic reactions. *J Investig. Allergol. Clin. Immunol.* **2007**, *17*, 168–172. [[PubMed](#)]
25. Audicana, M.T.; Ansotegui, I.J.; de Corres, L.F.; Kennedy, M.W. *Anisakis simplex*: Dangerous—Dead and alive? *Trends Parasitol.* **2002**, *18*, 20–25. [[CrossRef](#)]

26. Audicana, M.T.; Kennedy, M.W. *Anisakis simplex*: From obscure infectious worm to inducer of immune hypersensitivity. *Clin. Microbiol. Rev.* **2008**, *21*, 360–379. [[CrossRef](#)]
27. Bucci, C.; Gallotta, S.; Morra, I.; Fortunato, A.; Ciacci, C.; Iovino, P. *Anisakis*, just think about it in an emergency! *Int. J. Infect. Dis.* **2013**, *17*, e1071–e1072. [[CrossRef](#)] [[PubMed](#)]
28. Nieuwenhuizen, N.; Lopata, A.L.; Jeebhay, M.F.; De'Broski, R.H.; Robins, T.G.; Brombacher, F. Exposure to the fish parasite *Anisakis* causes allergic airway hyperreactivity and dermatitis. *J. Allergy Clin. Immunol.* **2006**, *117*, 1098–1105. [[CrossRef](#)]
29. Uña-Gorospe, M.; Herrera-Mozo, I.; Canals, M.L.; Martí-Amengual, G.; Sanz-Gallen, P. Occupational disease due to *Anisakis simplex* in fish handlers. *Int. Marit. Health* **2018**, *69*, 264–269. [[CrossRef](#)]
30. Mazzucco, W.; Lacca, G.; Cusimano, R.; Provenzani, A.; Costa, A.; Di Noto, A.M.; Massenti, M.F.; Leto-Barone, M.S.; Lorenzo, G.D.; Vitale, F. Prevalence of sensitization to *Anisakis simplex* among professionally exposed populations in Sicily. *Arch. Environ. Occup. Health* **2012**, *67*, 91–97. [[CrossRef](#)]
31. Barbarroja-Escudero, J.; Sanchez-Gonzalez, M.; Antolin-Amerigo, D.; Rodriguez-Rodriguez, M.; Alvarez-Mon, M. Nonoccupational airborne-induced anaphylaxis caused by *Anisakis simplex*. *J. Investig. Allergol. Clin. Immunol.* **2016**, *26*, 196–197. [[CrossRef](#)]
32. Meseguer, J.; Navarro, V.; Sánchez-Guerrero, I.; Bartolomé, B.; Negro Alvarez, J.N. *Anisakis simplex* allergy and nephrotic syndrome. *Allergol. Immunopathol.* **2007**, *35*, 216–220. [[CrossRef](#)]
33. Cuende, E.; Audicana, M.; Garcia, M.; Anda, M.; Fernández, L.C.; Jimenez, C.; Vesga, J. Rheumatic manifestations in the course of anaphylaxis caused by *Anisakis simplex*. *Clin. Exp. Rheumatol.* **1998**, *16*, 303–304. [[PubMed](#)]
34. Moneo, I.; Carballeda-Sangiao, N.; González-Muñoz, M. New Perspectives on the Diagnosis of Allergy to *Anisakis* spp. *Curr. Allergy Asthma Rep.* **2017**, *17*, 27. [[CrossRef](#)] [[PubMed](#)]
35. Caballero, M.L.; Moneo, I. Several allergens from *Anisakis simplex* are highly resistant to heat and pepsin treatments. *Parasitol. Res.* **2004**, *93*, 248–251. [[CrossRef](#)]
36. Armentia, A.; Martin-Gil, F.; Pascual, C.; Martin-Esteban, M.; Callejo, A.; Martinez, C. *Anisakis simplex* allergy after eating chicken meat. *J. Investig. Allergol. Clin. Immunol.* **2006**, *16*, 258–263. [[PubMed](#)]
37. Verhoeckx, K.C.; Vissers, Y.M.; Baumert, J.L.; Faludi, R.; Feys, M.; Flanagan, S.; Herouet-Guicheney, C.; Holzhauser, T.; Shimojo, R.; van der Bolt, N.; et al. Food processing and allergenicity. *Food Chem. Toxicol.* **2015**, *80*, 223–240. [[CrossRef](#)] [[PubMed](#)]
