Original Paper

In Silico Identification and *in Vitro* Analysis of B and T-Cell Epitopes of the Black Turtle Bean (*Phaseolus Vulgaris* L.) Lectin

Shudong He^{a,b,c} Jinlong Zhao^a Walid Elfalleh^d Mohamed Jemaà^{e,g} Hanju Sun^a
Xianbao Sun^a Mingming Tang^a Qian He^a Zeyu Wu^a Florian Lang^{f,g}

^aSchool of Food Science and Engineering, Hefei University of Technology, Hefei, ^bSchool of Food Science and Technology, National Engineering Research Center of Seafood, Dalian Polytechnic University, Dalian, ^cBeijing Advanced Innovation Center for Food Nutrition and Human Health, Beijing Technology and Business University (BTBU), Beijing, China, ^dUR Catalyse et Matériaux pour l'Environnement et les Procédés URCMEP (UR11ES85), Faculté des Sciences de Gabès, Université de Gabès, Gabès, Tunisia, ^eDepartment of Laboratory Medicine, Translational Cancer Research, Lund University, Lund, Sweden, ^fDepartment of Molecular Medicine II, Medical Faculty, Heinrich Heine University, Duesseldorf, ^gDepartment of Physiology I, Tübingen University, Tübingen, Germany

Key Words

Lectin • Black turtle bean • *In silico* identification • Epitopes • Peptide synthesis • QSAR

Abstract

Background/Aims: The incidence of lectin allergic disease is increasing in recent decades, and definitive treatment is still lacking. Identification of B and T-cell epitopes of allergen will be useful in understanding the allergen antibody responses as well as aiding in the development of new diagnostics and therapy regimens for lectin poisoning. In the current study, we mainly addressed these questions. **Methods:** Three-dimensional structure of the lectin from black turtle bean (*Phaseolus vulgaris* L.) was modeled using the structural template of Phytohemagglutinin from *P. vulgaris* (PHA-E, PDB ID: 3wcs.1.A) with high identity. The B and T-cell epitopes were screened and identified by immunoinformatics and subsequently validated by ELISA, lymphocyte proliferation and cytokine profile analyses. **Results:** Seven potential B-cell epitopes (B1 to B7) were identified by sequence and structure based methods, while three T-cell epitopes (T1 to T3) were identified by the predictions of binding score and inhibitory concentration. The epitope peptides were synthesized. Significant IgE binding capability was found in B-cell epitopes (B2, B5, B6 and B7) and T2 (a cryptic B-cell epitope). T1 and T2 induced significant lymphoproliferation, and the release of IL-4 and IL-5 cytokine confirmed the validity of T-cell epitope prediction. Abundant hydrophobic amino acids were found in B-cell epitope and T-cell epitope regions by amino acid analysis. Positively charged

S. He and J. Zhao contributed equally to this work.

Hanju Sun School of Food Science and Engineering, Hefei University of Technology, Hefei 230009 (China); Department of
and Florian Lang Molecular Medicine II, Medical Faculty, Heinrich Heine University, Duesseldorf (Germany); Department of
Physiology I, Tübingen University, Tübingen (Germany); E-Mail sunhanjv@yeah.net; florian.lang@uni-tuebingen.de

amino acids, such as His residue, might be more favored for B-cell epitope. **Conclusion:** The present approach can be applied for the identification of epitopes in novel allergen proteins and thus for designing diagnostics and therapies in lectin allergy.

© 2018 The Author(s)
Published by S. Karger AG, Basel

Introduction

Black turtle bean (*Phaseolus vulgaris* L.), also known as black soup bean and black kidney bean, is a common legume consumed worldwide [1, 2]. However, due to its rich content of lectin, the black turtle bean is held to be responsible for several food associated allergic reactions such as nausea, diarrhea, vomiting and abdominal swelling in both children and adults [3]. In recent decades, a series of damage mass food poisoning suffered from raw or under-cooked *P. vulgaris* beans have been reported around the world [4]. During 1976 and 1980, seven incidents involving forty-three persons on *P. vulgaris* beans poisoning were reported by Noah *et al.* [5] in Britain. In 2006, more than 1, 000 people had acute gastrointestinal symptoms and 100 people were hospitalized in Japan due to the intake of undenatured lectins following consumption of insufficiently heated kidney beans (*P. vulgaris* L.) [6]. From 2004 to 2013, lectin-intake accounted for 53.3% of legumes (*P. vulgaris* L.) food poisoning events in China [7]. Furthermore, in South Africa, the cultivation or consumption of a variety of the *P. vulgaris* beans by importation has been prohibited because of the potential toxicity caused by lectins [8]. But, black turtle bean (*P. vulgaris* L.) related allergies are still lacking definitive cure [9, 10].

Lectins are carbohydrate-binding proteins presenting in most plants and a wide range of vegetables. Owing to the relative resistance of the allergen protein to proteolytic degradation, primary epitope regions of undigested lectins can bind to the epithelium in the gut and result in the nutrient deficiencies via damage to enterocyte brush border and/or interaction with digestive enzymes, and it can pass the epithelial barrier to stimulate the production of antibodies and/or induce T cell immunity [11, 12]. Then, nausea, emesis, diarrhea and other disrupting digestion clinical symptoms would be induced by IgE mediated immediate-type hypersensitivity reactions and non IgE mediated reactions (cell-mediated immunity) [2, 7, 13]. However, little information on legume lectin epitopes is hitherto available [11].

Known as antigenic determinants, it is important to identify the epitopes to notarize the type and intensity of the immune response, and to serve for vaccine design, disease prevention, diagnosis and treatments [14-16]. Thus, a large number of rapid, fairly accurate and cost effective *in silico* algorithms are being developed with an enormous expansion in protein structures [17]. In a recent study, combination strategy with Antigenic Peptides, BepiPred 1.0 Server and DNASTar analytical tools was accurately applied in the screen and recognition of B-cell epitopes of tropomyosin in *Penaeus monodon* [18]. Three B-cell epitopes and two T-cell epitopes from Per a 10 allergen of *Periplaneta Americana*, were successfully identified in conjunction with various B-cell prediction (ABCpred, Antigenic, BepiPred 1.0b, Bcepred, BCPREDS, DNASTAR, ElliPro and Epitepia) and T-cell prediction (MHCpred, SVRMHC and SMM-Alig, MULTIPRED, ProPred, RANKPEP and SVMHC) softwares, respectively [19].

