Immunology Letters 213 (2019) 46–54

Contents lists available at ScienceDirect

Immunology Letters

journal homepage: www.elsevier.com/locate/immlet

IgE binding activities and *in silico* epitope prediction of Der f 32 in *Dermatophagoides farinae*

Yuwei Li[a,b,c,1], Yuwei Wang[a,1], Pixin Ran[c], Pingchang Yang[a,*], Zhigang Liu[a,b,*]

[a] State Key Laboratory of Respiratory Disease for Allergy at Shenzhen University, Shenzhen University School of Medicine, Shenzhen 518060, China
[b] The Third Affiliated Hospital of Shenzhen University, Shenzhen 518020, China
[c] State Key Laboratory of Respiratory Disease, Guangzhou Medical College, Guangzhou, 510006, China

ARTICLE INFO

Keywords:
Der f32
Allergen
Homology modeling
B-cell epitope
T-cell epitope
Epitope
Prediction

ABSTRACT

Dermatophagoides farinae is a common indoor allergen source that produces more than 30 allergens, which induces diverse allergic diseases such as allergic rhinitis, allergic asthma and atopic dermatitis. Der f 32 is an inorganic pyrophosphatase and an important allergen from *Dermatophagoides farinae*. In the present study, Der f 32 was cloned, expressed and purified in order to better understand its structure and immunogenicity. Immunoblotting analysis and ELISA showed 5 of 5 positive reactions to recombinant Der f 32 using serum from house dust mite (HDM)-allergic patients. We constructed homology modeling and predicted epitopes of Der f 32 via bioinformatic tools. The sequence and structural analysis indicated that Der f 32 belonged to the pyrophosphatase family and represented a special structure of external α-helices and internal antiparallel closed β-sheets. In addition, eight B-cell epitopes and four T-cell epitopes were predicted. B-cell epitopes were 24–31, 111–121, 135–140, 168–172, 200–207, 214–220, 237–243, and 268–274 and T-cell epitopes were 47–55, 78–90, 127–135 and 143–151. The B-cell epitopes were distributed completely on the surface of Der f 32 and were located largely in random coils of secondary structures. Hydrophobic and charged amino acids comprised more than 80% of the residues of B-cell epitopes and may participate in IgE binding. The T-cell epitopes were located primarily in the interior of Der f 32 and, to a certain extent avoided degradation by proteases. The structures of T-cell epitopes were surrounded by B-cell epitopes, and this arrangement may have important biological significance for maintaining the immunogenicity of allergens.

1. Introduction

In 1922, Cooke reported that house dust encompassed multiple allergens connected to asthma [1]. Several decades later, Voorhorst and Miyamoto identified that house dust mites (HDMs) produced allergens as a causative antigen in bronchial asthma [2,3]. HDMs represent an indoor major allergen source for humans, which induce allergic diseases such as asthma, dermatitis, and rhinitis [4,5]. Thirty-five allergens from the Der f group (Der f 1–4, 6–8, 10–11, 13–18, and 20–36) have been identified and named in the Allergen Nomenclature Database (http://www.allergen.org). Among the novel identified *D. farinae* allergens, Der f 32 is a pyrophosphatase (PPase) with a molecular weight of 34 kDa, and a previous study showed that 8/52 individuals had positive reactions to r–Der f 32 via a skin prick test (http://www.allergen.org/viewallergen.php?aid=817). Der f 32 represents the only allergen possessing PPase activity in the Allergen Nomenclature Database.

PPase is an enzyme (EC 3.6.1.1) that catalyzes the conversion of one molecule of pyrophosphate to two phosphate ions [6], and plays a critical role in lipid metabolism, calcium absorption, bone formation, and DNA synthesis [7–9]. PPases has been found in nearly every organism. Some PPases have been characterized from bacteria, as well as from eukaryotes such as plants [10,11]. However, few PPases have been identified as allergens, except for those found in *D. farinae*. Given that PPases are essential for multiple metabolic processes in living cells, as well as lipid metabolism and DNA synthesis, immunologic characterization of PPases were essential.

Currently, allergen–specific immunotherapy (AIT) is the only approach for determining the cause of allergic diseases, for which subcutaneous injections of increasing doses of allergen extracts to allergic patients has been the most commonly applied traditional method.

* Corresponding authors at: State Key Laboratory of Respiratory Disease for Allergy at Shenzhen University, Shenzhen University School of Medicine, Shenzhen 518060, China.
E-mail addresses: pcy2356@szu.edu.cn (P. Yang), lzg@szu.edu.cn (Z. Liu).
[1] These authors equally contributed to this work.

https://doi.org/10.1016/j.imlet.2019.07.006
Received 2 June 2019; Received in revised form 20 July 2019; Accepted 29 July 2019
Available online 02 August 2019

However, HDM crude extracts can induce severe anaphylactic side effects or lead to sensitization toward new allergens [12,13]. AIT development strategies are now centered on identifying epitopes responsible for allergic responses and designing of appropriate hypoallergenic AIT vaccines, such as T-cell peptides or allergen-derived B-cell peptides [14,15]. However, no study has yet determined the epitope of the Der f 32 allergen. In the present study, we cloned, expressed, and purified Der f 32. IgE binding activities of Der f 32 were analyzed by immunoblotting analysis and ELISA. We used bioinformatics to predict the secondary and tertiary protein structures of Der f 32 and to identify the B-cell and T-cell epitopes. Our results suggest that these identified Der f 32 epitopes may have potential utility in a peptide-based vaccine design for mite allergies.

2. Materials and methods

2.1. Materials

The sera of five patients with dust-mite allergic disorders were obtained from the First Affiliated Hospital of Guangzhou Medical University. The sera from two non-allergic individuals were used as normal controls. A written informed consent was obtained from each human participant for the use of blood samples. The using human sera in the present study was approved by the Human Ethic Committee at Shenzhen University (Shenzhen, Guangdong, China). *Escherichia coli* BL21 and pET-24a(+) were provided by Institute of Allergy and Immunology, Shenzhen University. Peroxidase-labeled mouse anti-human IgE antibody were from Southernbiotech, USA (9160-05). All other chemicals were all purchased from Sigma-Aldrich (St. Louis, MO, USA).

2.2. Sequence retrieval and phylogenetic analysis

The Der f 32 protein sequence was obtained from the GenBank Nucleotide with the accession number of KM009993.1. The homologous amino acid sequences were selected by BLAST in NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). A phylogenetic tree was generated by using Neighbor Joining (NJ) phylogenetic analysis on the basis of the JTT amino acid sequence distance implemented in MEGA 5.1 [16]. The multiple alignment of homologous amino acid sequences was performed using Clustal X 2.1 [17]. The reliability was evaluated by the bootstrap method with 1000 replications. The sera of patients with dust-mite allergic disorders were obtained from the First Affiliated Hospital of Guangzhou Medical University. The sera from non-allergic individuals were used as normal controls. A written informed consent was obtained from each human participant for the use of blood samples. The using human sera in the present study was approved by the Human Ethic Committee at Shenzhen University (Shenzhen, Guangdong, China).