38. Vojdani, A. Detection of IgE, IgG, IgA and IgM antibodies against raw and processed food antigens. *Nutr. Metab.* **2009**, *6*, 22. [[CrossRef](#)]
39. Sathe, S.K.; Teuber, S.S.; Roux, K.H. Effects of food processing on the stability of food allergens. *Biotechnol. Adv.* **2005**, *23*, 423–429. [[CrossRef](#)]
40. Pedreschi, R.; Nørgaard, J.; Maquet, A. Current challenges in detecting food allergens by shotgun and targeted proteomic approaches: A case study on traces of peanut allergens in baked cookies. *Nutrients* **2012**, *4*, 132–150. [[CrossRef](#)]
41. Di Girolamo, F.; Muraca, M.; Mazzina, O.; Lante, I.; Dahdah, L. Proteomic applications in food allergy: Food allergenomics. *Curr. Opin. Allergy Clin. Immunol.* **2015**, *15*, 259–266. [[CrossRef](#)]
42. Kochanowski, M.; González-Muñoz, M.; Gómez-Morales, M.Á.; Gottstein, B.; Dąbrowska, J.; Różycki, M.; Cencek, T.; Müller, N.; Boubaker, G. Comparative analysis of excretory-secretory antigens of *Anisakis simplex*, *Pseudoterranova decipiens* and *Contracaecum osculatum* regarding their applicability for specific serodiagnosis of human anisakidosis based on IgG-ELISA. *Exp. Parasitol.* **2019**, *197*, 9–15. [[CrossRef](#)] [[PubMed](#)]
43. Zhu, X.; Gasser, R.B.; Podolska, M.; Chilton, N.B. Characterisation of anisakid nematodes with zoonotic potential by nuclear ribosomal DNA sequences. *Int. J. Parasitol.* **1998**, *28*, 1911–1921. [[CrossRef](#)]
44. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675. [[CrossRef](#)] [[PubMed](#)]
45. Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D.J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; et al. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **2019**, *47*, D442–D450. [[CrossRef](#)]
46. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676. [[CrossRef](#)]

47. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)]
48. Kobayashi, Y.; Ikeda, K.; Shiomi, K. Elucidation of IgE-binding epitopes of Ani s 1: The major *Anisakis simplex* allergen. *Mol. Biochem. Parasitol.* **2010**, *174*, 128–131. [[CrossRef](#)]
49. Garcia Alonso, M.; Caballero, M.; Umpierrez, A.; Lluch-Bernal, M.; Knaute, T.; Rodríguez-Pérez, R. Relationships between T cell and IgE/IgG4 epitopes of the *Anisakis simplex* major allergen Ani s 1. *Clin. Exp. Allergy* **2015**, *45*, 994–1005. [[CrossRef](#)]
50. García-Mayoral, M.F.; Treviño, M.A.; Perez-Pinar, T.; Caballero, M.L.; Knaute, T.; Umpierrez, A.; Bruix, M.; Rodríguez-Perez, R. Relationships between IgE/IgG4 epitopes, structure and function in *Anisakis simplex* Ani s 5, a member of the SXP/RAL-2 protein family. *PLoS Negl. Trop. Dis.* **2014**, *8*, e2735. [[CrossRef](#)]
51. Carballeda-Sangiao, N.; Olivares, F.; Rodríguez-Mahillo, A.I.; Careche, M.; Tejada, M.; Moneo, I.; Gonzalez-Muñoz, M. Identification of autoclave-resistant *Anisakis simplex* allergens. *J. Food Prot.* **2014**, *77*, 605–609. [[CrossRef](#)]
52. Klapper, R.; Carballeda-Sangiao, N.; Kuhn, T.; Jensen, H.M.; Buchmann, K.; Gonzalez-Muñoz, M.; Karl, H. Anisakid infection levels in fresh and canned cod liver: Significant reduction through liver surface layer removal. *Food Control* **2018**, *92*, 17–24. [[CrossRef](#)]
