Journal of Medical Sciences

ISSN 1682-4474

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS Journal of Medical Sciences

ISSN 1682-4474
DOI: 10.3923/jms.2018.34.47

Research Article
B and T-cell Epitopes Based Vaccine Design in Api m3 Allergen of *Apis mellifera*: An Immunoinformatics Approach

[1]Nayem Zobayer and [2]Abul Bashar Mohammad Aowlad Hossain

[1]Department of Biomedical Engineering, Khulna University of Engineering and Technology, 9203 Khulna, Bangladesh
[2]Department of Electronics and Communications Engineering, Khulna University of Engineering and Technology, 9203 Khulna, Bangladesh

Abstract

Background and Objective: Api m3 is one of the major allergen in honeybee allergen family. Honeybee allergens can cause severe anaphylaxis, even leads to death. Being an economically important insect, human keeps close contact with it and often becomes life threatening not only in beekeepers but also for mass population. Therefore, vaccination can be a preferred method to counter this issue. This study aimed to identify effective epitopes for successful vaccine design. **Methodology:** In current study, immunoinformatics approach was applied to find out potential B-cell and T-cell epitopes of Api m3. **Results:** In quest of this, physicochemical properties of Api m3 were analyzed and found relative thermostable nature of Api m3 allergen and only 10.46% of the allergen consisted of beta turn region. Five MHC class I, T-cell epitopes were identified and through scrutiny of these T-cell epitopes led to find out YTEESVSAL as the best epitope. For MHC class II, T-cell epitopes YPKDPYLYYDFYPLE and GGPLLRIFTKHMLDV were found as most prominent T-cell epitopes of Api m3 allergen. Linear B-cell epitopes were predicted by using BCPREDS, ABCpred, BepiPred and Bcepred and results were validated by means of hydrophilicity, antigenicity, surface accessibility, flexibility and beta turn region. Current study also revealed that GDRIPDEKN and PHVPEYSSS, two 9 mer peptides could be the most effective B-cell epitopes of Api m3. **Conclusion:** These results can be very effective in designing potential therapeutics and venom allergen diagnosis of honeybee.

Key words: Epitopes, B-cell, T-cell, Api m3, allergen

Citation: Nayem Zobayer and Abul Bashar Mohammad Aowlad Hossain, 2018. B and T-cell epitopes based vaccine design in Api m3 allergen of *Apis mellifera*: An immunoinformatics approach. J. Med. Sci., 18: 34-47.

Corresponding Author: Nayem Zobayer, Department of Biomedical Engineering, Khulna University of Engineering and Technology, 9203 Khulna, Bangladesh

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Honey bee (*Apis mellifera*) is an economically important insect, which has diverse and prolong history of close interactions with human. Honey bee is widely distributed in different geographic regions all over the world and is native to South and Southeast Asia[1]. According to the WHO database of allergen[2], honey bee consists 12 allergen, (Api m1-Api m12), among these allergens Api m1, Api m2, Api m3, Api m5 and Api m10 showed more than 50% sensitization and considered as major allergens of honey bee venom[3].

These allergens of honey bee also cause severe anaphylaxis upon ingestion in human immune system. A study conducted with 92 patients below the age of 18, was admitted to hospital for medical attention due to anaphylactic reaction and among those patients food allergens were causal agents for 43%, drug/chemical allergens in 22% and hymenoptera sting allergens for 14%[4].

Most injuries results either from honey bee stinger or its venom. Honey bee venom can induce severe to acute allergic reactions in different parts of human body, such as the heart, kidney, nervous systems and eye[5]. Patients with venom allergy usually suffer from large local reactions. The proportion belonging to this group usually ranged from 2.4-26.4% of mass population and this range can be as high as 38% in beekeepers. Around 0.3-7.5% of studied population suffered from systemic anaphylaxis and severe systemic reaction was 14-43% among beekeepers. According to the study, mortality rate was 0.03-0.45 per million but this number might not be significant as many deaths caused by anaphylactic reactions to insect stings probably remained unreported[6].

Major honey bee allergens includes hyaluronidase, phosphatase and phospholipase-A2, these enzymes activate cells of immune system and via the mast cell degranulation influence histamine release and produce Immunoglobulin E (IgE)[3] can induce IgE-mediated hypersensitivity reactions can be induced by stings of hymenoptera in venom allergic patients and this reaction might ranged from local up to severe systemic reactions and even acute anaphylaxis[7].

Traditional medicines are available for treating honey bee venom allergy. The most frequently used medicine is epinephrine and usually provided through injection. Different types of histamine blocker or anti-histamines are also used. Ranitidine, famotidine and cimetidine are histamine type 2 blockers used to enhance the effects of diphenhydramine. Corticosteroids like prednisone or methylprednisolone are often applied to reduce the hyper immune responses. Venom specific immunotherapy is one of the most potential treatments of honey bee venom allergy as it is said to be effective and enhance quality life of patients. Moreover, venom specific immunotherapy economic analysis showed that it is not cost effective for preventing future systemic reactions[5].

However, there has been no report of Api m3 B-cell and T-cell epitopes spectra till to date. Therefore, the current study sought to analyze Api m3 protein sequences finding out its potential immunogenic B and T-cell epitopes using in immunoinformatics approaches. The main objective of epitope prediction was to design and bring up with hypoallergenic molecules that can replace venom related treatments and other traditional healing techniques. Therefore, the outcome of current study may aid value through providing devising new therapeutic modalities for immunotherapy of honey bee venom allergy.

MATERIALS AND METHODS

Sequence retrieval: The protein sequence of Api m3 was retrieved from UniProtKB database (http://www.uniprot.org/). The accession number of this protein assigned as AAY57281.1. Lastly, the sequence was saved in FASTA format for further analysis.

Primary and secondary structure analysis: Physiochemical properties of Api M3 were analyzed by using ExPASy ProtParam tool[8] and another ExPASy tool, Self-Optimized Prediction Method with Alignment (SOPMA)[9] was used to analyze secondary structure. ProtParam tool was applied to attain data on different physiological and chemical properties of Api M3. Amino acid composition, molecular weight, theoretical isoelectric point (pI), grand average hydropathicity (GRAVY), estimated half-life, extinction coefficient[10], instability index[11] and aliphatic index[12] of the protein were estimated using the default parameters by ProtParam tool. SOPMA was applied to analyze properties such as, helical pattern, globular regions, transmembrane helices, solvent accessibility, bend region, random coil and coiled-coil region.

Identification of T-cell epitope: Identification of consistent Cytotoxic T-Lymphocytes (CTL) epitopes was of utmost importance for contiguous vaccine design. Computational analysis of CTL epitopes reduced strenuous efforts of wet lab experiments, also mitigated unnecessary cost. A web-based server, NetCTL-1.2 [13] was used in present study for predicting human CTL epitopes. This epitope prediction tool provides combined scores based on three parameters, such as proteasomal cleavage, TAP transport efficiency and MHC

class I affinity. For identifying potential T-cell epitopes, the threshold value was set, which have sensitivity and specificity values of 0.89 and 0.94, respectively. Five epitopes carrying highest combined score were selected for further experimentation. The MHC-1 binding prediction[14] was used, the Stabilized Matrix Base Method (SMM)[15] to calculate IC_{50} values of peptide binding to different class 1 MHC molecules. Prior to this analysis, peptide length was set to 9 amino acids for both frequent and non- frequent allele. The resulted alleles having binding affinity IC_{50} less than 200 nm were considered for further analysis.