2.3. Cloning, expression, purification of Der f 32

The gene of Der f 32 (KM009993.1) was subcloned into pET-24a(+) and was expressed in *E. coli* BL21 cells. Expression of Der f 32 was induced by adding 0.1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG) at 30 °C for 16 h, when the cells exhibited an absorbance of 0.8 at an optical density (OD) of 600 nm. The bacterial cells were harvested by centrifugation at 5000 rpm at 4 °C for 20 min and resuspended in lysis buffer (50 mM Tris-HCl, 100 mM NaCl, pH7.5). The cells were sonicated at an amplitude of 30% for 10 min (3 s pulse-on and 3 s pulse-off) followed by centrifugation at 12,000 rpm at 4 °C for 20 min. The supernatant was loaded on a Ni-NTA-agarose column (Qiagen, Valencia, CA, USA) equilibrated in lysis buffer, washed with lysis buffer. The eluted fractions washed with lysis buffer containing 300 mM imidazole. The recombinant protein was analyzed by 12% SDS-PAGE stained with Coomassie Blue. For immunoblotting analysis, the r-Der f 32 was fractioned by SDS-PAGE, electro-transferred onto a nitrocellulose membrane, blocked for 2 h with 3% bovine serum albumin (BSA) at 25 °C, and was then incubated with the patient's sera. After washing three times with TBS-T (Tris-buffered saline containing triton X-100), the membrane was incubated with peroxidase-labeled mouse anti-human IgE for 60 min at 25 °C. After washing three times with TBS-T, the membrane was developed with a DAB kit (Invitrogen, Carlsbad, CA, USA).

2.4. Enzyme-linked immunosorbent assay (ELISA)

The specific IgE in patient serum to Der f 32 was also detected by indirect ELISA kits (Ebioscience, USA). Briefly, 100 µL of mixture including carbonate buffered solution (pH 9.5) and 100 ng of purified Der f 32 was adsorbed in 96-well microtiter plates for 12 h at 4 °C. Then, 200 µL of 3% BSA in PBS was used for blocking the plate at 37 °C for 2 h. Subsequently, the allergic serum (diluted to five times with 0.05 M PBS containing 3% BSA) was incubated overnight at 4 °C. The plates were incubated at 37 °C with a peroxidase-labeled mouse anti-human IgE (1:2000) monoclonal antibody for 60 min and were then washing three times with PBST. Tetramethylbenzidine (TMB; 100 µL/well) was added to each well as a color-change indicator and the reaction was stopped with 50 µl of 2 M H_2SO_4. An ELx808 absorbance microplate reader (BioTek, Shanghai, China) was used to measure the absorbance value at 450 nm. All of the data are expressed as mean ± SEM and processed with Graphpad software.

2.5. Physiochemical analysis and post-translational patterns and motifs

Physicochemical analysis of the Der f 32 protein sequence was performed by ProtParam (http://web.expasy.org/protparam/) and included the molecular weight, amino acid composition, negatively charged residues (Asp + Glu), positively charged residues (Arg + Lys), theoretical isoelectric point (pI), aliphatic index, grand average of hydropathicity (GRAVY), and the instability index of Der f 32 [18,19]. Characteristic patterns and functional motifs of Der f 32 were checked by Pfam v27.0 [20], Prosite [21], and InterPRO v46.0 [22].

2.6. Prediction of secondary structures

Secondary structures of Der f 32 were predicted by PSIPRED, which compared sequence segments with a template protein structure of high-sequence similarity to identify conserved substructures; PSIPRED has the highest published score among secondary structural prediction methods [23]. Furthermore, secondary structural elements were also identified by NetSurfP ver. 1.1, with high accuracy and precision for predicting secondary-structures [24].

2.7. Tertiary structure prediction and validation

Homology modeling was used for constructing the tertiary structure of Der f 32. The homologous templates suitable for Der f 32 were selected by the PSI-BLAST server and SWISS-MODEL server [25,26]. The best template was retrieved from the results of previous methods, which were based on the high score, lower e-value, and maximum sequence identity. Tertiary structure was constructed by MODELLER v9.20 [27], which was submitted to Chiron to correct disallowed clashes and to enhance the distribution of stereochemistry [28]. Chiron efficiently resolves severe clashes in homology models with minimal perturbation in the protein backbone [29]. PROCHECK [30], ERRAT [31], and VERIFY 3D programs [32] in the structural model analysis and verification server (SAVES) were used to check the initial structural model of Der f 32. The ultimate tertiary structure of Der f 32 was evaluated by QMEAN [33], by assessing protein stereology through ProSA [34]. A Ramachandran plot for all of the models was generated, and it showed that the majority of the protein residues were in the favored, allowed,

and/or disallowed regions.

2.8. In silico prediction of B-cell epitopes

BPAP, BcePred, DNAStar protean system, and the BepiPred 1.0 server were used for predicting the B-cell epitopes of Der f 32. Hydrophilicity, flexibility, accessibility, and antigenicity of the amino acid sequence were chosen as parameters for B-cell epitopes in the DNAStar protean system [35]. BcePred predicted B-cell epitopes using the same physicochemical properties [36]. BPAP predicted B-cell epitopes in antigen sequences using an artificial neural network [37]. The amino acid sequence was only provided for the BepiPred 1.0 server. Epitope prediction was combined with physicochemical properties of amino acids, such as hydrophilicity, flexibility, accessibility, turns, and exposed surfaces [35]. The consensus epitopes were accurate through combining the results of the four tools, as the ultimate predicting results of B cell epitopes with a previously published method [38]. If the predicting results of all four tools were non epitope, then the consensus result was 0% B cell epitopes : if the results were only one or no non-epitope, the consensus result was 75 or 100% B cell epitopes, respectively [39]. Ultimately, the results that consensus epitope result was 75 or 100% were determined as the ultimate epitope results [38].

2.9. In silico prediction of T-cell epitopes

T-cell epitopes are principally predicted on the basis of identifying the binding of amino acid fragments to the MHC complexes that can activate T-cells. The binding strength of each peptide to the given MHC was estimated by NetMHCIIpan-3.0 and NetMHCII 2.2 at a set threshold level. The NetMHCIIpan-3.0 (http://www.cbs.dtu.dk/services/NetMHCIIpan/) was applied for HLA-DR-based T-cell epitope prediction [40]. The NetMHCII-2.2 (http://www.cbs.dtu.dk/services/NetMHCII/) was used for HLA-DQ alleles [41]. In the present study, HLA-DR 101, HLA-DR 301, HLA-DR 401, and HLADR 501 were used to predict HLA-DR-based T-cell epitope prediction through IC50 values below 50 nM. After assembling and integrating the four results, if three of them was epitope that was the consensus epitope, the HLA-DR-based T-cell epitope results were obtained. This method was also used for HLA-DQ-based T-cell epitope prediction. HLA-DQA10101-DQB10501, HLADQA10501-DQB10201, HLA-DQA10501-DQB10301, and HLA-DQA10102-DQB10602 were used to predict HLA-DQ-based T-cell epitopes. As a result, the ultimate consensus T-cell epitope results were obtained by combining the results of the HLA-DR allele epitopes and HLA-DQ allele epitopes. B-cell and T-cell epitopes predicted by computational tools were mapped onto a linear sequence and a three-dimensional model of Der f 32 to confirm their positions and secondary structural elements.