53. Cuadrado, C.; Sanchiz, A.; Vicente, F.; Ballesteros, I.; Linacero, R. Changes induced by pressure processing on immunoreactive proteins of tree nuts. *Molecules* **2020**, *25*, 954. [[CrossRef](#)] [[PubMed](#)]
54. Lasekan, A.O.; Nayak, B. Effects of buffer additives and thermal processing methods on the solubility of shrimp (*Penaeus monodon*) proteins and the immunoreactivity of its major allergen. *Food Chem.* **2016**, *200*, 146–153. [[CrossRef](#)] [[PubMed](#)]
55. Cuadrado, C.; Cabanillas, B.; Pedrosa, M.M.; Varela, A.; Guillamón, E.; Muzquiz, M.; Crespo, J.F.; Rodríguez, J.; Burbano, C. Influence of thermal processing on IgE reactivity to lentil and chickpea proteins. *Mol. Nutr. Food Res.* **2009**, *53*, 1462–1468. [[CrossRef](#)] [[PubMed](#)]
56. Viñas, M.; Postigo, I.; Suñén, E.; Martínez, J. Urticaria and silent parasitism by Ascaridoidea: Component-resolved diagnosis reinforces the significance of this association. *PLoS Negl. Trop. Dis.* **2020**, *14*, e0008177. [[CrossRef](#)] [[PubMed](#)]
57. Daschner, A.; de La Osada, F.V.; Leza, B.G. Specific IgG and IgG4 in patients with IgE-antibodies against *Anisakis simplex*: Acute urticaria in gastro-allergic Anisakiasis versus chronic urticaria. *J. Allergy Clin. Immunol.* **2002**, *109*, S126. [[CrossRef](#)]
58. Vidaček, S.; de las Heras, C.; Solas, M.T.; Mendizábal, A.; Rodríguez-Mahillo, A.I.; Tejada, M. Antigenicity and viability of *Anisakis* larvae infesting hake heated at different time-temperature conditions. *J. Food Prot.* **2010**, *73*, 62–68. [[CrossRef](#)] [[PubMed](#)]
59. Tejada, M.; Olivares, F.; de las Heras, C.; Careche, M.; Solas, M.T.; García, M.L.; Fernandez, A.; Mendizábal, A.; Navas, A.; Rodríguez-Mahillo, A.I.; et al. Antigenicity of *Anisakis simplex* s.s. L3 in parasitized fish after heating conditions used in the canning processing. *J. Sci. Food Agric.* **2015**, *95*, 922–927. [[CrossRef](#)]
60. Rodríguez-Mahillo, A.I.; González-Muñoz, M.; de las Heras, C.; Tejada, M.; Moneo, I. Quantification of *Anisakis simplex* allergens in fresh, long-term frozen, and cooked fish muscle. *Foodborne Pathog. Dis.* **2010**, *7*, 967–973. [[CrossRef](#)] [[PubMed](#)]
61. Vidaček, S.; De Las Heras, C.; Solas, M.T.; García, M.L.; Mendizabal, A.; Tejada, M. Viability and antigenicity of *Anisakis simplex* after conventional and microwave heating at fixed temperatures. *J. Food Prot.* **2011**, *74*, 2119–2126. [[CrossRef](#)]
62. Carrera, M.; Gallardo, J.M.; Pascual, S.; González, Á.F.; Medina, I. Protein biomarker discovery and fast monitoring for the identification and detection of Anisakids by parallel reaction monitoring (PRM) mass spectrometry. *J. Proteom.* **2016**, *142*, 130–137. [[CrossRef](#)] [[PubMed](#)]
63. Ubeira, F.M. Travelling with *Anisakis* allergens. *Int. Arch. Allergy Immunol.* **2014**, *163*, 243–244. [[CrossRef](#)] [[PubMed](#)]
64. Guarneri, F.; Guarneri, C.; Benvenga, S. Cross-reactivity of *Anisakis simplex*: Possible role of Ani s 2 and Ani s 3. *Int. J. Dermatol.* **2007**, *46*, 146–150. [[CrossRef](#)]