MHC class II epitopes were predicted by using Immune Epitope Database 3.0 from IEDB T-cell prediction tool[16]. The current study predicted MHC II HLA-DQA1*01:01/ DQB1*05:01, HLA-DRB1*11:01, HLA-DRB1*03:01 and HLA-DRB1*15:01 restricted T-cell epitopes by using this tool. This tool assessed antigenic portion of a query amino acid sequence by providing a score. Higher score enhanced greater probability of that region forming T-cell epitope.

Epitope conservancy prediction: Epitope conservancy played pivotal role in selecting effective epitopes. Epitope conservancy can be defined as the part of the protein sequences that restrain the epitope measured at or exceeding a specific level of identity[17].

Analysis of population coverage: Epitopes those were selected on the basis of IC_{50} value, further analyzed with IEDB population coverage tool for predicting population coverage for each epitope. IEDB population coverage tool is a web based tool that works on the basis of MHC binding and/or T-cell restriction data. This tool is highly efficient in determining the portion of individuals' response to a set of epitopes with known MHC restrictions. The current study applied this tool for attaining the following information on population coverage: (1) Prediction of population coverage, (2) Recognition of Human Leukocyte Antigen (HLA) combinations by the population and (3) HLA combinations recognized by 90% of the population (PC90). One by one, five epitopes and their MHC-I molecules were pasted in the query box and population coverage area selected before submission to the server for analysis[17].

Design of the three-dimensional (3D) T-cell epitope structure: In this study, PEP-FOLD server was used[18] for designing 3D structures of T-cell epitopes. PEP-FOLD applies a de novo approach to design 3D structures from a given amino acid sequence. This server was capable of designing 3D structure of 9-36 amino acid long peptide residues. PEP-FOLD worked on the principle of structural alphabet (SA) letters to explain the structural annotations of four consecutive amino acid residues, couples the predicted series of structural alphabet (SA) letters to a greedy algorithm as well as a coarse- grained force field[19,20]. PEP-FOLD server provided five models as proposed 3D structure and best model was selected for further analysis.

B-cell epitope prediction: B-cell epitope prediction tools of IEDB were applied to find out linear B-cell epitopes. For predicting quality B-cell epitopes, some significant features were counted. Most crucial features of B-cell properties were flexibility, antigenicity, surface accessibility and hydrophilicity[21]. In current study, four immunoinformatics tools, such as BCPREDS[22], ABCpred[23], BepiPred[24] and Bcepred[25] were applied for predicting linear the B-cell epitopes through full protein sequence of Api m3 by using their default threshold values. Outputs from four B-cell prediction tools were aligned and most common result was predicted as potential B-cell epitope. Lastly, flexibility, antigenicity, surface accessibility and hydrophilicity of predicted epitope was checked by using the Karplus and Schulz flexibility prediction[26], Kolaskar and Tongaonkar antigenicity scale[27], Emini surface accessibility prediction[28], Parker hydrophilicity prediction[29] and beta turn regions turn[30] , respectively of IEDB B-cell analysis resource.

Homology modelling and validation of Api m3: Homology modeling of Api m3 was conducted by using SWISS-MODEL[31]. Predicted model usually contains some errors in their primitive structure. Trouble shooting step is prerequisite to overcome with these issues. Therefore, validation of predicted model was performed by using different software those were frequently used for model verification, such as RAMPAGE, ERRAT, PROCHECK[32-34]. Lastly, the model was illustrated with its T-cell and B-cell epitope with the Pymol 2.0 software (https://pymol.org/2/).

RESULTS AND DISCUSSION

Physicochemical properties of Api m3: Analysis of physicochemical properties of Api m3 protein was conducted by using ExPASy protparam tool. Api m3 consisted of 373 amino acids in its protein sequence and it poses molecular weight around 44 KDa. From the total number of positively and negatively charged residues, it can be concluded that Api m3 is hydrophobic in nature. The instability index and aliphatic index was found 53.69 and 92.76, respectively. Instability index value greater than 40 is regarded as the instability of the protein.

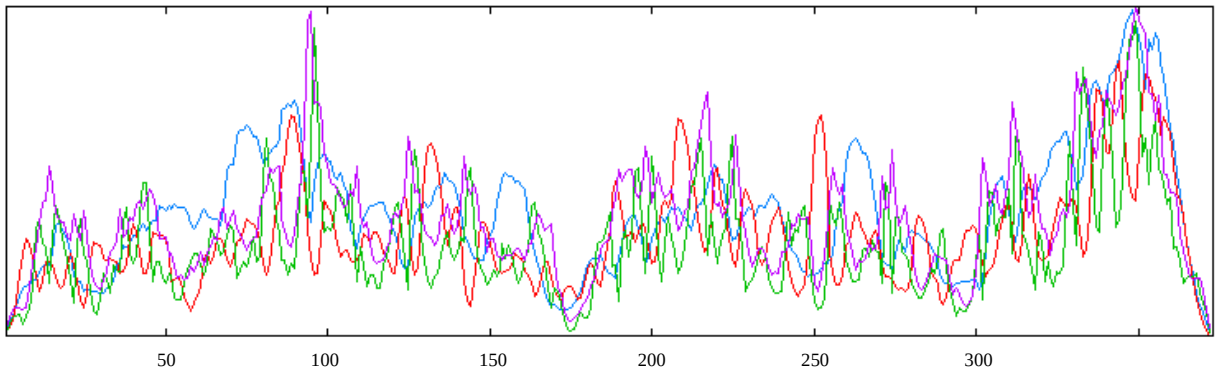

Fig. 1: Secondary structure plot of Api m3 of *Apis mellifera*
Helix is indicated by blue, random coils are shown in purple, while extended strands and beta turns are indicated by red and green, respectively

Table 1: Different physicochemical properties of Api m3 protein

Number of amino acids (Nos.)	373.000
Molecular weight (Da)	43905.190
Theoretical pI	5.630
Total number of negatively charged residues (Asp+Glu)	51.000
Total number of positively charged residues (Arg+Lys)	43.000
Instability index	53.690
Aliphatic index	92.760
Grand average of hydropathicity (GRAVY)	-0.430

Table 2: Representation of secondary structure data of Api m3

Features of secondary structure	Percentage
Alpha helix (Hh)	43.97
310 helix (Gg)	0.00
Pi helix (Ii)	0.00
Beta bridge (Bb)	0.00
Extended strand (Ee)	14.75
Beta turn (Tt)	10.46
Bend region (Ss)	0.00
Random coil (Cc)	30.83

So, Api m3 is fairly instable in nature and aliphatic index value suggests relative thermostability of Api m3 protein. It was also found that, the GRAVY value of protein was -0.43. The parameters calculated by ExPASy protparam tool is presented in Table 1.

Determination of secondary structure: Secondary structural features were calculated by SOPMA tool. Secondary structure of Api m3 dominated by alpha helix (43.97%) and random coil (30.83%). Minority part of the secondary structure poised by extended strand (14.75%) and beta turn (10.46%).

The features calculated by SOPMA tool is presented in Table 2. The secondary structure plot is shown in Fig. 1.

