Homology modeling and prediction of B-cell and T-cell epitopes of the house dust mite allergen Der f 20

FEIXIANG TENG[1], LILI YU[1], JINXIA SUN[1], NAN WANG[1] and YUBAO CUI[2]

[1]Department of Basic Medicine, Jiangsu Vocational College of Medicine, Yancheng, Jiangsu 224005; [2]Department of Clinical Laboratory, Wuxi People's Hospital Affiliated to Nanjing Medical University, Wuxi, Jiangsu 214023, P.R. China

Received May 23, 2016; Accepted April 6, 2017

DOI: 10.3892/mmr.2017.8066

Abstract. House dust mite allergens can cause allergic diseases, including asthma, atopic dermatitis and rhinitis. Der f 20 is a novel allergen of *Dermatophagoides farina* (Der f), which is an arginine kinase. In the present study, the B-cell and T-cell epitopes of Der f 20 were predicted. The protein attribution, patterns, physicochemical properties and secondary structure of Der f 20 were also predicted. Der f 20 is a member of the ATP:guanido phosphotransferase family and contains a phosphagen kinase pattern. Using homology modeling, the present study constructed a reasonable tertiary structure of Der f 20. Using BcePred, ABCpred, BCPred and BPAP systems, B-cell epitopes at 20-25, 41-49, 111-118, 131-141, 170-174 and 312-321 were predicted. Using NetMHCIIpan-3.0 and NetMHCII-2.2, T-cell epitopes were predicted at 194-202, 239-247 and 274-282. These results provide a theoretical basis for the design off Der f 20 epitope-based vaccines.

Introduction

Allergic diseases, including bronchial asthma, atopic dermatitis and rhinitis, affect 30-40% of the global population (1,2). Allergens from house dust mites (HDMs); in particular those from the most common HDMs, *Dermatophagoides farinae* (Der f) and *Dermatophagoides pteronyssinus* (Der p), are major environmental factors for allergic diseases (3-5). At least 34 groups of HDM allergens have been identified and listed in the Allergen Nomenclature database (http://www.allergen.org). Der f 20, identified and denominated from *D. farinae*, belongs to the group 20 allergens. Der f 20 is a 40-kDa arginine kinase, however, its physiological function remains to be fully elucidated.

Allergen extracts of various mite species, including mite bodies, eggs and culture media, have been used to diagnose and treat IgE-mediated allergic diseases. Certain patients may be sensitized to one or two mite allergens, whereas others respond to a spectrum of allergens (6-10). However, these extracts have limitations in terms of safety and validity in allergen-specific immunotherapy (SIT) (6-10). SIT is the only etiological therapy, which suppresses allergic responses in rhinitis and asthma (8,11). By contrast, pure and standardized recombinant allergens, containing the majority of the IgE-binding epitopes of an allergen source, can be used to replace natural extracts, offering a safer and more valid approach to SIT (9,10).

Several SIT-based studies have focused on using recombinant allergens to develop epitope-based vaccines (12,13). These vaccines contain multiple B-cell and/or T-cell linear antigen epitopes and can thus overcome virulence return or spread, and induce more efficient presentation when detected and combined by host major histocompatibility complex (MHC) molecules (14,15). These findings suggest that B-cell and T-cell epitopes from one major component of an allergen may be necessary for immunotherapy of allergic diseases. Therefore, the identification of exact epitopes of HDM allergens can benefit the preparation of epitope-based vaccines and treatment of allergic diseases.

Previous studies have identified several HDM allergen epitopes (16), although no Der f 20 epitopes have been reported. Therefore, the present study used bioinformatics approaches to identify B-cell and T-cell epitopes of Der f 20.

Materials and methods

Sequence retrieval and analyses. The amino acid sequence of Der f 20 (accession no. AIO08850.1) was obtained from the International Union of Immunological Societies nomenclature database and the protein database of the National Center for Biotechnology Information (www.allergen.org). The family classification of Der f 20 was analyzed using Pfam v29.0

Correspondence to: Dr Yubao Cui, Department of Clinical Laboratory, Wuxi People's Hospital Affiliated to Nanjing Medical University, 299 Qingyang Road, Wuxi, Jiangsu 214023, P.R. China
E-mail: ybcui1975@hotmail.com

Abbreviations: HDM, house dust mite; PDB, Protein Data Bank; 3D structure, tertiary structure

Key words: Der f 20, homology modeling, B-cell epitopes, T-cell epitopes

(pfam.xfam.org) (17), Superfamily v1.75 (supfam.cs.bris.ac.uk/SUPERFAMILY/hmm.html) (18) and InterPro v56.0 (www.ebi.ac.uk/interpro/) (19). The TMHMM server 2.0 (www.cbs.dtu.dk/services/TMHMM/) was used to predict transmembrane protein helices (20).