65. Quiazon, K.M.A.; Zenke, K.; Yoshinaga, T. Molecular characterization and comparison of four *Anisakis* allergens between *Anisakis simplex sensu stricto* and *Anisakis pegreffii* from Japan. *Mol. Biochem. Parasitol.* **2013**, *190*, 23–26. [[CrossRef](#)] [[PubMed](#)]

66. Hauser, M.; Roulias, A.; Ferreira, F.; Egger, M. Panallergens and their impact on the allergic patient. *Allergy Asthma Clin. Immunol.* **2010**, *6*, 1. [[CrossRef](#)] [[PubMed](#)]
67. Goodman, R.E.; Ebisawa, M.; Ferreira, F.; Sampson, H.A.; van Ree, R.; Vieths, S.; Baumert, J.L.; Bohle, B.; Lalithambika, S.; Wise, J.; et al. AllergenOnline: A peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity. *Mol. Nutr. Food Res.* **2016**, *60*, 1183–1198. [[CrossRef](#)] [[PubMed](#)]
68. Aalberse, R.C. Structural biology of allergens. *J. Allergy Clin. Immunol.* **2000**, *106*, 228–238. [[CrossRef](#)]
69. Johansson, E.; Aponno, M.; Lundberg, M.; Van Hage-Hamsten, M. Allergenic cross-reactivity between the nematode *Anisakis simplex* and the dust mites *Acarus siro*, *Lepidoglyphus destructor*, *Tyrophagus putrescentiae*, and *Dermatophagoides pteronyssinus*. *Allergy* **2001**, *56*, 660–666. [[CrossRef](#)]
70. Carballeda-Sangiao, N.; Rodríguez-Mahillo, A.; Puente, S.; Gutiérrez, M.; Moneo, I.; González-Muñoz, M. *Anisakis/Ascaris* IgE ratio improves specificity for the diagnosis of *Anisakis simplex* sensitization in travellers and immigrants. *Acta Trop.* **2014**, *138*, 1–4. [[CrossRef](#)]
71. Maldonado, J.L.; Hita, L.M.; Sáez, V.D.; Almendros, I.M.; López, A.V.; Bueno, M.C. Cross-reactivity between antigens of *Anisakis simplex* s.l. and other ascarid nematodes. *Parasite* **2004**, *11*, 219–223. [[CrossRef](#)]
72. Sun, X.; Shan, X.; Yan, Z.; Zhang, Y.; Guan, L. Prediction and characterization of the linear IgE epitopes for the major soybean allergen β -conglycinin using immunoinformatics tools. *Food Chem. Toxicol.* **2013**, *56*, 254–260. [[CrossRef](#)]
73. Zhang, J.; Cui, Y.; Wu, Y.; Wang, H.; Ke, J. Prediction and identification of B-cell epitopes for tumor necrosis factor- α . *Mol. Med. Rep.* **2017**, *16*, 3439–3444. [[CrossRef](#)] [[PubMed](#)]
74. Yang, J.-m.; Wang, H.-j.; Du, L.; Han, X.-m.; Ye, Z.-y.; Fang, Y.; Tao, H.-q.; Zhao, Z.-s.; Zhou, Y.-l. Screening and identification of novel B cell epitopes in human heparanase and their anti-invasion property for hepatocellular carcinoma. *Cancer Immunol. Immunother.* **2009**, *58*, 1387–1396. [[CrossRef](#)] [[PubMed](#)]
75. Bodelón, G.; Palomino, C.; Fernández, L.Á. Immunoglobulin domains in *Escherichia coli* and other enterobacteria: From pathogenesis to applications in antibody technologies. *FEMS Microbiol. Rev.* **2013**, *37*, 204–250. [[CrossRef](#)] [[PubMed](#)]
76. Halaby, D.; Poupon, A.; Mornon, J.P. The immunoglobulin fold family: Sequence analysis and 3D structure comparisons. *Protein Eng.* **1999**, *12*, 563–571. [[CrossRef](#)] [[PubMed](#)]
77. Lopata, A.; Kleine-Tebbe, J.; Kamath, S. Allergens and molecular diagnostics of shellfish allergy. In *Molecular Allergy Diagnostics*; Springer: Berlin, Germany, 2017; pp. 399–414.