Prediction of T-cell epitope: T-cell epitopes were predicted by implying NetCTL 1.2 server which provides combined score based on each MHC class 1 supertypes, proteasomal C terminal cleavage and TAP (transport associated proteins) transport efficiency. The result generated by NetCTL 1.2 server for query sequence Api m3, we selected five epitopes based on the highest combinational score. Selected five epitopes were ITTPWDYYY, SVSALSSFY, FYDRTKMSL, YTEESVSAL and AEQSYGLTL, with the combinational score of 2.6286, 2.5280, 2.3858, 2.2899 and 2.0585, respectively.

Pursuant of IEDB MHC class1 binding prediction tool, the previously selected epitopes were found to be acquainted by wide range of MHC class 1 molecule. The tool is principled on Stabilized Matrix Method (SMM) and provides HLA binding affinity of selected epitopes based on IC_{50} value as an output. There is an inverse relationship between MHC-I molecules. In this study, threshold value of IC_{50} less than 200 nM was set for the section of MHC-I molecules. With this criteria, the selection of MHC-I molecules was ensured with higher affinity to bind with selected epitopes.

IEDB offers can excellent MHC-I processing efficiency tool, which predicts combined score for each of the previously selected epitopes based on three factors, such as proteasomal cleavage efficiency, TAP transport efficiency and MHC-I binding efficiency, accumulatively. Proteasomal cleavage efficiency stands for enzymatic digestion of a protein to form fragmented peptides. These peptide fragments are recognized by MHC-I molecules and simultaneously, forms MHC-Peptide complex. The complexes are then transported to the endoplasmic reticulum. This process rigorously influenced by TAP and it happens before being faced to T-cells on the plasma membrane of the cell. The higher combined score indicates more possibilities of accurate presentation of peptides and provoking a desired immune response. The combinational score attained from IEDB MHC-1 binding analysis and processing tools are presented in Table 3.

T-cell mediated immune response solely depends on successful recognition of designed peptides by HLA molecules with higher affinity. From that point, it can be concluded that the peptide recognized by most HLA alleles carry the high potentiality to induce a significant immune response. From selected epitopes, one epitope out smarted others by the mean of HLA allele interactions. The 9-mer peptide epitope YTEESVSAL showed interactive affinity for 15 MHC class I molecules, which includes HLA-A*02:06, HLA-B*39:01, HLA-C*02:02, HLA-C*02:09, HLA-C*03:03, HLA-C*03:02, HLA-C*03:04, HLA-C*05:01, HLA-C*08:01, HLA-C*08:02, HLA-C*12:02, HLA-C*12:03, HLA-C*14:02, HLA-C*15:02 and HLA-C*16:01 (Table 3).

Table 3: Five most potential T-cell epitope candidates with interacting MHC-1 alleles, total processing score and epitope conservancy result

Epitopes	Interacting MHC-1 allele with an affinity, 200 (total score of proteasome score, TAP score, MHC score, processing score and MHC-1 binding)	Epitope conservancy analysis result (%)
ITTPWDYYY	HLA-A*01:01 (1.12)	100
	HLA-A*29:02 (1.78)	
	HLA-A*30:02 (2.52)	
	HLA-A*68:01 (0.44)	
	HLA-B*35:01 (0.33)	
	HLA-B*57:01 (0.41)	
	HLA-B*58:01 (1.16)	
	HLA-C*16:01 (0.41)	
	HLA-C*03:02 (0.34)	
SVSALSSFY	HLA-A*01:01 (0.23)	100
	HLA-A*03:01 (0.29)	
	HLA-A*11:01 (0.81)	
	HLA-A*26:01 (0.23)	
	HLA-A*29:02 (1.40)	
	HLA-B*15:01 (0.42)	
	HLA-B*15:02 (0.43)	
	HLA-A*30:02 (1.16)	
	HLA-B*35:01 (0.50)	
	HLA-B*15:25 (0.68)	
	HLA-A*68:01 (0.60)	
	HLA-C*12:03 (0.10)	
FYDRTKMSL	HLA-C*14:02 (0.90)	100
	HLA-C*07:02(-0.16)	
YTEESVSAL	HLA-A*02:06 (-0.47)	100
	HLA-B*39:01 (-0.59)	
	HLA-C*02:02 (-0.46)	
	HLA-C*02:09 (-0.46)	
	HLA-C*03:03 (0.57)	
	HLA-C*03:02 (0.20)	
	HLA-C*03:04 (0.57)	
	HLA-C*05:01 (0.01)	
	HLA-C*08:01 (-0.01)	
	HLA-C*08:02 (-0.09)	
	HLA-C*12:02 (-0.33)	
	HLA-C*12:03 (0.24)	
	HLA-C*14:02 (-0.19)	
	HLA-C*15:02 (-0.49)	
	HLA-C*16:01 (0.10)	
AEQSYGLTL	HLA-B*13:01 (-0.27)	100
	HLA-B*37:01 (-0.40)	
	HLA-B*40:01 (0.96)	
	HLA-B*40:02 (0.85)	

Also, MHC class II T-cell epitopes were found based on HLA-DQA1*01:01/DQB1*05:01, HLA-DRB1*11:01, HLA-DRB1*03:01 and HLADRB1*15:01 restriction are presented in Table 4. YPKDPYLYYDFYPLE and GGPLLRIFTKHMLDV were found as prominent MHC-II type T-cell epitopes of Api m3 protein (Table 4).

Systematic selection of T-cell epitopes can be very useful in stimulation of clinical and immunological resistance to allergen in a refined treatment strategy. From that point, it is predicted that epitopes critically evaluated as per current theorems of potential allergen T-cell epitopes[35]. Researchers found effective T-cell epitopes these can effectively trigger immunogenic Th2 response in different food and insect allergen[36-39].

Prediction of epitope conservancy: In present study, five 9 mer T-cell epitopes for MHC class I molecules and two 15 mer T-cell epitopes for MHC class II molecules were predicted. Conservancy of these epitopes is expected as the more conserved epitopes provides better immunization. All seven of predicted epitopes including five MHC class I and two MHC class II peptides showed 100% conservancy, data is shown in Table 3 and 4, respectively.

Population coverage prediction: Suitable epitopes should cover wide range of ethnic population all over the world. For the purpose, predicted epitopes population coverage was checked through IEDB's population coverage analysis tool. From predicted MHC class I epitopes, MHC-I epitopes ITTPWDYYY, SVSALSSFY, FYDRTKMSL, YTEESVSAL and AEQSYGLTL were found, covered 9.84, 9.84, 3.42, 62.71 and 4.66% population of the world on an average, respectively (Table 5). YTEESVSAL covered around 62.71% population of

Table 4: Predicted two MHC class II T-cell epitopes with their position, length and conservancy score

Epitopes	Interacting MHC-II	Position	Number of amino acid residues	Epitope conservancy (%)
YPKDPYLYYDFYPLE	HLA-DQB1*05:01	23-37	15	100
	HLA-DRB1*03:01			
GGPLLRIFTKHMLDV	HLA-DRB1*11:01	226-240	15	100
	HLA-DRB1*15:01			