Physiochemical and patterns analyses. Physiochemical analyses, including molecular weight, negatively charged residues, positively charged residues, theoretical pI, aliphatic index, grand average of hydropathicity (GRAVY) and instability index of Der f 20, were predicted using ProtParam (web.expasy.org/protparam/) (21). The characteristic patterns, functional motifs and active sites of Der f 20 were assessed using Prosite (prosite.expasy.org/) (22).

Structure prediction and homology modeling. The TMHMM server 2.0 (www.cbs.dtu.dk/services/TMHMM/) was used to predict transmembrane protein helices (20). The PredictProtein server (www.predictprotein.org/) was used to predict the secondary structure of Der f 20 (23). Homology modeling was used to construct a tertiary structure of Der f 20. A BLASTP (blast.ncbi.nlm.nih.gov/Blast.cgi) search with default parameters was performed against the Protein Data Bank (PDB) (www.rcsb.org/pdb/) to identify suitable templates of Der f 20. The appropriate templates were selected based on the high score, low e-value, and maximum sequence identity. MODELLER v9.16 (salilab.org/modeller/) (24) was used to predict the tertiary structure of Der f 20. The predicted structure was imported into Chiron (redshift.med.unc.edu/chiron/login.php) (25) to rectify unfavorable clashes and improve stereochemistry quality.

Estimating the quality of the structural models is a vital step in protein structure construction. PROCHECK (services.mbi.ucla.edu/SAVES) (26) was used to verify the stereochemical quality of the structure of Der f 20. ERRAT (services.mbi.ucla.edu/SAVES) (27) was used to analyze the statistics of non-bonded interactions between different atom types. VERIFY_3D (services.mbi.ucla.edu/SAVES) (28) was used to determine the compatibility of an atomic model (3D) with its the amino acid sequence (1D) and to compare the results with favorable structures. ProSA (prosa.services.came.sbg.ac.at/prosa.php) (29) was used to analyze the Z-score to determine the degree of match between the template protein and Der f 20. QMEAN (swissmodel.expasy.org/qmean) (30) is a composite scoring function, which provides the global (for the entire structure) and local (per residue) error estimates on the basis of a single model. Superimposition of the query and template structure, and visualization of the generated models was performed using UCSF Chimera 1.10.2 (www.cgl.ucsf.edu/chimera/) (31).

Prediction of B-cell epitopes. BcePred (http://crdd.osdd.net/raghava/bcepred) (32), ABCpred (crdd.osdd.net/raghava/abcpred) (33), BCPreds (ailab.ist.psu.edu/bcpred) (34) and the Bioinformatics Predicted Antigenic Peptides (BPAP) system (imed.med.ucm.es/Tools/antigenic.pl) (35) were used to predict the B-cell antigenic epitopes of Der f 20. BcePred predicts B-cell epitopes using physicochemical properties, including hydrophilicity, flexibility/mobility, accessibility, polarity, exposed surfaceand turns, or a combination of properties. ABCpred predicts B-cell epitopes in antigen sequences using artificial neural networks. BCPREDS selects three prediction methods of the AAP method (35), BCPred (36) and FBCPred (37) predict B-cell antigenic epitopes. The BPAP system combines the physicochemical properties of amino acids to predict epitopes.

Prediction of T-cell epitopes. The T-cell epitopes were predicted by identifying peptide binding to MHC molecules. The binding significance of each peptide to the given MHC molecule was based on the estimated strength of binding exhibited by a predicted nested core peptide at a set threshold level. NetMHCII 2.2 (www.cbs.dtu.dk/services/NetMHCII) (38) predicted the binding of epitope peptides to HLA-DQ alleles using artificial neuron networks. NetMHCIIpan-3.1(www.cbs.dtu.dk/services/NetMHCIIpan) (39) was used for HLA-DR-based epitope prediction. In these two software programs, peptides with a high binding ability had a half maximal inhibitory concentration (IC50) value <50 nM and weak binding peptides had an IC50 value <500 nM. The ultimate T-cell epitopes were obtained by combining the results of the HLA-DR and HLA-DQ allele epitopes.