78. Emameh, R.Z.; Masoori, L.; Taheri, R.A.; Falak, R. Identification and characterization of parvalbumin-like protein in *Trichophyton violaceum*. *Fungal Biol.* **2020**, *124*, 592–600. [[CrossRef](#)] [[PubMed](#)]
79. Gangl, K.; Niederberger, V.; Davies, J.; Valenta, R.; Nandy, A. Marker allergens and panallergens in tree and grass pollen allergy. In *Molecular Allergy Diagnostics*; Springer: Berlin, Germany, 2017; pp. 203–226.
80. Sørensen, J.G.; Kristensen, T.N.; Loeschcke, V. The evolutionary and ecological role of heat shock proteins. *Ecol. Lett.* **2003**, *6*, 1025–1037. [[CrossRef](#)]
81. Moseley, P.L. Heat shock proteins and heat adaptation of the whole organism. *J. Appl. Physiol.* **1997**, *83*, 1413–1417. [[CrossRef](#)]
82. Feder, M.E.; Hofmann, G.E. Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. *Annu. Rev. Physiol.* **1999**, *61*, 243–282. [[CrossRef](#)]
83. Bannai, H.; Sakurai, T.; Inoue, N.; Sugimoto, C.; Igarashi, I. Cloning and expression of mitochondrial heat shock protein 70 of *Trypanosoma congolense* and potential use as a diagnostic antigen. *Clin. Diagn. Lab. Immunol.* **2003**, *10*, 926–933. [[CrossRef](#)]
84. Tazir, Y.; Steisslinger, V.; Soblik, H.; Younis, A.E.; Beckmann, S.; Grevelding, C.G.; Steen, H.; Brattig, N.W.; Erttmann, K.D. Molecular and functional characterisation of the heat shock protein 10 of *Strongyloides ratti*. *Mol. Biochem. Parasitol.* **2009**, *168*, 149–157. [[CrossRef](#)] [[PubMed](#)]
85. Aki, T.; Fujikawa, A.; Wada, T.; Jyo, T.; Shigeta, S.; Murooka, Y.; Oka, S.; Ono, K. Cloning and expression of cDNA coding for a new allergen from the house dust mite, *Dermatophagoides farinae*: Homology with human heat shock cognate proteins in the heat shock protein 70 family. *J. Biochem.* **1994**, *115*, 435–440. [[CrossRef](#)] [[PubMed](#)]
86. Chuang, J.G.; Su, S.N.; Chiang, B.L.; Lee, H.J.; Chow, L.P. Proteome mining for novel IgE-binding proteins from the German cockroach (*Blattella germanica*) and allergen profiling of patients. *Proteomics* **2010**, *10*, 3854–3867. [[CrossRef](#)]

87. Rao, K.; Eswaran, M.; Ravi, V.; Gnanasekhar, B.; Narayanan, R.; Kaliraj, P.; Jayaraman, K.; Marson, A.; Raghavan, N.; Scott, A.L. The *Wuchereria bancrofti* orthologue of *Brugia malayi* SXP1 and the diagnosis of bancroftian filariasis. *Mol. Biochem. Parasitol.* **2000**, *107*, 71–80. [[CrossRef](#)]
88. Chandrashekar, R.; Curtis, K.C.; Ramzy, R.M.; Liftis, F.; Li, B.W.; Weil, G.J. Molecular cloning of *Brugia malayi* antigens for diagnosis of lymphatic filariasis. *Mol. Biochem. Parasitol.* **1994**, *64*, 261–271. [[CrossRef](#)]