Table 5: Population coverage of our predicted five MHC class I T-cell epitopes

	ITTPWDYYY			SVSALSSFY			FYDRTKMSL			YTEESVSAL			AEQSYGLTL		
	Class I			Class I			Class I			Class I			Class I		
Areas	Coverage[a] (%)	Average hit[b]	pc 90[c]	Coverage[a] (%)	Average hit[b]	pc 90[c]	Coverage[a] (%)	Average hit[b]	pc 90[c]	Coverage[a] (%)	Average hit[b]	pc 90[c]	Coverage[a] (%)	Average hit[b]	pc 90[c]
Central Africa	2.84	0.03	0.10	2.84	0.03	0.10	3.21	0.03	0.10	56.62	0.66	0.23	0.53	0.01	0.10
East Africa	11.24	0.11	0.11	11.24	0.11	0.11	1.07	0.01	0.10	49.22	0.57	0.20	2.76	0.03	0.10
East Asia	2.55	0.03	0.10	2.55	0.03	0.10	11.28	0.11	0.11	81.27	1.20	0.53	0.06	0.00	0.10
Europe	25.67	0.26	0.13	25.67	0.26	0.13	2.01	0.02	0.10	69.56	0.89	0.33	0.18	0.00	0.10
North Africa	14.18	0.14	0.12	14.18	0.14	0.12	2.48	0.02	0.10	59.39	0.71	0.25	1.34	0.01	0.10
North America	12.72	0.13	0.11	12.72	0.13	0.11	3.55	0.04	0.10	65.58	0.83	0.29	9.59	0.10	0.11
Northeast Asia	3.44	0.03	0.10	3.44	0.03	0.10	6.91	0.07	0.11	71.42	0.94	0.35	21.57	0.22	0.13
Oceania	11.38	0.11	0.11	11.38	0.11	0.11	0.67	0.01	0.10	57.64	0.70	0.24	0.05	0.03	-59.00
South America	6.03	0.06	0.11	6.03	0.06	0.11	1.93	0.02	0.10	53.75	0.63	0.22	0.00	0.00	0.10
South Asia	13.80	0.14	0.12	13.80	0.14	0.12	4.23	0.04	0.10	59.80	0.75	0.25	2.60	0.03	0.10
Southeast Asia	2.10	0.02	0.10	2.10	0.02	0.10	4.34	0.04	0.10	73.65	1.00	0.38	10.93	0.11	0.11
Southwest Asia	14.66	0.15	0.12	14.66	0.15	0.12	2.75	0.03	0.10	47.69	0.55	0.19	1.65	0.02	0.10
West Africa	7.31	0.07	0.11	7.31	0.07	0.11	0.00	0.00	0.10	69.66	0.86	0.33	0.00	0.00	0.10
Average	9.84	0.10	0.11	9.84	0.10	0.11	3.42	0.03	0.10	62.71	0.79	0.29	3.94	0.04	-4.40
Standard deviation	6.47	0.07	0.01	6.47	0.07	0.01	2.85	0.03	0.00	9.66	0.18	0.09	6.16	0.06	15.60

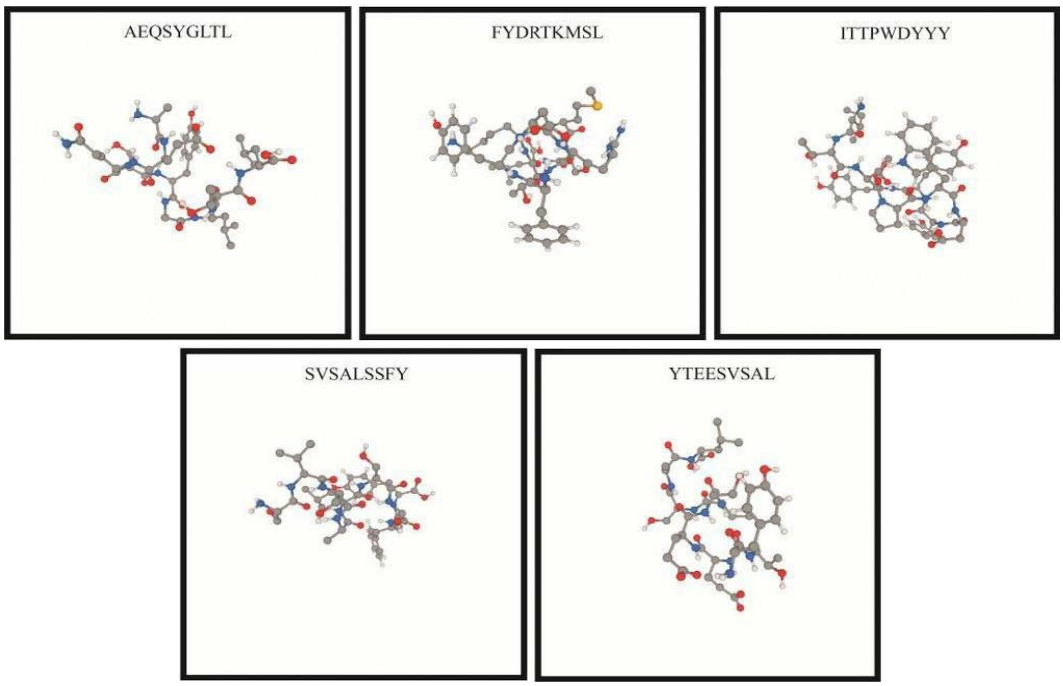

Fig. 2: Ball and stick structure of five MHC class I T-cell epitopes

Table 6: Prediction of B-cell epitopes of Api m3 of *Apis mellifera* by using BepiPred

Start	End	Peptide	Length (amino acids)
12	30	GDRIPDEKNEMYPKDPYLY	19
35	50	PLERGELTNSGKMREY	16
61	63	GDF	3
66	75	DIYTEESVSA	10
78	82	SFYDR	5
95	114	YPPNKLQQWNEDLNWQPIAT	20
117	119	LRR	3
124	125	IF	2
140	173	VLESPRGKYEFSKYDKLKKKLEEWTGKNITTPWD	34
183	184	AE	2
186	200	SYGLTLPSWTNNIFP	15
202	205	GELF	4
213	220	NITNSTPL	8
241	248	VSGTQKKK	8
256	261	GHESNI	6
273	280	PHVPEYSS	8
289	292	IEGT	4
303	312	IPSEARELQL	10
331	339	VIPSNEELI	9
344	354	FVDESANNLSI	11

the world on an average. This epitope showed 81.27% coverage in East Asia, 71.42% coverage in Northeast Asia, 73.65% coverage in Southeast Asia, 69.66% coverage in West Africa, 59.39% coverage in North Africa, 69.56% coverage in Europe and 65.58% coverage in North America. In case of MHC class II epitopes, much lower population coverage was found, comparing to the MHC class I epitope population coverage. YPKDPYLYYDFYPLE and GGPLLRIFTKHMLDV showed average 17.5 and 18.85% world population coverage.

Designing of 3D structure of T-cell epitopes: 3D structures of T-cell epitopes were predicted for better understanding the molecular actions of these peptides. Designed models are illustrated in Fig. 2.

Prediction of B-cell epitopes: Primarily, linear B-cell epitopes were predicted by using BepiPred, Bcepred, ABCpred and BCPREDS are shown in Table 6-9.