3. Results

3.1. Sequence retrieval and sequence analysis of Der f 32

The amino acid sequence of Der f 32 was obtained from the nucleotide database of NCBI. In order to determine the relationships between Der f 32 and its homologous sequences, 36 homologous sequences were obtained from Uniprot and tBLASTn. The evolutionary tree was inferred through the NJ method in MEGA. Phylogenetic analysis showed that Der f 32 and PPases were clustered into the same group (Fig. 1). Moreover, domain analysis showed that Der f 32 belongs to the inorganic pyrophosphatase superfamily (IPR036649) and PPase family (IPR008162). Multiple alignments of the other five PPases were performed using Clustal X (Fig. 2). After searching for characteristic motifs and/or patterns, we found that Der f 32 exhibited the specific PPase pattern, PS00387 (121–127, A). The primary structure of Der f 32 contained 296 amino acids, among which there were 47 negatively charged amino acids (D and E) and 38 positively charged amino acids

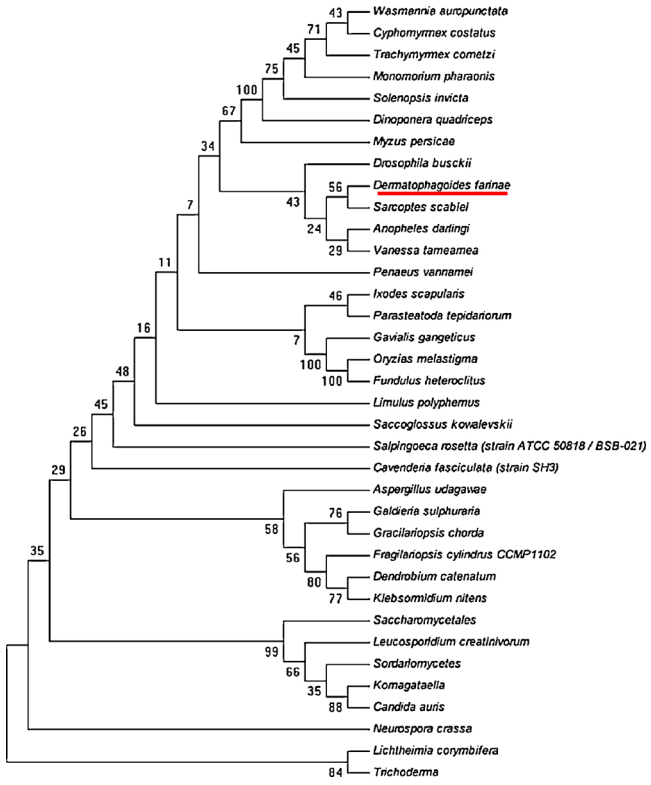

Fig. 1. Evolutionary tree of Der f 32. Phylogenetic relationship of Der f 32 with other homologs. Der f 32 are highlighted by red underline.

(R and K). ProtParam analysis showed that Der f 32 had a calculated molecular weight of 33972.39 Da, a theoretical pI of 5.54, an aliphatic index of 78.72, an instability index of 27.21, and grand average of hydropathicity (GRAVY) of -0.636.

3.2. Allergenicity assessment of Der f 32

Der f 32 was cloned, expressed and purified. The gene of Der f 32 was subcloned into pET-24a(+) and transformed into *E. coli* BL21. The purified r-Der f 32 was examined by SDS-PAGE (Fig. 3A). Immunoblotting and ELISA were performed to evaluate the allergenicity of r-Der f 32 using the sera from dust-mite allergic patients and normal individuals. The results showed that all five allergic patients had positive reactions to r-Der f 32, as indicated from immunoblotting assays and ELISAs (Fig. 3B and C). The IgE reactivity of r-Der f 32 from dust-mite allergic patients was more than five-fold that of healthy subjects (Fig. 3C). These results indicated that r-Der f 32 was an important allergen from *D. farinae*.

3.3. Homology modeling and validation

Homo sapien PPase, Protein Data Bank (PDB) accession number: 6C45.B, was searched as a template and showed the highest sequence identity (52.3%) and coverage (96%) with Der f 32 in the PDB yielded in the SWISS-MODEL server. Thus, the 6C45.B was used as template for homology modeling. The overall three-dimensional (3D) structure of Der f 32 is shown in Fig. 4A. The overall folding of Der f 32 was found to consist of six external α-helices, as well as an internal antiparallel closed β-sheet. Sequence polymorphism was responsible for the changes in the spatial distribution of the skeleton alpha carbons, which was reflected in differences between the structures of Der f 32 and 6C45.B. (Fig. 4B). As a PPase, the characteristic pattern of Der f 32 predicted by ScanProsite tool is shown in Figs. 4C and 2 . The characteristic pattern of PPase was a special structure in the PPase family.

Fig. 2. Sequences and two-dimensional structure analysis of PPases from different species. Sequence comparisons of different PPases (*D. farinae*, XP_011686954.1; *Ixodes scapularis*, XP_002409166.1; *Wasmannia auropunctata*, XP_011686954.1; *Candida auris*, XP_018169651.1; *Lichtheimia corymbifera*, CDH55747.1; *Gracilariopsis chorda*, PXF45408.1) performed by multiple alignments. Residues conserved among the six species are highlighted by blue font, while those only conserved in the first five species are highlighted by gray font. Two-dimensional elements are depicted as blue barrels (α-helices) and green arrows (β-sheets). The characteristic pattern of PPases are highlighted by orange background. Sequences of T-cell and B-cell epitopes are framed in different colors and are labeled as T1–T4 and B1–B8, respectively.

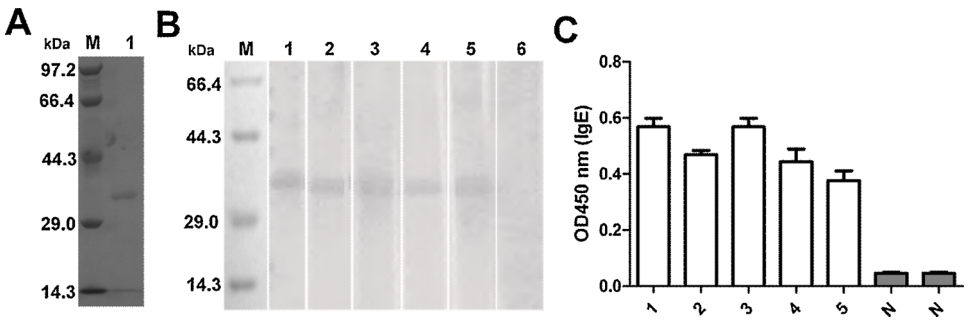

Fig. 3. Immunological characterization of r-Der f 32. (A) SDS-PAGE analysis of purified r-Der f 32. Lane M denotes the protein marker, whereas Lane 1 shows purified r-Der f 32. (B) Immunoblotting analysis of specific IgE reactivity to allergen r-Der f 32. Lane M denotes the protein marker, Lane 1–5 show the sera from dust-mite allergic patients, Lane 6 shows the serum from a healthy subject. (C) The specific IgE reactivity to Der f 32 by ELISA. N, the serum from healthy subjects; 1–5, the sera from dust-mite allergic patients.