Results

Amino acid sequence analysis. Family classification showed that Der f 20 belongs to the ATP: guanido phosphotransferase family (InterPro no. IPR000749) and arginine kinase superfamily (InterPro no. IPR023660). Prosite was used to analyze characteristic motifs or patterns and revealed that Der f 20 contained a PHOSPHAGEN_KINASE pattern (PS00112; 271-277; **CPTNLGT**) and active site at residue 271 (Fig. 1). The phosphorylation sites of Der f 20 included Ser residues 20, 260 and 282; Thr residues 44, 49, 177, 269, 278, 311 and 334; and Tyr residues 75 and 145. DNA-PK kinase (T 177; LLGMDKA**T**QQQLIDD) was predicted as phosphorylated for Der f 20.

Der f 20 includes 356 amino acids and has a molecular weight of 40177.8 Da. The protein contains 49 negatively charged residues (Asp and Glu) and 46 positively charged residues (Arg and Lys). Der f 20 had a theoretical pI of 6.24 and an aliphatic index of 95.06. GRAVY was -0.103, indicating that Der f 20 exhibits a hydrophilic characteristic. The instability index was 30.57, indicating that the amino acid sequence of Der f 20 was stable.

Structural analysis and homology modeling. The Der f 20 protein sequences were entered into the TMHMM Server 2.0 to predict transmembrane helices. The computed results showed that Der f 20 had no transmembrane helices, and all protein sequences were located outside of the membrane (Fig. 1). The percentages of overall amino acids located in α-helices, β-sheets and random coils were 41.57% (12 domains), 15.73% (10 domains) and 42.70%, respectively (Table I).

The tertiary structure of *Litopenaeus vannamei* arginine kinase (PDB accession no. 4BG4) had a high sequence identity (78%) with Der f 20 and was therefore used as the template for homology modeling. Following homology modeling,

Table I. Secondary and tertiary structural elements of Der f 20.

Structure	α-helices (%)	β-sheets (%)	Random coils (%)
Secondary	41.57 (12 domains)	15.73 (10 domains)	42.70
Tertiary	41.01 (14 domains)	13.48 (10 domains)	45.51

Table II. Evaluation parameters for the tertiary structure of Der f 20.

Protein	Structural assessment method	Ramachandran plot %	ERRAT value	VERIFY 3D (%)	Z-score	Q-value
Der f 20	PROCHECK	89.4 core 9.9 allowed 0.6 generously allowed 0.0 disallowed	96.264	96.63		
	ProSa				-10.17	
	QMEAN				-0.12	0.753
4BG4	PROCHECK	92.0 core 7.6 allowed 0.5 generously allowed 0.0 disallowed	98.127	95.21		
	ProSa				-10.53	
	QMEAN				-0.14	0.748

Figure 1. Secondary structure elements for Der f 20. α-helices underlined; β-sheets are shaded grey; random coils are unlabeled; epitopes are shown in boxes; and the black triangle indicates the active site.

a Ramachandran plot showed that 89.4% of the amino acid residues within the tertiary structure of Der f 20 were within the most favored regions; 9.9% of residues were in additional allowed regions; 0.6% of residues were in generously allowed regions; and 0% of residues were in disallowed regions. The ERRAT results showed that the overall quality factor was 96.264, indicating that the tertiary structure of Der f 20 had high resolution. The VERIFY 3D results showed that 96.63% of residues had an average 3D-1D score ≥0.2, indicating that the structures were favorable. As indicated by ProSa, the Z-scores of template and Der f 20 showed high matching between the tertiary structures of the protein. The QMEAN server results showed that the QMEAN Z-score was -0.82 and the standard deviation value was <1, indicating that the protein model variation rate was low, overall folding and local structures had high accuracy rates and stereochemistry was reasonable. In addition, the Q value was 0.753, indicating that the predicted model of Der f 20 was reliable. Therefore, based on these results, the tertiary structure model of Der f 20 was reliable and suitable for use in the present study (Table II; Fig. 2A).