Thus, in order to improve the *in silico* accuracy, different immunoinformatic tools in B and T-cell epitope identification have been adopted in the present study to explore the lectin from black turtle bean (*P. vulgaris* L.). The potential epitope peptides were synthesized and validated by ELISA, lymphocyte proliferation and cytokine profile analyses. Furthermore, the effect of amino acids on lectin epitopes was analysed by quantitative structure-activity relationship (QSAR) method.

Materials and Methods

*Purification of the lectin from black turtle beans (*P. vulgaris* L.)*

Black turtle beans (*P. vulgaris* L.) cultivated in Heilongjiang Province of China were obtained from a local market (Harbin, Heilongjiang, China). The lectin from the black turtle beans (*P. vulgaris* L.) with at least 94% purity was obtained according to our previous study [20]. The extracted lectins were freeze-dried and stored in a -80 °C refrigerator.

Synthesis of epitopes

The predicted epitope peptides and an unrelated peptide (LSREFLLGLE) were synthesized using standard Fmoc (9-fluorenylmethoxycarbonyl) solid phase peptide synthesis, and purified using an preparative high performance liquid chromatography (Pre-HPLC, HP Plus, Lisure Science Co., Ltd., Suzhou, China) equipping with a Galaksil EP C18 column (10 µm, 50×250 mm, Galak Chromatography Technology Co., Ltd., Wuxi, China) monitored at 214 nm using 0.1% trifluoroacetic acid in water as mobile phase A and 0.1% trifluoroacetic acid in acetonitrile as mobile phase B. A gradient from 5-30% B was applied for the separation at a flow rate of 1.0 ml/min, and the target peak was collected by a SIL-20A autosampler (Shimadzu Co., Kyoto, Japan). The purity of the synthesized peptide was assessed using the LC-MS system. HPLC (LC-20A, Shimadzu Co., Kyoto, Japan) was equipped with a UV-Vis detector and an Inertsil ODS-SP C18 column (5 µm, 4.6×250 mm, Shimadzu Co., Kyoto, Japan) with the limit of quantification (LOQ) and limit of detection (LOD) of 10 ng/ml and 5 ng/ml, respectively. The molecule weight and purity were further evaluated by the MS analysis (MS-2020A, Shimadzu Co., Kyoto, Japan).

*Homology modeling of lectin from black turtle beans (*P. vulgaris* L.)*

The allergen protein sequence can be retrieved from the NCBI protein database (<https://www.ncbi.nlm.nih.gov/>) using the ID: AHB17899.1. The homologous template suitable for lectin from black turtle beans (*P. vulgaris* L.) was selected and employed for homology modeling the lectin 3-D structure by SWISS-MODEL (<https://www.swissmodel.expasy.org/>). The model was visualized using Discovery Studio Visualizer Client 2017R2 version (Accelrys Inc., San Diego, CA, USA). Primary structure analysis was performed utilizing the ProtParam online tool (<http://web.expasy.org/protparam/>). The content of secondary structures was predicted by PSIPRED online tool (<http://bioinf.cs.ucl.ac.uk/psipred/>).

Stereochemical analysis and model evaluation

Structural evaluation and stereochemical analysis were performed using various evaluation and validation tools. Backbone conformation was evaluated by analyzing the Psi/Phi Ramachandran plot (<http://services.mbi.ucla.edu/SAVES/Ramachandran/>) [21]. The Z-score of the predicted structure was performed on ProSA-web server (<https://prosa.services.came.sbg.ac.at/prosa.php>) [22]. The model was further evaluated through ERRAT (<http://services.mbi.ucla.edu/ERRAT/>) [23]. Furthermore, visualization of the generated model was performed using Discovery Studio Visualizer Client 2017R2 version (Accelrys Inc., San Diego, CA, USA).

Prediction of B-cell epitope

In order to enhance the accuracy of prediction towards the B-cell epitopes, the complete amino acid sequence of lectin from black turtle beans (*P. vulgaris* L.) was analyzed using seven immunoinformatics-based computational servers including DNASTAR, ABCPred, BCPREDS, BepiPred-2.0, Antigenic, Bcepred and EpiToPIA. DiscoTope and ElliPro were used to identify discontinuous B-cell epitopes by submitting the homology modelled 3-D structure of the allergen protein for the secondary and 3-D structural analyses. The overlapped consensus epitope regions in the various computational results combined with the structural analysis of the protein were selected as the potential epitopes for the further analysis [24, 25].

Prediction of T-cell epitope

The prediction of T-cell epitopes of the lectin allergen was based on inhibitory concentration (IC50) values and binding score values obtained by quantitative prediction methods, including ProPred, SYFPEITHI, NetMHCII, NN-align, SMM-align, NetMHCIIpan and RANKPEP [19]. For HLA class II alleles, 4 most common HLA-DRB alleles (HLA-DRB1*03:01, HLA-DRB1*04:01, HLA-DRB1*11:01 and HLA-DRB1*15:01) were

selected, and the principle of more than three simultaneous HLA-DR-based T-cell epitope results predicted by at least five tools was applied to select the potential T-cell epitopes.

Enzyme linked immunosorbent assay (ELISA)

The antigenicity of the predicted epitopes was determined by ELISA method using the 96-well microtiter plates (Covalink Nunc TM Immunomodule, Roskilde, Denmark). After encapsulation with 1 µg/100 µl/well of lectin and 100 ng/100 µl/well of B or T-cell peptides/control peptides in carbonate-bicarbonate buffer (pH 9.6) and incubation overnight at 4 °C, the plates were washed with phosphate buffer saline (PBS) twice, and then were blocked with 5% defatted milk (200 µl/well) for 1 h at room temperature. After washing with PBS, the recombinant polyclonal anti-lectin antibodies obtained from Nanjing Senbeijia Biological Technology Co., Ltd. (Nanjing, Jiangsu, China) were diluted 1:100, 000 v/v in PBS and added to the wells except for the control, then the plate was incubated overnight at 4 °C. The plate was incubated with 1:1000 v/v diluted anti-human-IgE peroxidase antibody (Sigma-Aldrich Chemicals Co., St. Louis, MO, USA) at 37 °C for 2 h and washed with PBS. Prior to the detecting, the chromogenic reaction was developed using 3, 3',5, 5'-tetramethylbenzidine (TMB) for 10 min in the dark and was terminated with 0.2 M sulphuric acid. The absorbance was determined at 450 nm by a microtiter plate reader (Eon model, Inc., Winooski, Vermont, USA).