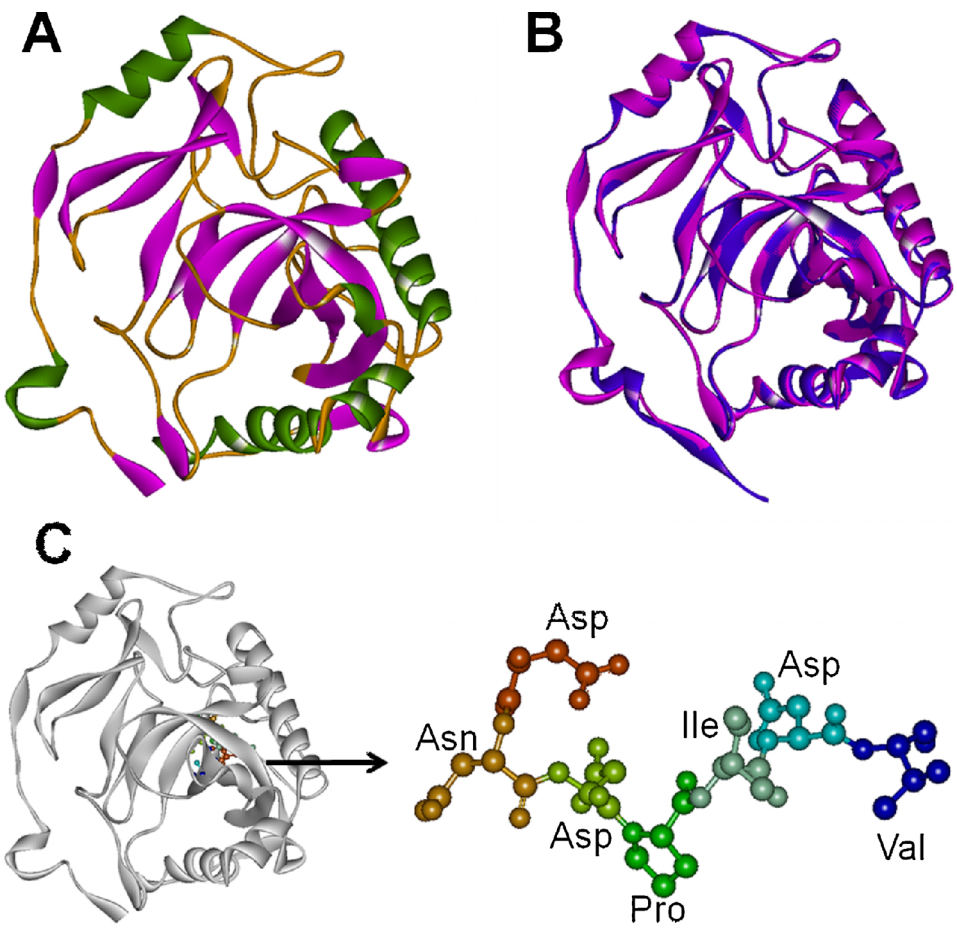

Fig. 4. Structural features of Der f 32. (A) Protein structure of Der f 32 homology model. Cartoon representation of Der f 32 shows an antiparallel closed β-sheet. The α-helices and β-strands in the Der f 32 conformation is highlighted in green and purple, respectively. (B) Superimposition between Der f 32 and the 6C45.B template. Der f 32 is depicted in purple and the 6C45.B template is depicted in blue. (C) Distribution of characteristic pattern in Der f 32.

Table 1
Secondary structural elements of Der f 32.

Structure	α-helices (%)	β-sheets (%)	Random coils (%)
Secondary structure	19.93 (6 domains)	25.34(11 domains)	54.73
Tertiary structure	16.55 (6 domains)	24.66(15 domains)	58.78

The secondary structure of Der f 32 was predicted to have 6 *α*-helices and 11 *β*-sheets by PSIPRED. However, 6 *α*-helices and 15 *β*-sheets were identified by NetSurfP v1.1 (Table 1).

The Ramachandran plot of the Der f 32 model indicated that 85% of the amino acid residues of Der f 32 were located in the most favored regions, 13.8% were within the additional allowed region, 0.8% were in generously allowed regions, and 0.4% were in the disallowed region (Table 2 and Fig. 5A). Comparatively, 88.3% of the residues in the 6C45.B template were within the most favored regions, whereas 10.9%, 0.8% and 0% of the 6C45.B residues were in the additional allowed region, generously allowed regions, and the disallowed region, respectively (Table 2 and Fig. 5B). The distribution of the Der f 32 stereo–chemical parameters was calculated by *G*-factor, and included dihedral bonds (-0.24), covalent bonds (0.21), and overall score (-0.19), in comparison with the template 6C45.B (Table 2). The overall quality factor of the ERRAT program was 90.7 for Der f 32 and 93.9 for 6C45.B. As indicated by the VERIFY 3D program, 88.54% of Der f 32 residues had an average 3D (atomic model) one-dimensional (amino acid sequence) score > 0.2, which indicated that those two structures were available. ProSa server showed that the Z-scores of Der f 32 and 6C45.B were -7.38 and -7.82, respectively. The QMEAN Q value of Der f 32 and 6C45.B was 0.781 and the Q value was 0.719. The tertiary structure of Der f 32 is shown in Fig. 4A.

Table 2
Parameters used for proteins structural assessment.

Protein	Structural assessment methods	Ramachandran plot (%)	*G*-factor	*Z*-score	*Q* value
Der f 32	PROCHECK analysis	85.0[E] 13.8[F] 0.8[G] 0.4[H]	−0.24[I] 0.21[J] −0.19[K]		
	ProSa QMEAN			−7.38	0.781
6C45.B	PROCHECK analysis	88.3[E] 10.9[F] 0.8[G] 0.0[H]	−0.25[I] 0.60[J] 0.05[K]		
	ProSa QMEAN			−7.82	0.719

[E] Residues in favorable regions; [F] residues in allowed regions; [G] residues in generally allowed regions; [H] residues in disallowed regions; [I] *G*-factor score of the dihedral bonds; [J] *G*-factor score of the covalent bonds; [K] overall *G*-factor score.

3.4. B-cell epitopes prediction

Hydrophobicity, surface accessibility, and segmental flexibility are important features for epitope identification. The antigenic index directly indicated the epitope-forming capacity of the Der f 32 sequence. Based on these sequence properties, the final epitope regions of the DNAStar prediction were 10–15, 23–30, 44–47, 58–61, 109–126, 135–140, 154–158, 168–172, 175–179, 199–207, 214–220, 237–243,