The tertiary structure of Der f 20 was also found to contain α-helices, β-sheets and random coils, although the amino acid numbers of these three elements were marginally different, compared with the secondary structure. The percentages of overall amino acids located in α-helices, β-sheets and random coils were 41.01% (14 domains), 13.48% (10 domains) and 45.51%, respectively (Table I).

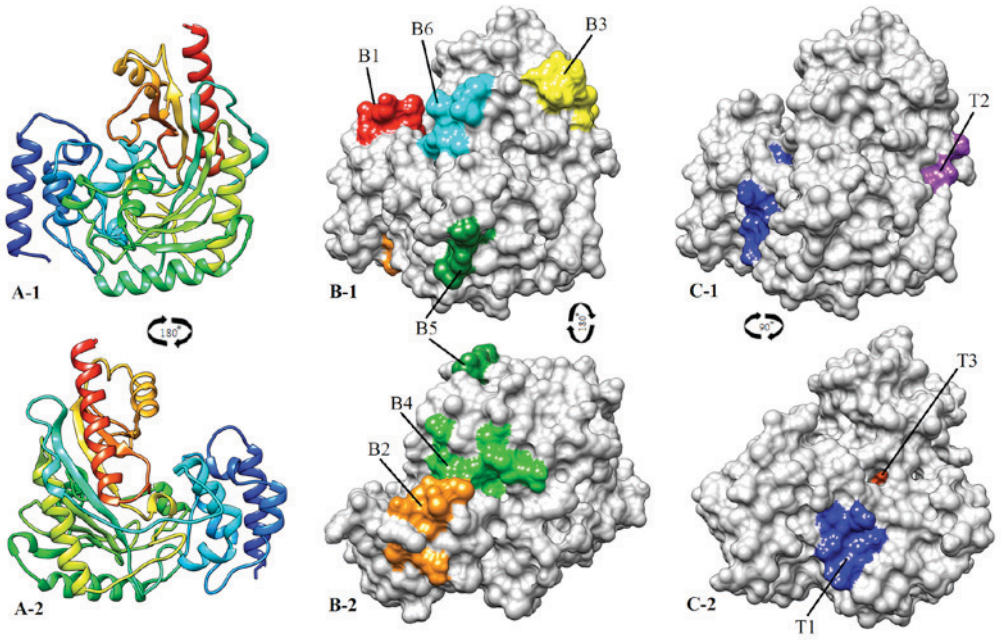

Figure 2. B-cell and T-cell epitopes on the 3D structure of Der f 20. (A-1 and A-2) 3D structure of Der f 20. (B-1 and B-2) B-cell epitopes on the 3D structure of Der f 20. (C-1 and C-2) T-cell epitopes on the 3D structure of Der f 20.

B-cell epitope prediction. Hydrophobicity, fragment flexibility/mobility, surface accessibility, polarity, exposed surface and turns are important features for B-cell antigenic epitope identification. These antigenic indices were used to determine the epitope forming capacity of the Der f 20 amino acid sequence. Based on antigenic indices, BcePred, ABCpred, BCPred and BPAP were used in the present study to predict B-cell epitopes. Ultimately, seven antigenic epitope peptides were predicted, including 20-25, 43-49, 110-118, 131-142, 170-174, 203-210 and 311-321 (Tables III and IV; Fig. 2B).

T-cell epitope prediction. NetMHCIIpan 3.1 was used to predict T-cell epitopes in the regions of HLA-DR DRB101, HLA-DRB301, HLA-DRB40 and HLA-DRB501. NetMHCII 2.2 was used to predict T-cell epitopes in the regions of HLA-DQA10101-DQB10501, HLA-DQA10102-DQB10602, HLA-DQA10301-DQB10302, HLA-DQA10401-DQB10402, HLA-DQA10501-DQB10201 and HLA-DQA10501-DQB10301. Combined with the software results, two T-cell epitope peptides were ultimately predicted, including 194-202 and 274-282 (Tables III and IV; Fig. 2C).

Discussion

Type I allergic diseases, including rhinitis, asthma and atopic dermatitis, are increasing worldwide. HDM antigens are responsible for the sensitization of >50% of patients with airway allergic disease (7,40). Therefore, the prediction and characterization of specific B-cell and T-cell epitopes of HDM allergens, including Der f 20, can assist in mechanistic investigations of immune responses and the design of epitope-based vaccines.