Cytokine profile and lymphocyte proliferation by identified T-cell epitopes

Peripheral blood mononuclear cells (PBMC) were isolated from heparinized blood of the lectin hypersensitive rats by Ficoll-histopaque®-1077 (Sigma-Aldrich Chemical Co., St. Louis, MO, USA) centrifugation. Heparinized blood (5 ml) diluted 1:1 with PBS was layered onto an equal volume of histopaque and centrifuged at 1000 g for 30 min at 25 °C. PBMCs were recovered and washed twice with PBS and resuspended in complete RPMI-1640 media with 10% fetal calf serum (FCS) and cultured 10⁵ cells/well in a 24 well plate. The cells were stimulated with the lectin (10 µg), T-cell epitope peptides (10 µg). Cells cultured with an unrelated peptide (10 µg) were served as control. The plates were cultured for 48 h with 5% CO₂. Culture supernatant was collected and centrifuged at 1200 g for 5 min, and then kept at -20°C for further analysis. IL-4 and IL-5 were determined in culture supernatants using the BD Pharmingen Opt EIA kits (BD Pharmingen, Heidelberg, Germany). The 3-[4, 5-dimethylthiazol-2-yl]-2, 5-diphenyl tetrazolium bromide (MTT) colorimetric assays on the cultured PBMCs were performed to measure lymphocyte proliferation. Cells were lysed in acidic isopropanol, and absorbance was read at 570 nm using an ELISA reader (Bio-Rad 550, Tokyo, Japan).

Amino acid analysis of identified epitopes

Amino acid descriptors (predictors, X) and antigenicity activity of B and T-cell epitopes represented by specific IgE binding ability (OD_{450 nm}) and inducing lymphocyte proliferation capacity (OD_{570 nm}) (dependent, Y) were applied for the quantitative structure-activity relationship (QSAR) analysis by SIMCA-P software (version 12.0.1, Umetrics, Umeå, Sweden) using partial least squares regression (PLS). The amino acid descriptor of 3 z-scale (Supplementary Table S1 - For all supplemental material see 10.1159/000493496/) was previously calculated by principal component analysis (PCA) from a matrix consisting of 29 physicochemical variables [26]. The z1 descriptor represents the hydrophobicity, z2 represents bulk and z3 represents electronic properties. In the peptide descriptor variable matrix X, the amino acid at the first position from the C-terminus was designated as c1, and its properties were described as c1z1, c1z2 and c1z3. Similarly, the amino acid at the second position from the C-terminus was designated as c2, and its properties were described as c2z1, c2z2, c2z3, etc.

Statistical analysis

Origin 8.5 and GraphPad Prism 7.01 softwares were employed for the statistical analysis. For multiple comparisons, Tukey's procedures were applied to identify the significant differences between the means. p < 0.0001 was considered significant.

Results

Lectin structural and quality assessments

The lectin from black turtle bean (*P. vulgaris* L.) was confirmed to belong to the PHA family. Its sequence was first identified in our previous study [27]. Lectin sequence (NCBI accession no. AHB17899.1) can be retrieved from NCBI database.

Primary structure analysis according to Protparam online tool (<http://web.expasy.org/protparam/>) is presented in Supplementary Table S2. The allergen protein consisting of 275 amino acids has a molecular weight of 29774.53 Da, and can be formulated by $C_{1346}H_{2095}N_{341}O_{418}S_1$. The theoretical isoelectric point (PI) is 4.84 and the protein thus negatively charged at physiological pH. According to instability index value of 19.64 (< 40), the allergen protein might be fairly stable. The grand average of hydropathicity (GRAVY) value of 0.068 (> 0) reveals the hydrophilicity of the allergen protein molecule. Approximately 60% of the amino acid sequence is confirmed as the hydrophilicity regions using the method of Kyte and Doolittle [28] at ProtScale server (<http://web.expasy.org/protscale/>), which also indicate the hydrophilicity of the allergen protein (Supplementary Fig. S1a). Secondary structure prediction with online server PSIPRED reveals that the lectin protein has 4.36% alpha helices regions (6-16: LLSLALFLVLL and 269-271: AFL), 38.18% beta sheets regions (25-29: TSFSE, 38-41: ILQR, 43-55: ATV, 50-55: QLRLT, 67-72: LGRAFY, 76-77: IQ, 79-80: WD, 89-98: FATSFTFNID, 108-114: GLAFVL, 127-128: LG, 141-145: TVAVE, 161-164: IGID, 169-178: KSIKTTTWDF, 183-192: NAEVLITYD, 196-203: LLVASLLVY, 209-217: SFIVSDTVD, 226-231: VIVGFT and 241-254: VETNDVLSWSFASK), and 57.45% random coils (Fig. 1a).

A typical workflow involving template selection, sequence alignment, model building and quality assessment was carried out to perform the homology modeling [29]. The 3-D structure of lectin from black turtle bean (*P. vulgaris* L.) was modeled based on the template of 3wcs.1.A (PHA-E from red kidney bean) by SWISS-MODEL (<https://www.swissmodel.expasy.org/>) [30]. The homology modeled structure of lectin was visualized by Discovery Studio Visualizer Client (2017R2 version, Accelrys Inc., San Diego, CA, USA), and the 3-D structures including homo-tetramer and monomer are presented in Fig. 1b and Fig. 1c, respectively. The modeled 3-D structure of lectin from black turtle bean (*P. vulgaris* L.) is in good agreement with the alpha helices, beta sheets and coils predicted by PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>).

Stereochemical quality and reliability of the structure were further checked by several structure assessment methods, including Ramachandran plots, Z-score and ERRAT. As shown in Fig. 1d, the majority of residues (94.5%) are in the favorable region in the Ramachandran plot. The Z-score of the protein was computed to be -7.02, which was within the range of scores usually found for native proteins of similar size (Fig. 1e) [22]. A high score of 91.365 was obtained by ERRAT (Fig. 1f).

B-cell epitope prediction

Seven immunoinformatic-based computational servers (DNASTAR, ABCPred, BCPREDS, BepiPred-2.0, Antigenic, Bcepred and Epitopia) and two discontinuous B-cell epitopes servers (DiscoTope and ElliPro) were employed in the B-cell epitope prediction. The predicted results from different immunoinformatic servers are shown in Supplementary Table S3-11. Since a wide range of predictions for epitopes were obtained, the protein structural screenings further employed basing 3-D structure. As shown in Supplementary Table S12, five potential epitope sequences, including 55-66: NVNDNGEPTLSS, 98-107: DVPNNSGPAD, 116-125: VGSEPKDKGG, 133-141: NNYKYDSNAHT and 149-160: LYNVHWDPKPRH with nearly 100% coils and fully surface exposed, were obtained by combining the analyses of secondary structure and spatial position.