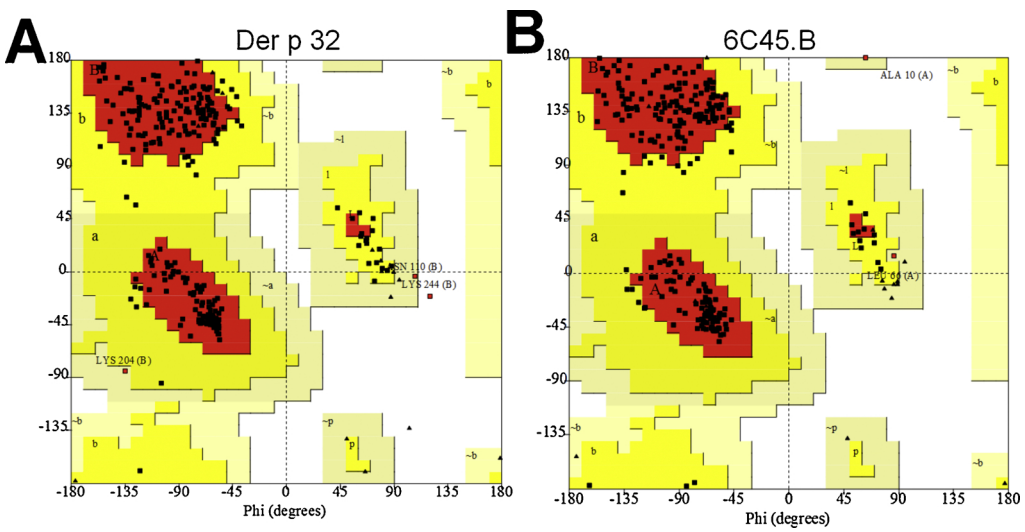

Fig. 5. Ramachandran plot of Der f 32 and the 6C45.B template. (A) Ramachandran plot of Der f 32. (B) Ramachandran plot of template 6C45.B. The residues in the most favored regions, additional allowed regions, generously allowed regions and disallowed regions are shown in red, yellow, grey and white in the Ramachandran plot, respectively.

Table 3
The results of B-cell epitopes and T-cell epitopes predictions.

	Tools	Location of the prediction results
B-cell epitope prediction	DNAStar protean	10-15, 23-30, 44-47, 58-61, 109-126, 135-140, 154-158, 168-172, 175-179, 199-207, 214-220, 237-243, 257-264, 268-278.
	BPAP	4-10, 15-21, 31-45, 48-55, 79-93, 124-134, 136-151, 178-199, 243-249, 272-278, 280-291.
	BepiPred	1-9, 26-31, 57-60, 63-79, 101-125, 135-140, 154-158, 168- 172, 175- 179, 200-220, 237-243, 257-264, 268-284.
	Bcepred	7-13, 22-29, 56-60, 72-79, 111-121, 131-137, 164-178, 199-204, 211-218, 234-241, 267-273.
T-cell epitope prediction (HLA-DR)	DRB1∗01:01	49-57, 78-86, 83-91, 95-103, 127-135, 142-150, 143-151, 144-152, 180-188, 181-189, 183-191, 192-200.
	DRB3∗01:01	6-14, 14-22, 15-23, 21-29, 22-30, 34-42, 40-48, 63-71, 82-90, 95-103, 159-167, 161-169, 162-170, 187-195, 224-232, 256-264, 285-293.
	DRB4∗01:01	50-58, 51-59, 78-86, 82-90, 83-91, 127-135, 143-151, 144-152, 145-153, 159-167, 182-190, 183-191, 187-195, 192-200, 194-202.
	DRB5∗01:01	14-22, 40-48, 49-57, 78-86, 79-87, 82-90, 85-93, 127-135, 128-136, 143-151, 159-167, 180-188, 181-189, 188-196, 194-202.
T-cell epitope prediction (HLA-DQ)	HLA-DQA10101-DQB10501	17-25, 19-27, 31-39, 32-40, 33-41, 35-43, 47-55, 95-103, 96-104, 97-105, 144-152, 156-164, 172-180, 187-195, 192-203, 228-236, 229-237.
	HLA-DQA10102-DQB10602	47-55, 60-68, 61-69, 62-70, 63-71, 145-153, 192-200.
	HLA-DQA10501-DQB10201	34-42, 47-55, 48-56, 50-58, 62-70, 63-71, 95-103, 98-106, 137-145, 145-153, 146-154, 148-156, 159-167, 161-169, 162-170, 173-181, 180-188, 187-195, 219-227, 220-228, 224-232, 228-236, 255-263, 280-288, 282-290.
	HLA-DQA10501-DQB10301	26-34, 27-35, 63-71, 65-73, 66-74, 78-86, 79-87, 92-100, 97-105, 98-106, 128-136, 133-141, 135-143, 136-144, 183-191, 184-192.

Table 4
Predicted B-cell and T-cell epitopes of Der f 32.

Peptide	Type of epitope	Position	Sequence
P1	B	24-31	**KD***NSNGKI*
P2	B	111-121	*HI***DQDTKTKG***D*
P3	B	135-140	***AKRG****DV*
P4	B	168-172	***DE****LAS*
P5	B	200-207	*IP***DGK***PAN*
P6	B	214-220	***EAKDRE****F*
P7	B	237-243	***ENKSGEH***
P8	B	268-274	***ETRPSSD***
P9	T	47-55	***KH****YNMVVEI*
P10	T	78-90	*IKKGAL****RYVKNVF***
P11	T	127-135	*VIEIGS****RVA***
P12	T	143-151	*VKILGTIAL*

Charged residues are shown in a bold font; hydrophobic residues are depicted in italic.

257–264, and 268–278 (Table 3). The predicted epitope region of the BPAP system were 4–10, 15–21, 31–45, 48–55, 79–93, 124–134, 136–151, 178–199, 243–249, 272–278, 280–291. The predicted epitope regions in the BepiPred 1.0 server were 1–9, 26–31, 57–60, 63–79,101–125, 135–140, 154–158, 168–172, 175–179, 200–220, 237–243, 257–264, and 268–284. The predicted epitope regions in the Bcepred server were 7–13, 22–29, 56–60, 72–79, 111–121, 131–137, 164–178, 199–204, 211–218, 234–241 and 267–274. Through integrating the results from the four programs, the final potential B-cell epitopes of Der f 32 were selected. The final results of the three-tool prediction yielded eight peptides (24–31, 111–121, 135–140, 168–172, 200–207, 214–220, 237–243, and 267–274; Table 4), and these peptides are also shown in Figs. 2 and 6.

3.5. T-cell epitopes prediction

T-cell epitopes of Der f 32 were identified by NetMHCIIpan-3.0 and NetMHCII-2.2. For HLA-DR-based T-cell epitope prediction, the final prediction results of HLA-DR 101, HLA-DR 301, HLA-DR 401, and HLA-DR 501 are listed in Table 3. The results of the HLA-DR-based T-cell epitope prediction yielded three peptides (78–90, 127–135 and 143–151). For HLA-DQ-based T-cell epitope prediction, the final prediction results of HLADQA10101-DQB10501, HLA-DQA10501-DQB10201, HLADQA10501-DQB10301, and HLA-DQA10102-DQB10602 are also shown in Table 3. The results of the HLA-DQ-based

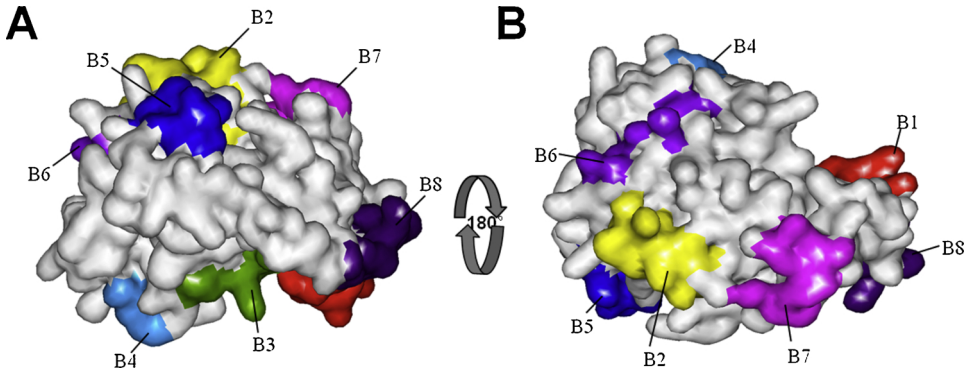

Fig. 6. B-cell epitopes in the tertiary structure of Der f 32.