Table 7: Prediction of B-cell epitopes of Api m3 of *Apis mellifera* by using Bcepred

Start position	Sequence	End position
24	PKDPYLYYDFYP	35
69	TEESVSALSS	78
85	MSLQLVLAALYPPN	98
111	PIATKYLR	118
122	DNIFLPEDCLLFT	134
136	ELDRVLE	142
172	WDYYYIYHTLVAEQSYGLTL	191
205	FDATVFT	211
216	NSTPLLKKLYGGPLLRI	232
236	HMLDVVSG	243
249	RKIYLFS	255
260	NIASVLHALQLYYPHVPEYSSSIIME	285
291	GTHYVKIVYYLGI	303
308	RELQLPGCEVLCPLYKYLQLIENVIP	333
336	EELICDK	342
353	SIEELDFVKLNLIR	366

Table 8: Prediction of B-cell epitopes of Api m3 of *Apis mellifera* by using ABCPred

Rank	Sequence	Start position	Score
1	GDRIPDEKNE	12	0.86
2	NEMYPKDPYL	20	0.79
3	TPLLKKLYGG	218	0.76
3	NITTPWDYYY	167	0.76
4	LNWQPIATKY	107	0.75
5	ALSSFYDRTK	75	0.74
5	PLLRIFTKHM	228	0.74
5	DYYYIYHTLV	173	0.74
6	KYDKLKKKLE	152	0.72
7	IYHTLVAEQS	177	0.71
7	PEDCLLFTIE	127	0.71
8	VIFRHGDRIP	7	0.70
8	LFTIELDRVL	132	0.70
8	YEDNIFLPED	120	0.70
9	RERYGDFLGD	57	0.69
9	EWTGKNITTP	162	0.69
10	KDPYLYYDFY	25	0.68
10	ITNSTPLLKK	214	0.68
10	LFDATVFTYN	204	0.68
10	YGLTLPSWTN	187	0.68
10	AEQSYGLTLP	183	0.68
11	YYDFYPLERG	30	0.67
11	VFTYNITNST	209	0.67
12	KQINVIFRHG	3	0.66
13	LQLVLAALYP	87	0.64
13	YPLERGELTN	34	0.64
13	NIFPRGELFD	197	0.64
14	GKMREYQLGQ	45	0.62
14	IATKYLRRYE	112	0.62
15	KKLYGGPLLR	222	0.61
15	RVLESPRGKY	139	0.61
16	DIYTEESVSA	66	0.60
17	RTKMSLQLVL	82	0.59
18	LYPPNKLQQW	94	0.57
19	KYEFSKYDKL	147	0.54
20	EYQLGQFLRE	49	0.52

Validation of predicted B-cell epitopes: Validation of B-cell epitopes were preformed based on the data from

Table 9: Prediction of B-cell epitopes of Api m3 of *Apis mellifera* by using BCPREDS

Position	Epitope	Score
10	RHGDRIPDEKNE	0.975
270	LYYPHVPEYSSS	0.943
210	FTYNITNSTPLL	0.927
66	DIYTEESVSALS	0.863
144	PRGKYEFSKYDK	0.862
96	PPNKLQQWNEDL	0.827
53	GQFLRERYGDFL	0.765
163	WTGKNITTPWDY	0.749
195	TNNIFPRGELFD	0.691
334	SNEELICDKRFV	0.665
246	KKKRKIYLFSGH	0.621
313	PGCEVLCPLYKY	0.616
27	PYLYYDFYPLER	0.363

Table 10: Emini surface accessibility prediction

Start	End	Peptide	Length
14	29	RIPDEKNEMYPKDPYL	16
44	51	SGKMREYQ	8
79	85	FYDRTKM	7
96	108	PPNKLQQWNEDLN	13
115	121	KYLRRYE	7
143	161	SPRGKYEFSKYDKLKKKLE	19
244	250	TQKKKRK	7
273	278	PHVPEY	6

Table 6-9. Potential B-cell epitopes poise certain features which are essential for effectual recognition by B-cells[40]. These characteristics comprise hydrophilicity, antigenicity, surface accessibility, flexibility and beta turn prediction. Api m3 protein sequence was analyzed through IEDB tools in order to select best epitope from Table 6-9.

Surface accessibility is an essential feature of B-cell epitopes. Determination of surface accessibility is prerequisite to select potent B-cell epitope because it indicates most exposed area on the surface of a protein and likely to trigger B-cell mediated immune response. Also, hydrophilic regions of a protein are also generally exposed on the surface. The Emini surface accessibility prediction tool and Parker hydrophilicity prediction tools were used to attain data on surface accessibility and hydrophilicity of Api m3 protein. In Table 10, data of Emini surface accessibility prediction is presented and Fig. 3 shows hydrophilicity of Api m3 protein.

Kolaskar and Tongaonkar antigenicity prediction tool evaluated assessed antigenic propensity value of the protein for B cell epitopes. The Kolaskar and Tongaonkar antigenicity prediction tool subsequently provided the result, shown in Table 11. The average antigenic propensity value of Api m3 was calculated 1.034,

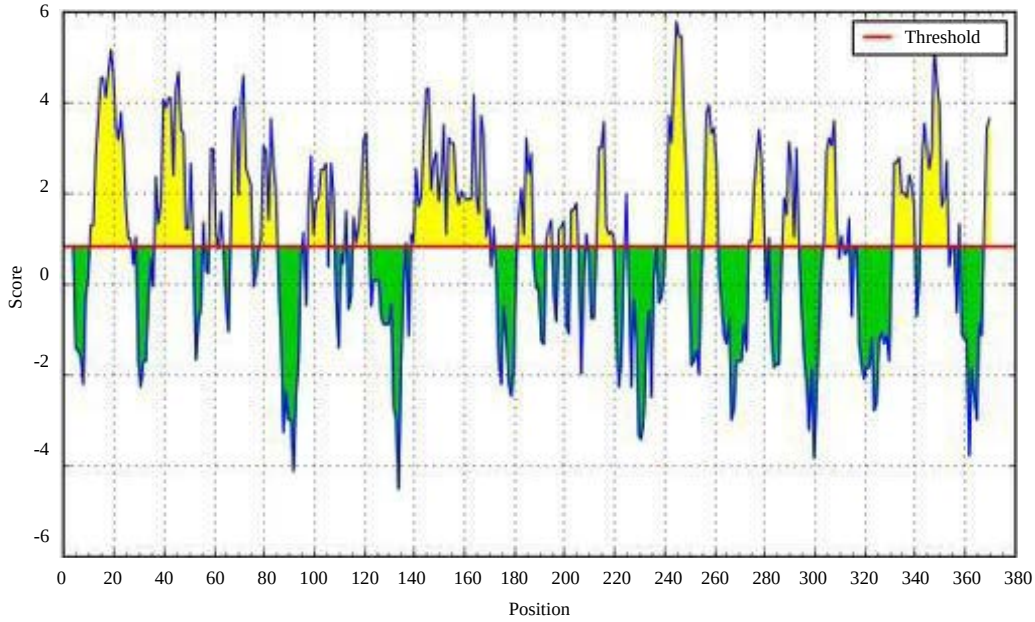

Fig. 3: Parker hydrophilicity prediction of Api m3 protein of *Apis mellifera*. The threshold value is 0.822
In the graph, hydrophilic regions in the protein are shown in yellow color and above is the threshold value