Seven potential epitope regions were identified by the consensus combining various immunoinformatics tools (Supplementary Table S13), which were highly consistent with the above structural analysis results, except the two epitope regions of 20-31: NSATETSFSFQR and 68-79: GRAFYSAPIQIW. Each of the seven B-cell antigenic regions (Table 1), named

Fig. 1. (a): The secondary structure prediction of the lectin from black turtle bean (*Phaseolus vulgaris* L.) by PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>). (b): Homology modelling the tetramer of the lectin by SWISS-MODEL (<https://www.swissmodel.expasy.org/>). (c): Monomer of the lectin. (d): Ramachandran plot (<http://services.mbi.ucla.edu/SAVES/Ramachandran/>) for the validation of the lectin model. (e): ProSA analysis of the lectin model with Z-score (<https://prosa.services.came.sbg.ac.at/prosa.php>). (f): Error values for residues as predicted by ERRAT (<http://services.mbi.ucla.edu/ERRAT/>).

Table 1. B-cell epitopes of the lectin from black turtle bean (*Phaseolus vulgaris* L.) ultimately predicted by the consensus combining various immunoinformatics tools and the 3-D structure analysis

Number	Sequence	Position	Length	Secondary structure	Spatial position
B1	NSATETSFQQR	20-31	12	41.67% β -sheets, 58.33% coils	Inner
B2	NVNDNGEPTLSS	55-66	12	100% coils	Exposed
B3	GRAFYSAPIQIW	68-79	12	50% β -sheets, 50% coils	Inner
B4	DVPNNSGPAD	98-107	10	10% β -sheets, 90% coils	Exposed
B5	VGSEPKDKGG	116-125	10	100% coils	Exposed
B6	NNYKYDSNAHT	133-141	11	100% coils	Exposed
B7	LYNVHWDPKPRH	149-160	12	100% coils	Exposed

as B1 (20-31: NSATETSFSFQR), B2 (55-66: NVNDNGEPTLSS), B3 (68-79: GRAFYSAPIQIW), B4 (98-107: DVPNNSGPAD), B5 (116-126: VGSEPKDKGG), B6 (133-141: NNYKYDSNAHT) and B7 (149-160: LYNVHWDPKPRH), respectively, were applied for further identifications to prevent omissions.

Location distributions of these predicted epitopes on 3-D model of the lectin are visualized using Discovery Studio Visualizer Client (2017R2 version, Accelrys Inc., San Diego, CA, USA) (Fig. 2a, 2b and 2c). As presented, B1 and B3 epitopes are partially buried in the inner region of lectin tetramer, and B2, B4, B5, B6 and B7 are all situated on the surface of the allergen protein molecule. Furthermore, B1, B3 and B4 are found to be comprised of random coil and beta sheet structures, while B2, B5, B6 and B7 epitopes are all occurred in random coil regions.

The results of hydrophily and hydrophobicity characteristics using the method of Kyte and Doolittle [28] are presented in Supplementary Fig. S1a, which show that all the identified epitopes, except B3, lie in hydrophobic regions. As shown in Table 2, grand average of hydropathicity (GRAVY) was predicted by online tool ProtParam (<http://web.expasy.org/protparam/>). Then, B1, B2, B4, B5 and B6 were found to have negative GRAVY values (< 0), which were consistent with the results of hydrophily and hydrophobicity (Supplementary Fig. S1a). As shown in Supplementary Fig. S1b-2d and combined with the antigenicity predication in Table 1, the predicted epitope B1 may share high surface accessibility, but it is weak in average flexibility and antigenicity. Low scores in antigenicity, hydrophilicity, average flexibility and surface accessibility are also found in B3, B2 and B4, but, epitopes of B5 and B6 obtained higher scores.

T-cell epitope prediction

As presented in Table 3, three T-cell epitope regions, including T1 (39-47: LQRDATVSS), T2 (95-103: FNIDVPNNS), and T3 (256-264: SDGTTSEAL), were identified as promiscuous binders of MHC II by taking a consensus from the prediction data. T1 and T2 have high affinity for the HLA-DRB alleles of DRB1*0301, DRB1*0401, DRB1*1101, DRB1*1501,

Fig. 2. (a): Distribution of predicted B and T-cell epitopes in the lectin from black turtle bean (*Phaseolus vulgaris* L.) tetramer. (b) and (c): Predicted B-cell epitopes. (d): Predicted T-cell epitopes. B and T-cell epitope regions were highlighted with red.

Table 2. Grand average of hydropathicity (GRAVY) predicted by online tool ProtParam (<http://web.expasy.org/protparam/>)

Number	Sequence	Position	Length	GRAVY
B1	NSATETSFSFQR	20-31	12	-0.950
B2	NVNDNGEPTLSS	55-66	12	-1.150
B3	GRAFYSAPIQIW	68-79	12	0.200
B4	DVPNNSGPAD	98-107	10	-1.240
B5	VGSEPKDKGG	116-126	10	-1.420
B6	NNYKYDSNAHT	133-141	11	-2.127
B7	LYNVHWDPKPRH	149-160	12	-1.600
T1	LQRDATVSS	39-47	9	-0.444
T2	FNIDVPNNS	95-103	9	-0.544
T3	SDGTTSEAL	256-264	9	-0.533
Lectin	-	-	275	0.068
Unrelated Peptide	LSREFLLGLE	-	10	0.530

Table 3. T cell epitopes of the lectin from black turtle bean (*Phaseolus vulgaris* L.) ultimately predicted by combined HLA-DR allele, IC50 binding score and binding score

Number	Sequence	Position	Allele	ProPred MHC Class-II binding Score	SYFPEITHI Binding Score (%)	NetMHC II IC% Rank	NN-align IC50 nM binding Score	SMM-align IC50 nM binding Score	NetMHCIIpan IC50 nM binding score	RANKPEP Optimal score (%)
T1	LQRDATVSS	39-47	DRB1*0301	6.3	21	1.1	13.7	97	63.66	15.661
			DRB1*0401	4.98	26	1.1	17.4	66	15.62	10.473
			DRB1*1101	0.1	9	65	124.4	-	50.84	19.046
			DRB1*1501	-	24	85	59	74	42.79	19.046
T2	FNIDVPNNS	95-103	DRB1*0301	5.11	20	7.5	219.3	298	-	11.441
			DRB1*0401	5	22	4	19.6	92	69.6	6.52
			DRB1*1101	0.8	18	46	200.1	-	178.93	-
			DRB1*1501	-	18	48	-	-	188.33	12.471
T3	SDGTTSEAL	256-264	DRB1*0301	2.1	12	90	199.9	-	-	9.002
			DRB1*0401	0.4	18	85	-	344	147.24	8.73
			DRB1*1101	0.9	12	95	-	199	80.97	-
			DRB1*1501	-	14	75	-	97	25.03	16.829

DRB*0301, DRB*0401 and DRB*1101 with low IC50 binding values. As shown in Fig. 2d, the T-cell epitopes are localized both on the surface and interior of the allergen protein molecule. In detail, T1 is mainly in the β -sheets regions, and the T2 is made of the random coil and β -sheets structures, while T3 mainly lie in the random coil regions. Meanwhile, the GRAVY values of T-cell epitopes (T1, T2 and T3) are -0.444, -0.544 and -0.533, respectively (Table 2).