Der f 20 is a member of the ATP:guanido phosphotransferase family and arginine kinase superfamily, and the protein contains a phosphagen kinase motif pattern and an active site at residue 271. Der f 20 is a hydrophilic and stable protein with no transmembrane helices and all protein sequences located outside of the membrane. Homology modeling, or comparative protein modeling, construct a Der f 20 structure based on comparisons with data extracted from homologous sequences with known structures (parents or templates) (41).

The quality of a homology model is dependent on high quality sequence alignment and template structure. Therefore, the present study used the crystal structure of 4BG4 as a template, as it has 78% sequence identity with Der f 20. Following homology modeling with MODELLER, various additional parameters/programs were incorporated to establish a reliable model of Der f 20. Although Der f 20 contains α-helices, β-sheets and random coils, the numbers of amino acids in these three elements in the tertiary structure were found to vary marginally from the secondary structure. This discrepancy may be due to different structural prediction methods.

Epitopes or antigenic determinants, which represent the immune-active regions of antigen molecules, are the regions of an antigen, which are recognized by the immune system, specifically by antibodies and lymphocyte (B-cell or T-cell) surface antigen receptors. The properties of the antigen epitope, their number and their spatial configuration determine antigen specificity (42,43). Epitopes usually contain 6-8 amino acids residues and in general contain <20 amino acid residues.

In the present study, BcePred, ABCpred, BCPreds and BPAP were used to predict the B-cell epitopes of Der f 20. Secondary and tertiary protein structures also contain important information for B-cell epitope prediction. For example, α-helices and β-sheets have higher chemical bond energy and have difficulty forming epitope sequences. By contrast, β-turns and random coils are located in surface-exposed regions of a protein, which often contain

Table III. B-cell and T-cell epitope prediction.

Epitope prediction	Prediction server	Prediction result
B-cell epitopes	BcePred	12-18, 36-45, 143-152, 229-235, 256-263, 290-297, 321-330
	ABCpred	3-19, 24-40, 37-53, 61-77, 69-85, 83-99, 94-110, 127-143, 158-174, 164-180, 179-195, 185-201, 203-219, 228-244, 252-268, 309-325, 320-336, 329-345
	BCPred	99-119, 131-151, 311-331, 207-227
	BPAP system	11-29, 31-38, 47-56, 73-87, 106-113, 117-142, 164-172, 177-187, 193-210, 238-249, 264-273, 279-290, 296-308, 317-324, 345-351
T-cell epitopes (HLA-DR)	NetMHCIIpan-3.0 (DRB1*01:01)	14-22, 167-175, 194-202, 239-247, 243-251, 267-275
	NetMHCIIpan-3.0 (DRB3*01:01)	
	NetMHCIIpan-3.0 (DRB4*01:01)	
	NetMHCIIpan-3.0 (DRB5*01:01)	194-202, 279-287
T-cell epitopes (HLA-DQ)	NetMHCII-2.2 (HLA-DQA10101-DQB10501)	
	NetMHCII-2.2 (HLA-DQA10102-DQB10602)	228-236, 274-282
	NetMHCII-2.2 (HLA-DQA10301-DQB10302)	
	NetMHCII-2.2 (HLA-DQA10401-DQB10402)	217-225
	NetMHCII-2.2 (HLA-DQA10501-DQB10201)	
	NetMHCII-2.2 (HLA-DQA10501-DQB10301)	45-53, 66-74

BPAP, Bioinformatics Predicted Antigenic Peptides.

Table IV. Predicted B-cell and T-cell epitopes of *Dermatophagoides farina* 20.

Peptide	Type of epitope	Position	Sequence[a]
P1	B	20-25	SAECHS
P2	B	41-49	GRKTGMGAT
P3	B	111-118	CNVDPNNE
P4	B	131-141	LQGYPFNPCLT
P5	B	170-174	LLGMD
P6	B	312-321	AGEHTESVGG
P7	T	194-202	FLQAANACR
P8	T	239-247	LKQVFSRLI
P9	T	274-282	NLGTTIRAS

[a]Regions highlighted in black with white text represent hydrophobic amino acid residues.

epitope sequences (44). Integrating the shared results of the four servers, and combining information from secondary and tertiary structures, the present study ultimately predicted six B-cell epitope peptides: 20-25, 41-49, 111-118, 131-141, 170-174 and 312-321. A total of three T-cell epitope peptides were predicted: 194-202, 239-247 and 274-282. In addition, allergen epitopes usually contain high proportions of hydrophobic amino acid residues, including Ala, Ser, Asn, Gly and Lys (45). The predictions showed that the majority of the B-cell and T-cell epitopes identified in the present study contained multiple hydrophobic amino acids. However, these predicted epitopes require further experimental verification.