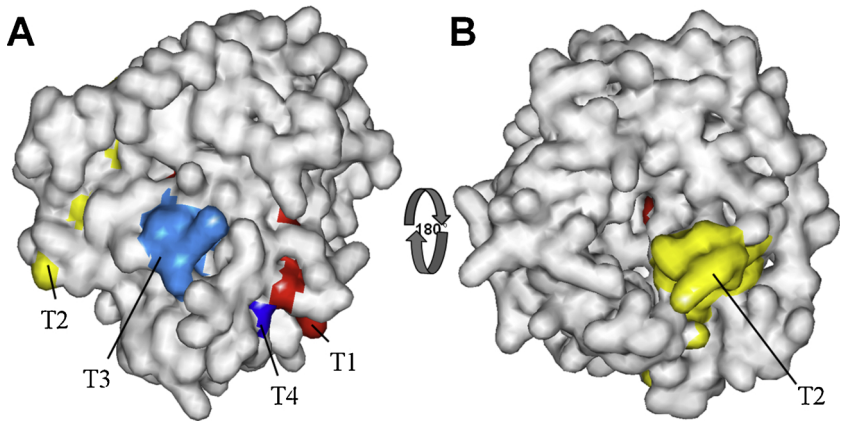

Fig. 7. T-cell epitopes in the tertiary structure of Der f 32.

T-cell epitope prediction yielded one peptide (47–55). As a result, Der f 32 was predicted to have four T-cell epitope sequences, including 47–55, 78–90, 127–135 and 143–151 (Table 4), and these peptides are also shown in Figs. 2 and 7.

4. Discussion

HDMs contribute to significant indoor allergens, that elicit allergic diseases such as asthma, rhinitis, and atopic dermatitis [42]. HDM allergies constitute more than 50% the allergies of allergic patients [12]. The relationship between house dust and HDMs in bronchial asthmatics has been demonstrated in many studies [43]. Although specific immunotherapy with HDM extracts is somewhat effective at present, allergen extracts have not been fully standardized and frequently cause severe adverse events in the course of treatment [44,45]. Immunologic characterization and epitope identification of single HDM allergens would be beneficial for the diagnosis and treatment of mite-induced atopic illnesses and for the design of effective epitope-based vaccines.

In the present study, Der f 32 was cloned, expressed and purified. Immunoblotting assays and ELISAs yielded 5 of 5 positive reactions to rDer f 32 using sera from dust-mite allergic patients. Park et al. recently reported that the rate of IgE-sensitisation to Der f 32 was 25% based on analysis of samples from 160 patients [46]. In order to better understand the structure and function of Der f 32, the basic sequence properties of Der f 32 were analyzed in the present study. The GRAVY of Der f 32 was -0.636, which indicated that Der f 32 exhibits hydrophilic characteristics. The instability index of Der f 32 was 27.21, which indicated that the sequence of Der f 32 was stable. Next, we successfully constructed the structure of Der f 32 through homology modeling. Homology modeling can construct a target structure on the basis of suitable templates extracted from homologous sequences [47]; this method has been successfully used for building several allergens structures, such as Der f 33 [18], Per a 9 [39], Ara h 2 [48], Ole e 2 [49] and Der f 5 [50]. In the present study, our Ramachandran plot showed that 90% of the residues present in the allowed region and that the model was reliable. A total of 98.8% residues of the Der f 32 model were in favored and allowed regions (Table 2), indicating that the distribution of the amino acids in the Der f 32 model were reasonable. The overall quality factor was 90.7% by the ERRAT program and 88.54% of the residues had an average 3D-1D score ≥ 0.2 by the VERIFY 3D program, indicating that the tertiary structures of Der f 32 were favorable and had high resolution. The Z-scores of ProSa indicated that there was high matching between the template and Der f 32 proteins. The Q-value (0.781) and Z-score (-7.38) of the QMEAN server results showed that the predicted model of Der f 32 was reliable and that the variation rate was low; both the overall folding and local structure had high accuracy, and the stereochemistry was reasonable (Table 2). The constructed model structure was feasible via the five-program analysis. All of the validations indicated that the constructed structural model of Der f 32 was available.

In silico prediction of B-cell epitopes are generally based on amino acid features of hydrophobicity, segmental flexibility, accessibility, polarity, antigenicity, exposed surfaces and turns. In the present study, BcePred, BCPreds, BPAP, and the DNAStar protean system were used to predict the B-cell epitopes of Der f 32. As a result, the present study ultimately predicted eight B-cell epitope peptides: 24–31, 111–121, 135–140, 168–172, 200–207, 214–220, 237–243 and 268–274. The B-cell epitopes were distributed completely on the surface of Der f 32 and were located largely in random coils of secondary structures (Figs. 2 and 6). Secondary protein structures are also important for predicting B-cell epitopes. Stable α-helices and β-sheets contain higher chemical-bond energies, which lead to increased difficulties in forming epitopes. B-cell epitopes are often located in β-turns and random coils of protein surface-exposed regions [51]. In addition, it has been reported that

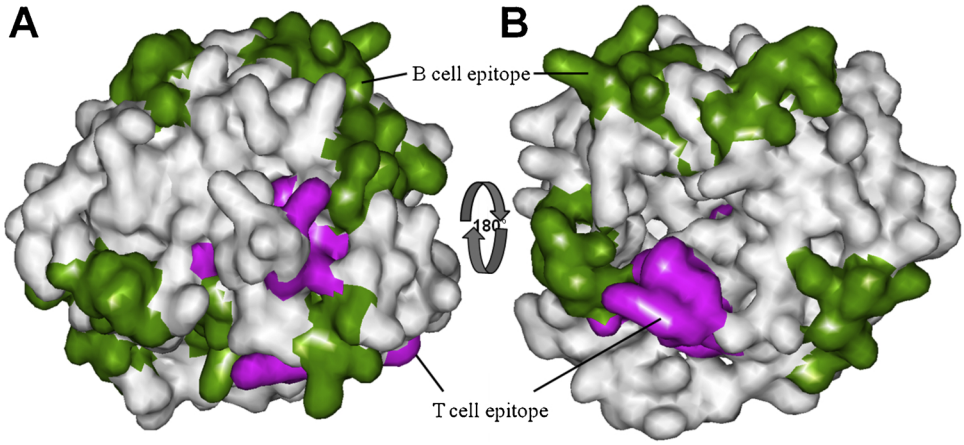

Fig. 8. Distribution of B-cell epitopes and T-cell epitopes in the tertiary structure of Der f 32.

allergen epitopes are comprised of a high proportion of hydrophobic amino acids [52]. The amino acids Lys and Asp play a key role in the IgE binding of allergenic epitopes [53]. In our present results, hydrophobic and charged amino acids comprised more than 80% of the residues of the predicted eight B-cell epitope peptides (Table 4). The common residues in all of the B-cell epitopes were Asp or Lys and each predicted that B-cell epitopes have one or more special Asp or Lys. The prediction results showed that four potential T-cell epitope sequences were predicted including 47–55, 78–90, 127–135, and 143–151. The T-cell epitopes were distributed largely in the interior of Der f 32 (Fig. 7). Four T-cell epitopes were located in antiparallel β-sheets, which suggests that they would not be easily degraded by proteases (Fig. 2). The epitope structure of Der f 32 had a typical structure of T-cell epitopes that were surrounded by B-cell epitopes (Fig. 8). The structure of T-cell epitopes being surrounded by B-cell epitopes may have biological significance for maintaining the immunogenicity of proteins. However, predicted results of each bioinformatic tools were different, indicating the limitations of *in silico* predictions approach. Thus, these predicted epitopes require experimental verification.