ELISA

The ten predicted epitopes (peptides) and an unrelated peptide (Table 2) were synthesized, and high purity > 95% was obtained as assessed by HPLC, while the molecular mass of the synthesized peptides was further confirmed by LC-MS (Supplementary Fig. S2).

The cutoff value for ELISA positive was considered as OD ≥ 3 times of the control value (OD value of 0.0940 ± 0.0035) [12]. As shown in Fig. 3a, compared to the control, five predicted epitopes (B2, B5, B6, B7 and T2) showed significantly high IgE binding ability ($p \leq 0.0001$). Identified B-cell epitopes (B1, B3 and B4) and T cell peptides (T1 and T3) showed almost negligible IgE binding.

Lymphocyte proliferation assay on PBMCs and cytokine profiling

A significant lymphocyte proliferation on stimulation with T-cell epitopes and lectin, respectively, was found in peripheral blood mononuclear cells (PBMC) (Fig. 3b). As shown in Fig. 3b, T1, T2 and black turtle bean lectin show significantly higher PBMC proliferation in relation to control ($p \leq 0.0001$), whereas no significant difference is observed in T3 group.

The releases of IL-4 and IL-5 cytokines in culture supernatant of PBMCs were measured. As shown in Fig. 3c and 3d, compared to control, significant increases ($p \leq 0.0001$) in IL-4 and IL-5 levels were found in incubation with T1, T2 and the extracted lectin, while no significant change ($p > 0.0001$) was induced by T3 cell epitopes. Furthermore, the IL-4 and IL-5 levels induced by T-cell epitopes were much lower than that of extracted lectin.

Analysis of amino acid composition of the B and T-cell epitopes

In order to explore the crucial amino acids making up the epitopes, the frequencies of amino acids presented in the whole lectin sequence and epitope regions were considered, respectively. As shown in Fig. 4a, nineteen amino acids are observed existing in the lectin allergen sequence (275 aa) with the absence of Cys residue. Moreover, the abundant quantities of Leu, Ser, Thr, Asn and Val residues are found to be 34, 30, 27, 22 and 22, respectively, in total sequence, but it is scarce in Met (1/275 aa), His (3/275 aa), Gln (5/275 aa), Arg (5/275 aa) and Trp (5/275 aa) residues.

In the four predicted B-cell epitope regions, fifteen different amino acids were observed with the absence of Met, Phe, Gln and Ile residues (Fig. 4a). The higher percentages of His (100%), Tyr (50%), Pro (40%), Asn (31.82%) and Lys (30.77%) residues are observed in identified B-cell epitope regions, and Ala (4.76%), Leu (5.88%) and Thr (7.4%) residues are seldom presented.

Fig. 3. (a): IgE binding of B and T-cell peptides were analyzed by ELISA with the recombinant polyclonal anti-lectin antibodies, and the wells without anti-lectin antibodies served as the control. (b): Lymphocyte proliferation assay. Cells stimulated with unrelated peptide were used as control. The proliferation was measured by MTT assay. (c): IL-4 and (d) IL-5 cytokines secreted in culture supernatants were detected by ELISA. Cells stimulated with unrelated peptide were used as control. The experiments were performed in five times and average was calculated. The data are presented as box and whiskers plot showing variation of IgE values from median. *** $p \leq 0.0001$ represented the significance of difference as compared to the control, ns: not significant.

In addition, twelve different amino acids were found in the two identified T-cell epitopes with the absence of Met, Glu, Lys, Gly, Tyr, Trp and His residues. The percentage of Gln (20%), Arg (20%), Asn (13.64%) and Asp (12.5%) residues were higher than other amino acids in T-cell epitope regions, which was different with the composition in B-cell epitope regions. Compared to the whole protein sequence, it seems that Gln, Arg, Asn and Asp residues are particularly abundant in the T-cell epitope regions.

Quantitative structure-activity relationship (QSAR) model was employed to analyze the relationship between potential antigenicity (Y) and the physicochemical characteristics (X) of amino acids that make up the epitopes. The PLS regression coefficients were computed and are shown in Fig. 4b-c. The QSAR model of B-cell epitopes using the 3 z-scale could explain 99% of the antigenicity variability (Y, R^2) with a predictive power of 75% (Q^2), indicating high model quality, while 99% variances of the T-cell epitope activity (Y, R^2) could be explained by the PLS components with cross-validation achieving 81%, suggesting the QSAR model was enough to reflect the changes of epitope activity.

As shown in Fig. 4b, positions of c1, c2 and c5 contain abundant hydrophobic amino acids in an active B-cell epitope peptide (decapeptide). Small chain groups and positively charged amino acids, and large chain groups with high hydrophilicity are more preferred in

Fig. 4. (a): The analysis of amino acid composition. Letter X1 represents that the number of amino acids presented in the lectin, Letter X2 and X3 represent that the percentage of amino acids (%) presented in B and T epitope regions, respectively. (b) and (c): PLS regression coefficients of the 3-z scale models of the identified B and T-cell epitopes dataset, respectively. The importance of a given X-variable is proportional to its distance (coefficient value) from the origin (zero). The bars illustrate 95% confidence intervals based on Jackknifing.

positions of c3, c4 and c9. Furthermore, high hydrophobicity and positively charged amino acids, such as Phe, Ile, Trp and Val residues, are preferred in positions of c6 and c7. Generally, the positively charged amino acids of Leu, Ala and His residues presenting in the epitope peptides positions of c1, c2, c3, c5, c6, c8 and c10 might significantly contribute to the B-cell epitope activity.