89. Gallin, M.Y.; Tan, M.; Kron, M.A.; Rechnitzer, D.; Greene, B.M.; Newland, H.S.; White, A.T.; Taylor, H.R.; Unnasch, T.R. *Onchocerca volvulus* recombinant antigen: Physical characterization and clinical correlates with serum reactivity. *J. Infect. Dis.* **1989**, *160*, 521–529. [[CrossRef](#)] [[PubMed](#)]
90. Faber, M.; Pascal, M.; El Kharbouchi, O.; Sabato, V.; Hagendorens, M.; Decuyper, I.; Bridts, C.; Ebo, D. Shellfish allergens: Tropomyosin and beyond. *Allergy* **2017**, *72*, 842–848. [[CrossRef](#)]
91. Wong, L.; Huang, C.H.; Lee, B.W. Shellfish and house dust mite allergies: Is the link tropomyosin? *Allergy Asthma Immunol. Res.* **2016**, *8*, 101–106. [[CrossRef](#)]
92. Reese, G.; Ayuso, R.; Lehrer, S.B. Tropomyosin: An invertebrate pan-allergen. *Int. Arch. Allergy Immunol.* **1999**, *119*, 247–258. [[CrossRef](#)]
93. Kamath, S.D.; Rahman, A.M.A.; Komoda, T.; Lopata, A.L. Impact of heat processing on the detection of the major shellfish allergen tropomyosin in crustaceans and molluscs using specific monoclonal antibodies. *Food Chem.* **2013**, *141*, 4031–4039. [[CrossRef](#)]
94. Bessot, J.; Metz-Favre, C.; Rame, J.; De Blay, F.; Pauli, G. Tropomyosin or not tropomyosin, what is the relevant allergen in house dust mite and snail cross allergies? *Eur. Ann. Allergy Clin. Immunol.* **2010**, *42*, 3–10. [[PubMed](#)]
95. Laclette, J.; Shoemaker, C.; Richter, D.; Arcos, L.; Pante, N.; Cohen, C.; Bing, D.; Nicholson-Weller, A. Paramyosin inhibits complement C1. *J. Immunol.* **1992**, *148*, 124–128. [[PubMed](#)]
96. Kalinna, B.; McManus, D. An IgG (Fc γ_2)-binding protein of *Taenia crassiceps* (Cestoda) exhibits sequence homology and antigenic similarity with schistosome paramyosin. *Parasitology* **1993**, *106*, 289–296. [[CrossRef](#)]
97. Park, T.-J.; Kang, J.-M.; Na, B.-K.; Sohn, W.-M. Molecular cloning and characterization of a paramyosin from *Clonorchis sinensis*. *Korean J. Parasitol.* **2009**, *47*, 359–367. [[CrossRef](#)]
98. Pérez-Pérez, J.; Fernández-Caldas, E.; Marañón, F.; Sastre, J.; Bernal, M.L.; Rodríguez, J.; Bedate, C.A. Molecular cloning of paramyosin, a new allergen of *Anisakis simplex*. *Int. Arch. Allergy Immunol.* **2000**, *123*, 120–129. [[CrossRef](#)]
99. Olivares, F.; González-Muñoz, M.; Carballeda-Sangiao, N.; Rodríguez-Mahillo, A.; Careche, M.; de las Heras, C.; Navas, A.; Tejada, M. Removal of *Anisakis simplex* allergens from infected fish during the washing step of surimi production. *J. Sci. Food Agric.* **2015**, *95*, 2626–2631. [[CrossRef](#)] [[PubMed](#)]