In conclusion, the present study constructed a reasonable tertiary structure of Der f 20. Using bioinformatics, the B-cell epitopes (20-25, 41-49, 111-118, 131-141, 170-174 and 312-321) and T-cell epitopes (194-202, 239-247 and 274-282) were predicted based on the secondary and tertiary structures of Der f 20. These results represent a significant step towards the design of Der f 20 epitope-based vaccines for allergic diseases.

References

1. Arlian LG: House-dust-mite allergens: A review. Exp Appl Acarol 10: 167-186, 1991.
2. Platts-Mills TAE, Thomas WR, Aalberse R, Vervloet D and Champman MD: Dust mite allergens and asthma: Report of a second international workshop. J Allergy Clin Immunol 89: 1046-1060, 1992.
3. Thomas WR, Hales BJ and Smith WA: House dust mite allergens in asthma and allergy. Trends Mol Med 16: 321-328, 2010.
4. Nadchatram M: House dust mites, our intimate associates. Trop Biomed 22: 23-37, 2005.
5. Tovey ER, Chapman MD and Platts-Mills TA: Mite faeces are a major source of house dust mite allergens. Nature 289: 592-593, 1981.
6. Marth K, Focke-Tejkl M, Lupinek C, Valenta R and Niederberger V: Allergen peptides, recombinant allergens and hypoallergens for allergen-specific immunotherapy. Curr Treat Options Allergy 1: 91-106, 2014.
7. Vrtala S, Huber H and Thomas WR: Recombinant house dust mite allergens. Methods 66: 67-74, 2014.
8. Jutel M, Solarewicz-Madejek K and Smolinska S: Recombinant allergens: The present and the future. Hum Vaccin Immunother 8: 1534-1543, 2012.
9. Valenta R, Niespodziana K, Focke-Tejkl M, Marth K, Huber H, Neubauer A and Niederberger V: Recombinant allergens: What does the future hold? J Allergy Clin Immunol 127: 860-864, 2011.
10. Focke-Tejkl M and Valenta R: Safety of engineered allergen-specific immunotherapy vaccines. Curr Opin Allergy Clin Immunol 12: 555-563, 2012.
11. Lee J, Park CO and Lee KH: Specific immunotherapy in atopic dermatitis. Allergy Asthma Immunol Res 3: 221-229, 2015.
12. Zhao J, Li C, Zhao B, Xu P, Xu H and He L: Construction of the recombinant vaccine based on T-cell epitope encoding Der p1 and evaluation on its specific immunotherapy efficacy. Int J Clin Exp Med 4: 6436-6443, 2015.
13. Koffeman EC, Genovese M, Amox D, Keogh E, Santana E, Matteson EL, Kavanaugh A, Molitor JA, Schiff MH, Posever JO, *et al*: Epitope-specific immunotherapy of rheumatoid arthritis: Clinical responsiveness occurs with immune deviation and relies on the expression of a cluster of molecules associated with T-cell tolerance in a double-blind, placebo-controlled, pilot phase II trial. Arthritis Rheum 11: 3207-3216, 2009.
14. Sharmin R and Islam AB: A highly conserved WDYPKCDRA epitope in the RNA directed RNA polymerase of human coronaviruses can be used as epitope-based universal vaccine design. BMC Bioinformatics 1: 161, 2014.
15. Alexander C, Kay AB and Larche M: Peptide-based vaccines in the treatment of specific allergy. Curr Drug Targets Inflamm Allergy 4: 353-361, 2002.
16. Cui Y: Immunoglobulin e-binding epitopes of mite allergens: From characterization to immunotherapy. Clin Rev Allergy Immunol 3: 344-353, 2014.
17. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, *et al*: The Pfam protein families database: Towards a more sustainable future. Nucleic Acids Res 44: D279-D285, 2016.
18. Gough J, Karplus K, Hughey R and Chothia C: Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. J Mol Biol 313: 903-919, 2001.
19. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, *et al*: The InterPro protein families database: The classification resource after 15 years. Nucleic Acids Res 43 (Database Issue): D213-D221, 2015.
20. Krogh A, Larsson B, von Heijne G and Sonnhammer EL: Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol 305: 567-580, 2001.
21. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD and Hochstrasser DF: Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 112: 531-552, 1999.