As presented in Fig. 4c, hydrophobic, positively charged amino acids with large bulk chains, such as Tyr, Phe, and Trp residues, are preferred in positions of c1, c2 and c3 in active T-cell peptides (nonapeptide). The amino acids with low hydrophobicity and charged properties as well as small chain groups, such as Asn and Asp residues, are favored in positions of c4 and c8. Hydrophobic and small chain groups with positively charged amino acids, such as Pro in positions of c5 and c7, would contribute to the positive T-cell epitope activity.

Discussion

The identified 3-D structure will potentially provide more insights into understanding the structure and function of allergen proteins [24, 25]. In the previous study, the epitope identification and display of Ara h 3 and some legumin allergens were perfectly performed by the accurate protein 3-D structure [31]. The structure of alcohol dehydrogenase (ADH) and Cur I 3 from *Curvularia lunata* were *in silico* predicted, and it was shown that computational tools applied for 3-D structure determination provide ideal approximation and can be employed for epitope identification [15, 32]. Recently, Per a 10 Allergen of *Periplaneta Americana* and Cur I 3 of *Curvularia lunata* were homology modelled, and their B and T-cell

epitope regions were identified successfully by *in silico* tools [18, 32].

In the present study, the lectin from black turtle beans (*P. vulgaris* L.), a homo-tetramer allergen protein, whose structure consisting of four identical or almost identical subunits, was modelled based on the template phytohemagglutinin from *P. vulgaris* (PHA-E) which has been solved experimentally by NMR [30], because of the highest sequence identity (98.43%) (Fig. 1b). The stereochemical evaluation by Ramachandram plots, Z-score and ERRAT demonstrated the homology-based model of lectin 3-D structure was reasonable (Fig. 1d-f).

Interactions among B and T cells play a vital role in the allergic response [15]. With an enormous expansion in knowledge and data about 3-D structures and epitopes of identified allergens, a large number of computational algorithms were developed [17, 33, 34]. The online tools are fast, precise and extremely efficient means of structure identification and mapping of epitope regions [35, 36]. Since the prediction accuracy of combined methods for B and T-cell epitope prediction is far higher than any single *in silico* tool, consensus method was used as it combined the strengths of different prediction algorithms [19, 35]. Nevertheless, due to a wide range of predictions for epitopes was obtained, further structural screenings were absolutely essential. Recent studies documented that the peptide sequences with low steric hindrance (surface exposed) and high flexibility would be most likely involved in IgE binding [25, 34]. It also suggested that the potential epitopes might be in random coil regions, as the structure was much looser than α -helix and β -sheet regions, which would be much prone to distribute and expose on the surface of the protein molecule [37].

Thus, B-cell epitopes were identified by combining the 3-D structure analysis and the consensus methods using various online tools and datasets (Table 1). Subsequently, validation of the identified epitopes by ELISA suggested that epitopes B2, B5, B6, B7 and T2 showed significant higher IgE binding than the control (Fig. 3a). In short, five B-cell epitopes were found to be antigenically effective in our present study, and it could be inferred that these epitopes are important in triggering of immune responses. Peptides with stronger IgE binding have low GRAVY values. Similarly, five identified IgE epitopes of the ovalbumin protein all occurred in relatively hydrophobic regions [38]. Besides, more than half of the total residues in the B-cell epitope regions of ADH allergen of *Curvularia lunata* were proven to be hydrophobic [15]. In the current study, the B6 epitope shows highest IgE binding among the predicted epitopes and has the highest grand average of hydrophobicity.

Furthermore, high average flexibility and surface accessibility of sequence could also be employed as the indicators for B-cell epitope identification, as the suggested sequence would be more flexible and easily exposed outside [19, 39]. Higher average flexibility and surface accessibility are found in epitopes B5 and B6, which displayed higher IgE binding capacity (Fig. 3a). All the predicted B-cell epitope were checked for their localization in the protein structure, and the validated epitopes (B2, B5, B6 and B7) were in the outside region of the protein (Table 1). The intrinsic physicochemical properties of the amino acids, such as hydrophilicity, hydrophobicity and overall charge, would affect IgE binding of the epitopes [15, 18]. The composition analyses revealed that the frequency of amino acids in B-cell epitopes was different with that in lectin, as Ala, Leu and Thr residues were abundant in the whole lectin sequence with fewer of His and Tyr residues (Fig. 4a). According to the previous study, the presence of Tyr and His residues would be attributed to the high IgE binding capacity due to the forming of protein interface hubs [40]. The amino acid composition analysis of ADH from *Curvularia lunata* showed that Pro and Lys residues had high propensities to lie in B-cell epitope regions [15]. Recently, Gly, Pro, Cys and Glu residues were found to be rich in IgE binding epitope peptides (Per a 10) of *Periplaneta americana* [19]. Amino acids in lectin B-cell epitopes were further analyzed by QSAR method, which indicated that the identified lectin B-cell epitopes tend to consist of hydrophobic residues, but positively charged residues such as His residue seemed to have a more important role in B-cell epitopes activity (Fig. 4b). According to the previous report, surface-exposed positively charged residues were usually found in most of IgE-binding regions of peanut (*Arachis hypogaea*) lectin [11]. An earlier study also showed that charged residues on the exterior of the allergen molecules, such as Ala and Leu residues, would make contribution to high solvent accessibility, which

probably raised the IgE binding capacity [15]. Since the identified epitope T2 overlapping with 6 amino acid residues of B4, the T-cell epitope presented stronger IgE binding capacity (OD value of 0.3784 ± 0.0385) (Fig. 3a), which might be a cryptic B-cell epitope. Similar results were obtained in the previous study as a T-cell epitope (p21–35) derived from Der P 2 could elicit strong antibody responses, and the results would remind us to check all predicted T-cell epitopes for the presence of B-cell epitope regions [32, 41].

Allergen-specific T cells can play an important role in the allergic reaction, and they are obvious targets for immunotherapeutic intervention in the disease [42, 43]. Immunotherapy with T-cell epitopes is more effective in relation to crude allergen mixtures or purified protein [44]. Knowledge of the dominant T-cell epitopes of allergens is a crucial step to develop T-cell targeted vaccine for lectin-specific allergen immunotherapy [17, 45]. Identification of promiscuous epitopes by computational approach capable of cross binding multiple HLA class II molecules were selected as potential T cell epitope. Although HLA II class binders can be identified by various publicly available online tools with various datasets, however, it is a daunting task to select potential T-cell epitopes as the abundant possibilities. Thus, four most common HLA-DRB alleles (HLA-DRB1*0301, HLA-DRB1*0401, HLA-DRB1*1101 and HLA-DRB1*1501) were employed for the prediction [15, 46]. Three T-cell epitope regions T1 (39–47: LQRDATVSS), T2 (95–103: FNIDVPNNS), and T3 (256–264: SDGTTSEAL) were finally identified by different prediction algorithms (ProPred, SYFPEITHI, NetMHCII, NN-align, SMM-align, NetMHCIIpan and RANKPEP) with multiple HLA II class binders (Table 3).