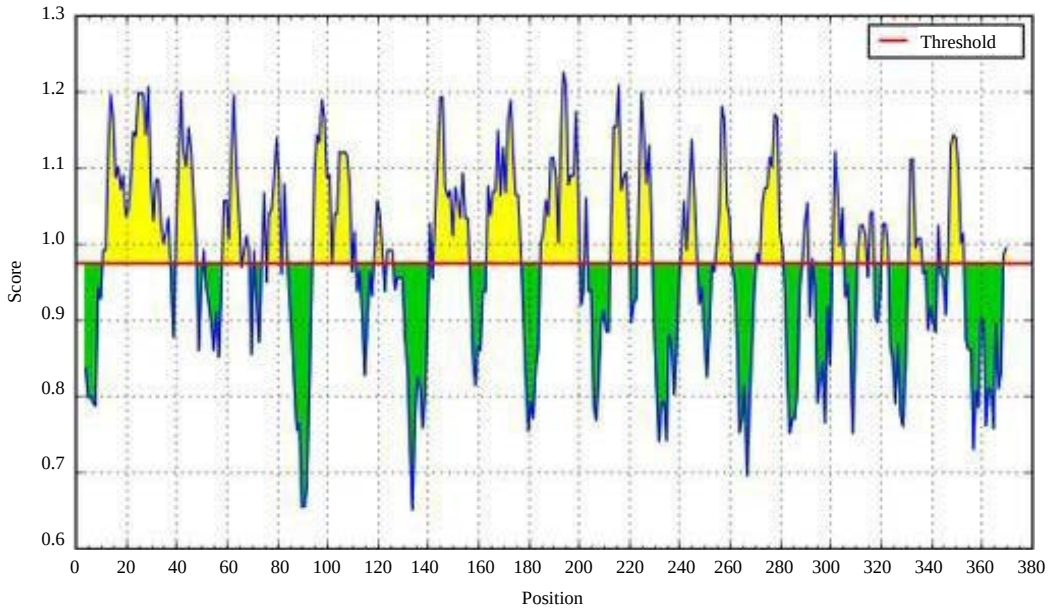

Fig. 4: Chou and Fasman Beta turn prediction of Api m3 protein of *Apis mellifera*. The threshold value is 0.975
In the graph, regions having beta turns in the protein are shown in yellow color and above is the threshold value

whereas the maximum value was 1.232 and minimum value was 0.869. Potent antigenic region was found from 261-281.

Another feature of B-cell epitope is the beta turns. Theoretically, beta turn regions of a protein are mostly hydrophilic and functionally surface accessible. Chou and Fasman Beta turn prediction was performed to find out potential antigenic region of a protein, as these two factors are positively correlated. Regions from 13-37 and 273-280 were found which are potential beta turn regions (Fig. 4) and 14-29 and 273-278 was most accessible region on the surface (Table 10). These regions were the best candidates to from B-cell epitopes.

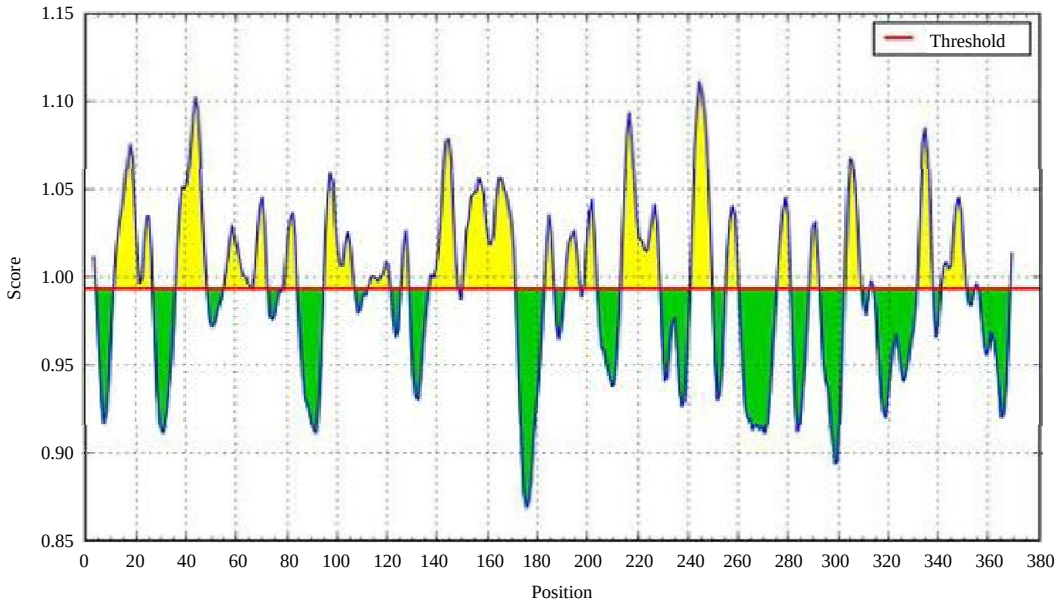

Fig. 5: Karplus Schulz flexibility prediction of Api m3 protein of *Apis mellifera*. The threshold value is 0.992
In the graph, regions having greater flexibilities in the protein are shown in yellow color and above is the threshold value

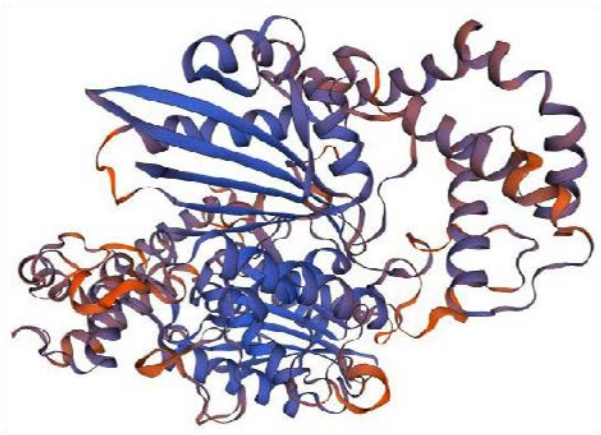

Fig. 6: 3D model of Api m3 protein of *Apis mellifera*

Table 11: Kolaskar and Tongaonkar antigenicity prediction

Start	End	Peptide	Length
4	10	QINVIFR	7
26	34	DPYLYYDFY	9
73	79	VSALSSF	7
87	98	LQLVLAALYPPN	12
127	135	PEDCLLFTI	9
137	143	LDRVLES	7
174	185	YYYIYHTLVAEQ	12
206	212	DATVFTY	7
226	233	GGPLLRIF	8
237	243	MLDVVSG	7
261	281	IASVLHALQLYYPHVPEYSSS	21
294	304	YVKIVYYLGIP	11
312	331	LPGCEVLCPLYKYLQLIENV	20
359	365	FVKLNLI	7

Flexible region of a protein aids in successful binding of an antibody to an epitope. Therefore, flexibility of Api m3 was predicted by Karplus Schulz flexibility prediction tool (Fig. 5).

Considering all the data available from different B-cell prediction tools, it is concluded that, the regions from 12-20 (GDRIPDEKN) and 273-281(PHVPEYSSS), two 9 mer peptides could be the most effective B-cell epitopes.

Homology modeling of Api m3: Homology modeling of Api m3 was performed by SWISS MODEL. The best model selected from SWISS MODEL (Fig. 6) checked for general errors by checking the model quality through ERRAT, RAMPAGE and PROCHEK.

ERRAT validates models by statistical relation of non-bonded interactions among different atom types based on characteristic atomic interaction[18]. ERRAT evaluates overall quality of a model at 0.01 and 0.05% level of significance and shows result as overall quality factor. Standard high resolution structures generally produce overall quality factor around 95% or higher. Low resolution structures produces overall quality factor around 91%. Selected model of Api m3 scored 81.505 as overall quality factor (Fig. 7).