5. Conclusion

In this study, Der f 32 was cloned, expressed and purified. Immunoblotting analysis and ELISA yielded 5 of 5 positive reactions to rDer f 32 using sera from dust-mite allergic patients. The 3D structures of Der f 32 were theoretically constructed, among which eight B-cell epitopes and four T-cell epitopes were predicted. The epitope structure of Der f 32 consisted of a typical structure of T-cell epitopes being surrounded by B-cell epitopes, which is indicative of conferring strong immunogenicity. Taken together, these results may be beneficial for developing allergen immunotherapies.

Author contributions

Yuwei Li wrote the main manuscript and performed the experiments presented in Figs. 4–8 and analyzed the data. Yuwei Wang performed the experiments presented in Figs. 1–3 and analyzed the data. Zhigang Liu, Pingchang Yang and Pixin Ran designed the project, supervised the experiments and wrote the manuscript. All authors reviewed the manuscript.

Funding

International Science and Technology Cooperation Project of Shenzhen Science and Technology Plan (No. GJHZ2018041819053), Shenzhen Scientific Technology Basic Research Projects (No. JCYJ20180306171550045), Shenzhen Science and Technology Peacock Team Project (No. KQTD20170331145453160), State Key Laboratory of Respiratory Diseases Open Topic: Basic analysis of standardization and clinical diagnosis and treatment of new allergen components in dust mites (No. SKLRD-OP-201909). Overseas, Hong Kong and Macao Scholars Cooperative Research Continuation Project of National Natural Science Foundation (No. 31729002), National Natural Science Foundation of China (NSFC-Guangdong Joint Fund) (No. U1801286), Guangzhou Scientific Research Plan (Key Projects) (No. 201804020043).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R.A. Cooke, New etiologic factors in bronchial asthma, J. Immunol. 7 (147) (1922) 90.
[2] T. Miyamoto, S. Oshima, T. Ishizaki, S. Sato, Allergenic identity between the common floor mite (Dermatophagoides farinae Hughes, 1961) and house dust as a causative antigen in bronchial asthma, J. Allergy (Cairo) 42 (1) (1968) 14–28.
[3] R. Voorhorst, F.T.M. Spieksma, H. Varekamp, M. Leupen, A. Lyklema, The house-dust mite (Dermatophagoides pteronyssinus) and the allergens it produces. Identity with the house-dust allergen, J. Allergy (Cairo) 39 (6) (1967) 325–339.
[4] C. Pastor, J. Cuesta-Herranz, B. Cases, M. Pérez-Gordo, E. Figueredo, M. De Las Heras, F. Vivanco, Identification of major allergens in watermelon, Int. Arch. Allergy Immunol. 149 (4) (2009) 291–298.
[5] S. Vrtala, H. Huber, W.R. Thomas, Recombinant house dust mite allergens, Methods 66 (1) (2014) 67–74.
[6] F.M. Harold, Inorganic polyphosphates in biology: structure, metabolism, and function, Bacteriol. Rev. 30 (4) (1966) 772.
[7] I.R. Orriss, M.L. Key, M.O. Hajjawi, J.L. Millán, T.R. Arnett, Acidosis is a key regulator of osteoblast ecto-nucleotidase pyrophosphatase/phosphodiesterase 1 (NPP1) expression and activity, J. Cell. Physiol. 230 (12) (2015) 3049–3056.
[8] M.S. Alam, H. Moriyama, M. Matsumoto, Enzyme kinetics of dUTPase from the planarian Dugesia ryukyuensis, BMC Res. Notes 12 (1) (2019) 163.
[9] Y. Usui, T. Uematsu, T. Uchihashi, M. Takahashi, M. Ishizuka, R. Doto, H. Tanaka, Y. Komazaki, M. Osawa, Inorganic polyphosphate induces osteoblastic differentiation, J. Dent. Res. 89 (5) (2010) 504–509.
[10] G. Serrano-Bueno, J.M. Madroñal, J. Manzano-López, M. Muñiz, J.R. Pérez-Castiñeira, A. Hernández, A. Serrano, Nuclear proteasomal degradation of Saccharomyces cerevisiae inorganic pyrophosphatase Ipp1p, a nucleocytoplasmic protein whose stability depends on its subcellular localization, Biochim. Biophys. Acta, Mol. Cell Res. 1866 (6) (2019) 1019–1033.
[11] Y. Yang, R.-J. Tang, B. Mu, A. Ferjani, J. Shi, H. Zhang, F. Zhao, W.-Z. Lan, S. Luan, Vacuolar proton pyrophosphatase is required for high magnesium tolerance in Arabidopsis, Int. J. Mol. Sci. 19 (11) (2018) 3617.
[12] A.O. Eifan, M.A. Calderon, S.R. Durham, Allergen immunotherapy for house dust mite: clinical efficacy and immunological mechanisms in allergic rhinitis and asthma, Expert Opin. Biol. Ther. 13 (11) (2013) 1543–1556.
[13] F. Frati, S. Scurati, P. Puccinelli, J.L. Justicia, T. Adamec, H.J. Sieber, L. Ras, M. David, F. Marcucci, C. Incorvaia, Development of an allergen extract for sublingual immunotherapy–evaluation of Staloral, Expert Opin. Biol. Ther. 9 (9) (2009) 1207–1215.