T1 and T2 displayed significant lymphocyte proliferation with increased IL-4 and IL-5 levels in lymphocyte proliferation in PBMCs and cytokine profiling (Fig. 2b–d), which could provide huge potential for the development of target vaccines, diagnostic purposes and for investigating the pathology of infectious diseases [16]. An earlier study well documented that oral immunotherapy with immunodominant T-cell epitope peptides could alleviate allergic reactions of the egg-white allergen of ovomucoid [42]. T-cell epitope peptides from major allergens could provide effective and safer (non-IgE-reactive) alternatives for the conventional immunotherapy treatment [44]. Five dominant CD4⁺ T-cell epitopes of Ara h 2 peptides have been identified as novel candidates for T-cell-targeted peanut allergy therapeutics [47].

As demonstrated in a previous study for Der P 2, the antigenicity could be provided by the net charge from the amino acids of Gln, Arg and His residues [41]. In our present study, amino acid composition analysis of T-cell epitopes presented that Gln, Arg, Asn and Asp residues are abundant in the T-cell epitope regions (Fig. 4a), thus may be of importance to trigger T-cell epitope activity. A previous study revealed that most of the T-cell epitopes of Cur I 3 (a major allergen of *Curvularia lunata*) were probably composed by the hydrophobic amino acids [32]. QSAR model revealed that high hydrophobic amino acids are preferred in the T-cell epitope regions, and favored in the majority of positions including c1, c2, c3, c5, c6 and c7 (Fig. 4c). The epitope peptides identified by specific IgE binding and PBMC proliferations from Per a 10 allergen of *Periplaneta Americana* were also found to be of high hydrophobicity [19].

Conclusion

In conclusion, four B-cell epitope regions (B2: NVNDNGEPTLSS, B5: VGSEPKDKGG, B6: NNYKYDSNAHT, B7: LYNVHWDPKPRH) and two T-cell epitope regions (T1: LQRDATVSS, T2: FNIDVPNNS) were validated combining bioinformatics computational tools and *in vitro* experiments in our present study. Positively charged amino acids with high hydrophobicity were found in the epitope regions of the lectin. The *in silico* approach provides an efficient method for epitope prediction, and those validated epitopes may be potentially applied in designing diagnostics and therapies for the lectin allergy.

Acknowledgements

The authors are grateful for the financial support from the National Natural Science Foundation of China (No. 31701524, No. 31771974), the Natural Science Foundation of Anhui Province (No. 1708085MC70), the Fundamental Research Funds for the Central Universities (No. JZ2018HGTB0245), the Anhui Provincial Science and Technology Major Project (No. 16030701081, No. 16030701084), and the Financial Grant from China Postdoctoral Science Foundation (No. 2017M611208, No. 2018T110211). The authors thank Mr. Jintao Gong and Miss Yingshuo Zhang, School of Food Science and Engineering, Hefei University of Technology, for helping in lymphocyte proliferation assay, and the authors would also like to thank Mr. Liangyu Sun, Bankpeptide Biological Technology Co., Ltd., in the technical service for peptide synthesis. This study was supported by the Deutsche Forschungsgemeinschaft and the Open Access Publishing Fund of the Eberhard-Karls- University of Tübingen, Germany.

Disclosure Statement

All authors have claimed that there is no existing conflict.

References

- 1 He S, Simpson BK, Sun H, Ngadi MO, Ma Y, Huang T: *Phaseolus vulgaris* Lectins: A Systematic Review of Characteristics and Health Implications. Crit Rev Food Sci 2018;58:70-83.
- 2 Kumar S, Verma AK, Das M, Jain SK, Dwivedi PD: Clinical complications of kidney bean (*Phaseolus vulgaris* L.) consumption. Nutrition 2013;29:821-827.
- 3 Vasconcelos IM, Oliveira JT: Antinutritional properties of plant lectins. Toxicon 2004;44:385-403.
- 4 Hamid R, Masood A: Dietary lectins as disease causing toxicants. PAK J Nutr 2009;8:293-303.
- 5 Noah ND, Bender AE, Reaidi GB, Gilbert RJ: Food poisoning from raw kidney beans. Brit Med J 1980;281:236-237.
- 6 Wang R, Ding F, Gao YJ, Wang XY, Qun LI: Epidemiological analysis for vegetal food poisoning in China. Chin J Food Hygiene 2016;5:580-584.
- 7 Ogawa H & Date K: The "White Kidney Bean Incident" in Japan, in *Lectins: Methods and Protocols*. ed. J. Hirabayashi Springer New York: New York, U.S. 2014;39-45.
- 8 Venter FS & Thiel PG: Red kidney beans-to eat or not to eat? S Afr Med J 1995;85:250-252.
- 9 Ballmer-Weber BK, Montserrat FR, Kirsten B, Marianne D, Matthew S, Mackie AR, Salt LJ, Jonathan O'BH, Riccardo A, Simona B: How much is too much? Threshold dose distributions for 5 food allergens. J Allergy Clin Immunol 2015;135:964-971.
- 10 Valenta R, Hochwallner H, Linhart B, Pahr S: Food Allergies: The Basics. Gastroenterology 2015;148:1120-1131.
- 11 Rougé P, Culerrier R, Granier C, Rancé F, Barre A: Characterization of IgE-binding epitopes of peanut (*Arachis hypogaea*) PNA lectin allergen cross-reacting with other structurally related legume lectins. Mo Immunol 2010;47:2359-2366.
- 12 Kumar S, Verma AK, Das M, Dwivedi PD: Molecular mechanisms of IgE mediated food allergy. Int Immunopharmacol 2012; 3:432-439.
- 13 Kumar S, Sharma A, Verma AK, Chaudhari BP, Das M, Jain SK, Dwivedi PD: Allergenicity potential of red kidney bean (*Phaseolus vulgaris* L.) proteins in orally treated BALB/c mice and passively sensitized RBL-2H3 cells. Cell Immunol 2013;284:37-44.
- 14 Yang X, Yu X: An introduction to epitope prediction methods and software. Rev Med Virol 2009; 9:77-96.
- 15 Nair S, Kukreja N, Singh BP, Arora N: Identification of B Cell Epitopes of Alcohol Dehydrogenase Allergen of *Curvularia lunata*. PLOS ONE 2011;6:e20020.
- 16 He Y, Rappuoli R, De Groot AS, Chen RT: Emerging vaccine informatics. J Biomed Biotechnol 2010;2010: 218590.