RAMPAGE is another 3D model validation tool, which presents result based on amino acids geometry and deviation.

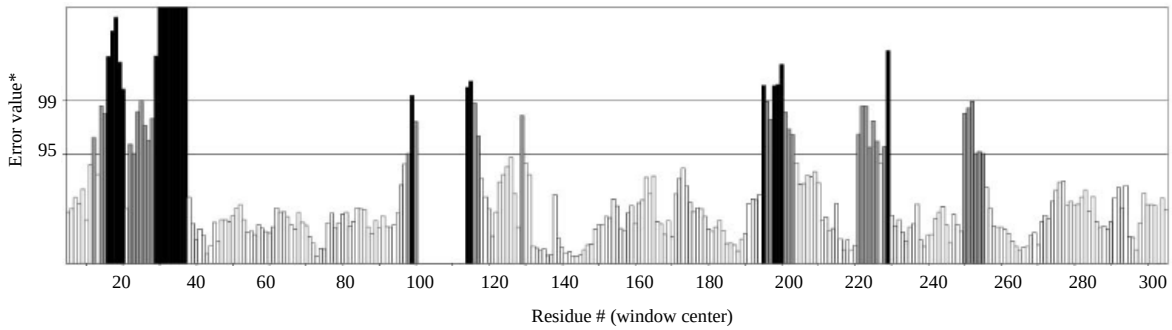

Fig. 7: ERRAT generated result of Api m3 protein of *Apis mellifera*, where 95% indicates rejection limit

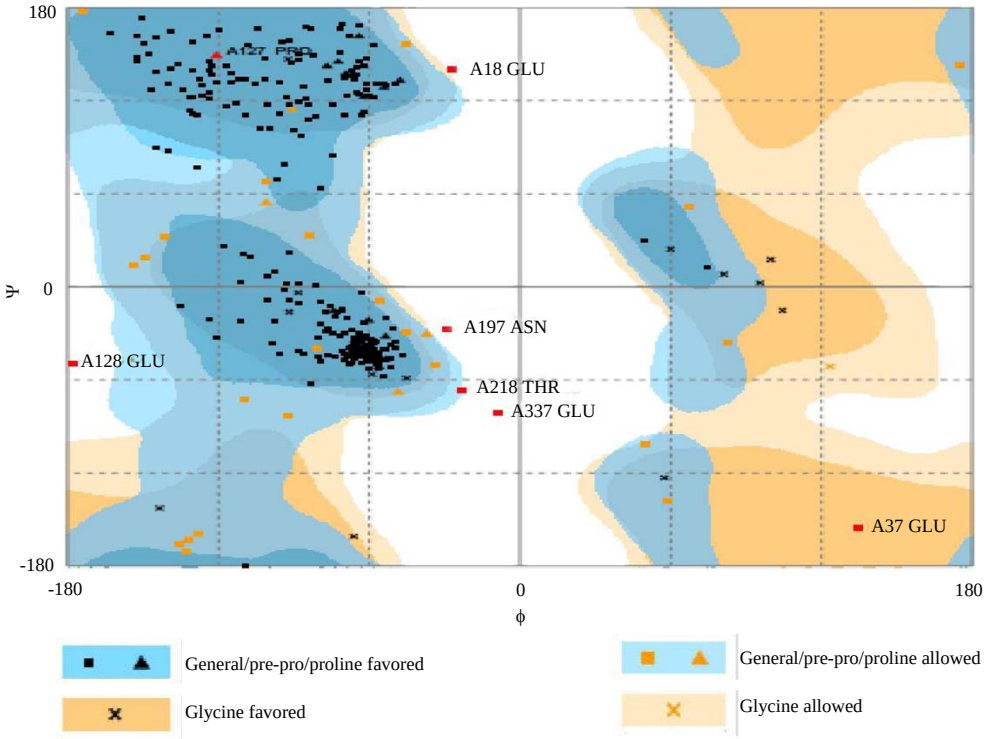

Fig. 8: Color coded results attained from RAMPAGE. Maximum cluster was observed in favored regions (90.0%)

It provides result in three major categories, such as number of residues in favoured region (expected value ~98.0%), number of residues in allowed region (~2.0% expected) and number of residues in outlier region. Selected model of Api m3 showed 90.0% amino acid residue resided in the favored regions (Fig. 8).

PROCHECK evaluates stereochemical quality of predicted protein structure by checking residue-by-residue geometry and overall structural geometry. It provides amino acid residues distribution on Ramachandran plot divided into four colour coated regions. Residues in most favored regions colored in red, residues in additional allowed regions colored in yellow, residues in generously allowed regions colored in pale yellow and residues in disallowed regions colored in white (Fig. 9). According to PROCHECK standard, a good quality model should posses over 90% amino acid resided in the most favored regions. Selected model of Api m3 showed 85.4% amino acid residue resided in the most favored regions.

After validation of selected predicted model, proposed T-cell and B-cell regions were denoted onto the model (Fig. 10-12).

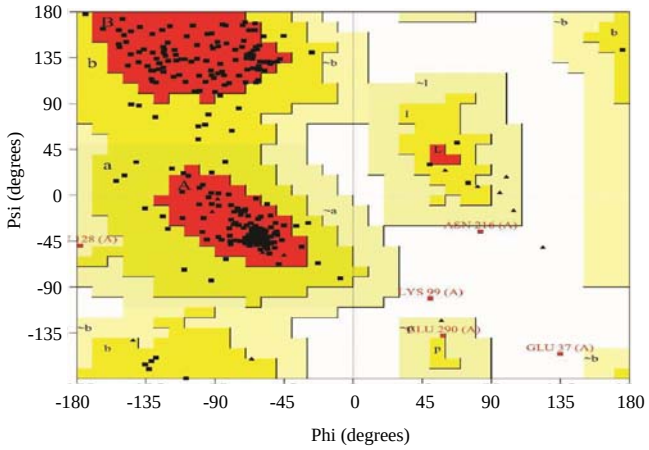

Fig. 9: PROCHECK analysis result of Api m3 protein of *Apis mellifera,* Color codes: Red color-most favorable regions, yellow color region-allowed region and pale yellow generously allowed region and white color-disallowed regions

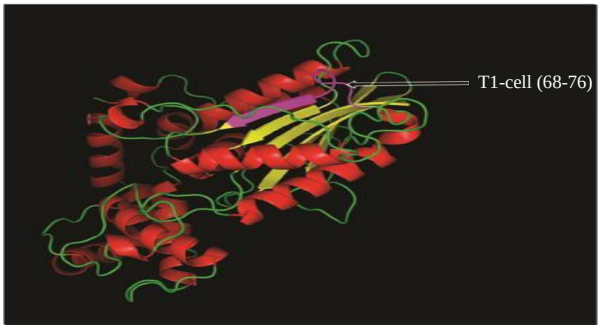

Fig. 10: MHC class I T-cell epitope is marked in magenta, which is positioned amino acid residue No. 68-76