[14] W.R. Thomas, The advent of recombinant allergens and allergen cloning, J. Allergy Clin. Immunol. 127 (4) (2011) 855–859.
[15] W.R. Reisacher, A. Wang, Novel strategies for allergy immunotherapy, Cur. Otorhinolaryngol. Rep. 1 (2) (2013) 119–126.
[16] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, Mol. Biol. Evol. 33 (7) (2016) 1870–1874.
[17] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (24) (1997) 4876–4882.
[18] F. Teng, J. Sun, L. Yu, Q. Li, Y. Cui, Homology modeling and epitope prediction of Der f 33, Braz. J. Med. Biol. Res. 51 (5) (2018).
[19] E. Gasteiger, C. Hoogland, A. Gattiker, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, the Proteomics Protocols Handbook, Springer, 2005, pp. 571–607.
[20] A. Bateman, L. Coin, R. Durbin, R.D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E.L. Sonnhammer, The Pfam protein families database, Nucleic Acids Res. 32 (suppl_1) (2004) D138–D141.
[21] C.J. Sigrist, E. De Castro, L. Cerutti, B.A. Cuche, N. Hulo, A. Bridge, L. Bougueleret, I. Xenarios, New and continuing developments at PROSITE, Nucleic Acids Res. 41 (D1) (2012) D344–D347.
[22] J. Gough, K. Karplus, R. Hughey, C. Chothia, Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure, J. Mol. Biol. 313 (4) (2001) 903–919.
[23] D.T. Jones, Protein secondary structure prediction based on position-specific scoring matrices, J. Mol. Biol. 292 (2) (1999) 195–202.
[24] M.S. Klausen, M.C. Jespersen, H. Nielsen, K.K. Jensen, V.I. Jurtz, C.K. Soenderby, M.O.A. Sommer, O. Winther, M. Nielsen, B. Petersen, NetSurfP-2.0: improved prediction of protein structural features by integrated deep learning, Proteins: Struct., Funct., Bioinf. (2019).
[25] F. Kiefer, K. Arnold, M. Künzli, L. Bordoli, T. Schwede, The SWISS-MODEL Repository and associated resources, Nucleic Acids Res. 37 (suppl_1) (2008) D387–D392.
[26] F. Melo, D. Devos, E. Depiereux, E. Feytmans, ANOLEA: a www server to assess protein structures, ISMB (1997) 187–190.
[27] N. Eswar, D. Eramian, B. Webb, M.-Y. Shen, A. Sali, Protein Structure Modeling With MODELLER, Structural Proteomics, Springer, 2008, pp. 145–159.
[28] H.A. Basta, J.-Y. Sgro, A.C. Palmenberg, Modeling of the human rhinovirus C capsid suggests a novel topography with insights on receptor preference and immunogenicity, Virology 448 (2014) 176–184.
[29] S. Ramachandran, P. Kota, F. Ding, N.V. Dokholyan, Automated minimization of steric clashes in protein structures, Proteins: Struct. Funct. Bioinf. 79 (1) (2011) 261–270.
[30] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (2) (1993) 283–291.
[31] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (9) (1993) 1511–1519.
[32] J.U. Bowie, R. Luthy, D. Eisenberg, A method to identify protein sequences that fold into a known three-dimensional structure, Science 253 (5016) (1991) 164–170.
[33] P. Benkert, S.C. Tosatto, D. Schomburg, QMEAN: a comprehensive scoring function for model quality assessment, Proteins: Struct. Funct. Bioinf. 71 (1) (2008) 261–277.
[34] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (suppl_2) (2007) W407–W410.
[35] L.-N. Zheng, H. Lin, R. Pawar, Z.-X. Li, M.-H. Li, Mapping IgE binding epitopes of major shrimp (Penaeus monodon) allergen with immunoinformatics tools, Food Chem. Toxicol. 49 (11) (2011) 2954–2960.
[36] S. Saha, G. Raghava, BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties, International Conference on Artificial Immune Systems, Springer, 2004, pp. 197–204.
[37] S. Saha, G. Raghava, Prediction of continuous B-cell epitopes in an antigen using recurrent neural network, Proteins: Struct. Funct. Bioinf. 65 (1) (2006) 40–48.
[38] X. Yang, X. Yu, An introduction to epitope prediction methods and software, Rev. Med. Virol. 19 (2) (2009) 77–96.
[39] H. Yang, H. Chen, M. Jin, H. Xie, S. He, J.-F. Wei, Molecular cloning, expression, IgE binding activities and in silico epitope prediction of per a 9 allergens of the American cockroach, Int. J. Mol. Med. 38 (6) (2016) 1795–1805.
[40] J.C. Jimenez-Lopez, S.O. Kotchoni, M.I. Rodríguez-García, J.D. Alché, Structure and functional features of olive pollen pectin methylesterase using homology modeling and molecular docking methods, J. Mol. Model. 18 (12) (2012) 4965–4984.
[41] D. Georgieva, K. Greunke, R.K. Arni, C. Betzel, Three-dimensional modelling of honeybee venom allergenic proteases: relation to allergenicity, Z. Fã¼r Naturforschung C 66 (5-6) (2011) 305–312.
[42] R. Sporik, S.T. Holgate, T.A. Platts-Mills, J.J. Cogswell, Exposure to house-dust mite allergen (Der p I) and the development of asthma in childhood: a prospective study, N. Engl. J. Med. 323 (8) (1990) 502–507.
[43] A. Bjerg, E. Rönmark, S. Hagstad, J. Eriksson, M. Andersson, G. Wennergren, K. Toren, L. Ekerljung, Gas, dust, and fumes exposure is associated with mite sensitization and with asthma in mite-sensitized adults, Allergy 70 (5) (2015) 604–607.
[44] W.R. Thomas, B.J. Hales, W.-A. Smith, House dust mite allergens in asthma and allergy, Trends Mol. Med. 16 (7) (2010) 321–328.
[45] H. Moed, R. Gerth van Wijk, R.W. Hendriks, J.C. van der Wouden, Evaluation of clinical and immunological responses: a 2-year follow-up study in children with allergic rhinitis due to house dust mite, Med. Inflamm. 2013 (2013).
[46] K. Park, J. Lee, J.Y. Lee, S. Lee, D. Sim, J. Shin, C. Park, J.H. Lee, K. Lee, K. Jeong, Sensitization to various minor house dust mite allergens is greater in patients with atopic dermatitis than in those with respiratory allergic disease, Clin. Exp. Allergy 48 (8) (2018) 1050–1058.
[47] A. Wong, C. Gehring, H.R. Irving, Conserved functional motifs and homology modeling to predict hidden moonlighting functional sites, Front. Bioeng. Biotechnol. 3 (2015) 82.
[48] A. Barre, J.-P. Borges, R. Culerrier, P. Rougé, Homology modelling of the major peanut allergen Ara h 2 and surface mapping of IgE-binding epitopes, Immunol. Lett. 100 (2) (2005) 153–158.
[49] J. Jimenez-Lopez, M. Rodriguez-Garcia, D. Alche, Olive tree genetic background is a major cause of profilin (Ole e 2 allergen) polymorphism and functional and allergenic variability, Commun. Agric. Appl. Biol. Sci. 78 (1) (2013) 213–219.
[50] S. Khemili, J.M. Kwasigroch, T. Hamadouche, D. Gilis, Modelling and bioinformatics analysis of the dimeric structure of house dust mite allergens from families 5 and 21: der f 5 could dimerize as Der p 5, J. Biomol. Struct. Dyn. 29 (4) (2012) 663–675.
[51] K. Sikic, S. Tomic, O. Carugo, Systematic comparison of crystal and NMR protein structures deposited in the protein data bank, Open Biochem. J. 4 (2010) 83.
[52] N. Wolff, S. Yannai, N. Karin, Y. Levy, R. Reifen, I. Dalal, U. Cogan, Identification and characterization of linear B-cell epitopes of β-globulin, a major allergen of sesame seeds, J. Allergy Clin. Immunol. 114 (5) (2004) 1151–1158.
[53] N. Oezguen, B. Zhou, S.S. Negi, O. Ivanciuc, C.H. Schein, G. Labesse, W. Braun, Comprehensive 3D-modeling of allergenic proteins and amino acid composition of potential conformational IgE epitopes, Mol. Immunol. 45 (14) (2008) 3740–3747.