- 17 Soria-Guerra RE, Nieto-Gomez R, Govea-Alonso DO, Rosales-Mendoza S: An overview of bioinformatics tools for epitope prediction: Implications on vaccine development. *J Biomed Inform* 2015;53:405-414.
- 18 Zheng LN, Lin H, Pawar R, Li ZX, Li MH: Mapping IgE binding epitopes of major shrimp (*Penaeus monodon*) allergen with immunoinformatics tools. *Food Chem Toxicol* 2011;49:2954-2960.
- 19 Govindaraj D, Sharma S, Gaur SN, Lavasa S, Prasad N, Arora N: Immunogenic peptides: B & T Cell Epitopes of Per a 10 Allergen of *Periplaneta americana*. *Mol Immunol* 2016;80:24-32.
- 20 He S, Shi J, Walid E, Ma Y, Xue SJ: Extraction and purification of a lectin from small black kidney bean (*Phaseolus vulgaris*) using a reversed micellar system. *Process Biochem* 2013;48:746-752.
- 21 Sheik SS, Sundararajan P, Hussain ASZ, Sekar K: Ramachandran plot on the web. *Bioinformatics* 2002;18:1548-1549.
- 22 Wiederstein M, Sippl MJ: ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 2007;35:W407-W410.
- 23 Colovos C, Yeates TO: Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci* 1993;2:1511-1519.
- 24 Dall'Antonia F, Pavlovkeller T, Zangger K, Keller W: Structure of allergens and structure based epitope predictions. *Methods* 2014;66:3-21.
- 25 Oezguen N, Zhou B, Negi SS, Ivanciuc O, Schein CH, Labesse G, Braun W: Comprehensive 3D-modeling of allergenic proteins and amino acid composition of potential conformational IgE epitopes. *Mol Immunol* 2008;45:3740-3747.
- 26 Hellberg S, Sjoestroem M, Skagerberg B, Wold S: Peptide quantitative structure-activity relationships, a multivariate approach. *J Med Chem* 1987;30:1126-1135.
- 27 He S, Shi J, Li X, Ma Y, Xue SJ: Identification of a lectin protein from black turtle bean (*Phaseolus vulgaris*) using LC-MS/MS and PCR method. *LWT-Food Sci Technol* 2015;60:1074-1079.
- 28 Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 1982;157:105-132.
- 29 Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T: Protein structure homology modeling using SWISS-MODEL workspace. *Nat Protoc* 2009;4:1-13.
- 30 Nagae M, Soga K, Moritamatumoto K, Hanashima S, Ikeda A, Yamamoto K, Yamaguchi Y: Phytohemagglutinin from *Phaseolus vulgaris* (PHA-E) displays a novel glycan recognition mode using a common legume lectin fold. *Glycobiology* 2014;24:368-378.
- 31 Barre A, Jacquet G, Sordet C, Culerrier R, Rougé P: Homology modelling and conformational analysis of IgE-binding epitopes of Ara h 3 and other legumin allergens with a cupin fold from tree nuts. *Mol Immunol* 2007;44:3243-3255.
- 32 Sharma V, Singh BP, Gaur SN, Pasha S, Arora N: Bioinformatics and Immunologic Investigation on B and T Cell Epitopes of Cur l 3, a Major Allergen of *Curvularia lunata*. *J Proteome Res* 2009;8:2650-2655.
- 33 Hayes M, Rougé P, Barre A, Herouet-Guicheney C, Roggen EL: In silico tools for exploring potential human allergy to proteins. *Drug Discov Today: Disease Models* 2015;s17-18:3-11.
- 34 He L, Zhu J: Computational tools for epitope vaccine design and evaluation. *Curr Opin Viro* 2015;11:103-112.
- 35 Yang X, Yu X: An introduction to epitope prediction methods and software. *Rev Med Virol* 2009;19:77-96.
- 36 Sancheztrincado JL, Gomezperosanz M, Reche PA: Fundamentals and Methods for T- and B-Cell Epitope Prediction. *J Immunol Res* 2017;1:1-14.
- 37 Mishra A, Jain A, Arora N: Mapping B-cell epitopes of major and minor peanut allergens and identifying residues contributing to IgE binding. *J Sci Food Agr* 2016;96:539-547.
- 38 Mine Y, Rupa P: Fine mapping and structural analysis of immunodominant IgE allergenic epitopes in chicken egg ovalbumin. *Protein Eng* 2003;16:747-752.
- 39 Pomés A: Relevant B Cell Epitopes in Allergic Disease. *Int Arch Allergy Imm* 2010;152:1-11.
- 40 Brinda KV, Vishveshwara S: Oligomeric protein structure networks: insights into protein-protein interactions. *BMC Bioinformatics* 2005;6:1-15.
- 41 Wu B, Toussaint G, Vander EL, Granier C, Jacquemin MG, Saintremy JM: Major T cell epitope-containing peptides can elicit strong antibody responses. *Eur J Immunol* 2000;30:291-299.
- 42 Rupa P, Mine Y: Oral immunotherapy with immunodominant T-cell epitope peptides alleviates allergic reactions in a Balb/c mouse model of egg allergy. *Allergy* 2015;67:74-82.

- 43 Wambre E, James EA, Kwok WW: Characterization of CD4+ T cell subsets in allergy. *Curr Opin Immunol* 2012;24:700-706.
- 44 Larché M: Peptide immunotherapy for allergic diseases. *Allergy* 2007;62:325-331.
- 45 Groot ASD: Immunomics: discovering new targets for vaccines and therapeutics. *Drug Discov Today* 2006;11:203-209.
- 46 Gededzha MP, Mphahlele MJ, Selabe SG: Prediction of T-cell epitopes of hepatitis C virus genotype 5a. *Virology* 2014;11:187-199.
- 47 Prickett SR, Voskamp AL, Dacumos-Hill A, Symons K, Rolland JM, O'Hehir RE: Ara h 2 peptides containing dominant CD4+ T-cell epitopes: candidates for a peanut allergy therapeutic. *J Allergy Clin Immunol* 2011;127:1-5.