22. De Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A and Hulo N: ScanProsite: Detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34 (Web Server issue): W362-W365, 2006.
23. Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, Hönigschmid P, Schafferhans A, Roos M, *et al*: PredictProtein-an open resource for online prediction of protein structural and functional features. Nucleic Acids Res 42 (Web Server issue): W337-W343, 2014.
24. Webb B and Sali A: Protein structure modeling with MODELLER. Methods Mol Biol 1137: 1-15, 2014.
25. Ramachandran S, Kota P, Ding F and Dokholyan NV: Automated minimization of steric clashes in protein structures. Proteins 1: 261-270, 2011.
26. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R and Thornton JM: AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. J Biomol NMR 4: 477-486, 1996.
27. Colovos C and Yeates TO: Verification of protein structures: Patterns of nonbonded atomic interactions. Protein Sci 9: 1511-1519, 1993.
28. Bowie JU, Lüthy R, Eisenberg D: A method to identify protein sequences that fold into a known three-dimensional structure. Science 253: 164-170, 1991.
29. Wiederstein M and Sippl MJ: ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35 (Web Server issue): W407-W410, 2007.
30. Benkert P, Tosatto SC and Schomburg D: QMEAN: A comprehensive scoring function for model quality assessment. Proteins 1: 261-277, 2008.
31. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC and Ferrin TE: UCSF Chimera-a visualization system for exploratory research and analysis. J Comput Chem 13: 1605-1612, 2004.
32. Saha S and Raghava GP: BcePred: Prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. Artificial Immune Systems 3239: 197-204, 2004.
33. Saha S and Raghava GP: Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 65: 40-48, 2006.
34. Chen J, Liu H, Yang J and Chou KC: Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. Amino Acids 3: 423-428, 2007.
35. Zheng LN, Lin H, Pawar R, Li ZX and Li MH: Mapping IgE binding epitopes of major shrimp (Penaeus monodon) allergen with immunoinformatics tools. Food Chem Toxicol 49: 2954-2960, 2011.
36. EI-Manzalawy Y, Dobbs D and Honavar V: Predicting linear B-cell epitopes using string kernels. J Mol Recognit 21: 243-255, 2008.
37. EI-Manzalawy Y, Dobbs D and Honavar V: Predicting flexible length linear B-cell epitopes. Comput Syst Bioinformatics Conf 7: 121-132, 2008.
38. Nielsen M and Lund O: NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinformatics 10: 296, 2009.
39. Andreatta M, Karosiene E, Rasmussen M, Stryhn A, Buus S and Nielsen M: Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. Immunogenetics 67: 641-650, 2015.
40. An S, Shen C, Liu X, Chen L, Xu X, Rong M, Liu Z and Lai R: Alpha-actinin is a new type of house dust mite allergen. Plos One 8: e81377, 2013.
41. Wong A, Gehring C and Irving HR: Conserved functional motifs and homology modeling to predict hidden moonlighting functional sites. Front Bioeng Biotechnol 3: 82, 2015.
42. Brusic V, Bajic VB and Petrovsky N: Computational methods for prediction of T-cell epitopes-a framework for modelling, testing, and applications. Methods 34: 436-443, 2004.
43. Zhao H, Verma D, Li W, Choi Y, Ndong C, Fiering SN, Bailey-Kellogg C and Griswold KE: Depletion of T-cell epitopes in lysostaphin mitigates anti-drug antibody response and enhances antibacterial efficacy in vivo. Chem Biol 5: 629-639, 2015.
44. Sikic K, Tomic S and Carugo O: Systematic comparison of crystal and NMR protein Structures deposited in the protein data bank. Open Biochem J 4: 83-95, 2010.
45. Oezguen N, Zhou B, Negi SS, Ivanciuc O, Schein CH, Labesse G and Braun W: Comprehensive 3D-modeling of allergenic proteins and amino acid composition of potential conformational IgE epitopes. Mol Immunol 45: 3740-3747, 2008.