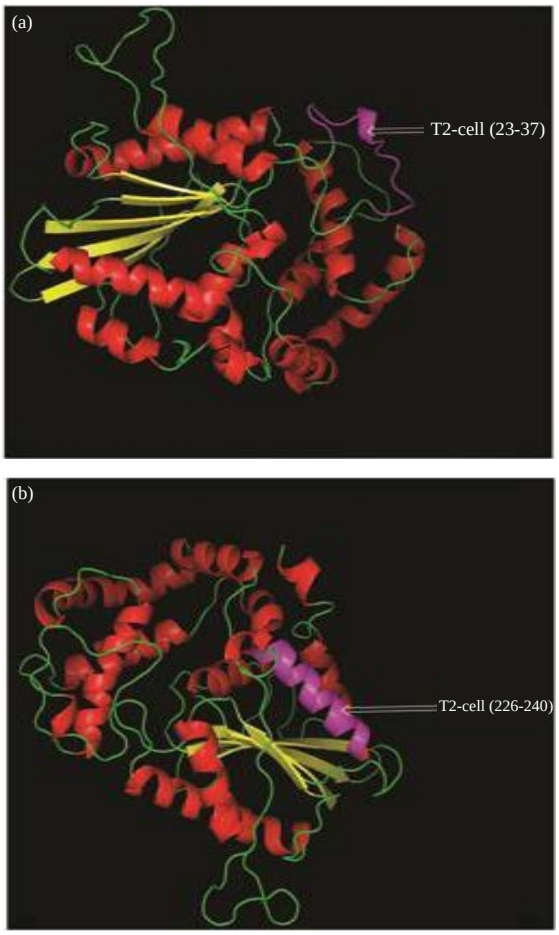

Fig. 11(a-b): Two MHC class I T-cell epitope is marked in magenta, which is positioned amino acid residue No. 23-37 and 226-240

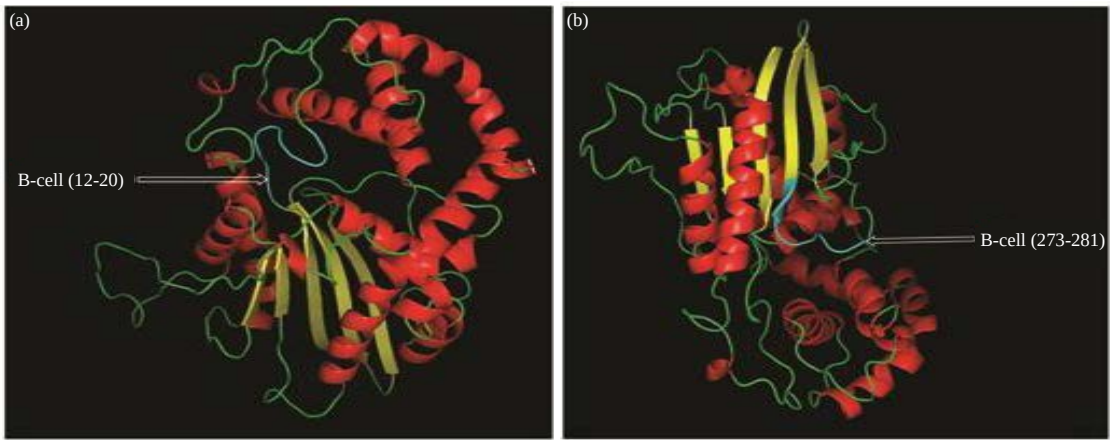

Fig. 12(a-b): B-cell epitopes are highlighted in Cyan. First, potential B-cell epitope is region 12-20 (Cyan) and another candidate is 273-281 (Cyan)

CONCLUSION

Identification of epitopes in Api m3 of honey bee is crucial for suitable vaccine assessment and development. The study predicted MHC Class I and Class II T-cell epitopes and B-cell epitopes in Api m3 *Apis mellifera*. These peptides should be assessed further for immunoreactivity through *in vivo* studies. The outcome of this study can certainly aid in designing new therapeutic modalities in Api m3 venom allergen of *Apis mellifera.*

SIGNIFICANCE STATEMENT

In this study, potential T-cell and B-cell epitopes of Api m3 allergen are reported. These predicted T-cell and B-cell epitopes of Api m3 allergen could help the researchers to test them further for immunoreactivity applying *in vivo* analysis. Still there is no report of T-cell and B-cell epitopes of *Apis mellifera*, this study can be the pioneer in finding effective vaccine against allergens of honey bee.

REFERENCES

1. Smith, D.R., L. Villafuerte, G. Otis and M.R. Palmer, 2000. Biogeography of *Apis cerana* F. and *A. nigrocincta* smith: Insights from mtDNA studies. Apidologie, 31: 265-279.
2. Radauer, C., A. Nandy, F. Ferreira, R.E. Goodman and J.N. Larsen *et al*., 2014. Update of the WHO/IUIS allergen nomenclature database based on analysis of allergen sequences. Allergy, 69: 413-419.
3. Nam, K.W., K.H. Je, J.H. Lee, H.J. Han, H.J. Lee, S.K. Kang and W. Mar, 2003. Inhibition of COX-2 activity and proinflammatory cytokines (TNF-α and IL-1β) production by water-soluble sub-fractionated parts from bee (*Apis mellifera*) venom. Arch. Pharm. Res., 26: 383-388.
4. Wiwatwongwana, D., N. Jariyapan and A. Wiwatwongwana, 2012. Eyelid bee sting with late migration onto the cornea after primary removal: The mystery of the bee stinger. Arch. Ophthalmol., 130: 392-393.
5. Ludman, S.W. and R.J. Boyle, 2015. Stinging insect allergy: Current perspectives on venom immunotherapy. J. Asthma Allergy, 8: 75-86.
6. Bilo, B.M., F. Rueff, H. Mosbech, F. Bonifazi and O.J.N.G. Elberink, 2005. Diagnosis of Hymenoptera venom allergy. Allergy, 60: 1339-1349.
7. Schiener, M., A. Graessel, M. Ollert, C.B. Schmidt-Weber and S. Blank, 2017. Allergen-specific immunotherapy of Hymenoptera venom allergy-also a matter of diagnosis. Hum. Vaccines Immunotherapeut., 13: 2467-2481.
8. Colovos, C. and T.O. Yeates, 1993. Verification of protein structures: Patterns of nonbonded atomic interactions. Prot. Sci., 2: 1511-1519.
9. Geourjon, C. and G. Deleage, 1995. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput. Applied Biosci., 11: 681-684.
10. Gill, S.C. and P.H. von Hippel, 1989. Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem., 182: 319-326.
11. Guruprasad, K., B.V.B. Reddy and M.W. Pandit, 1990. Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting *in vivo* stability of a protein from its primary sequence. Protein Eng. Des. Sel., 4: 155-161.
12. Ikai, A., 1980. Thermostability and aliphatic index of globular proteins. J. Biochem., 88: 1895-1898.
13. Larsen, M.V., C. Lundegaard, K. Lamberth, S. Buus, O. Lund and M. Nielsen, 2007. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinform., Vol. 8, No. 1. 10.1186/1471-2105-8-424
14. Buus, S., S.L. Lauemoller, P. Worning, C. Kesmir and T. Frimurer *et al*., 2003. Sensitive quantitative predictions of peptide MHC binding by a query by committee artificial neural network approach. HLA., 62: 378-384.
15. Peters, B. and A. Sette, 2005. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. BMC Bioinform., Vol. 6, No. 1. 10.1186/1471-2105-6-132.
16. Vita, R., J.A. Overton, J.A. Greenbaum, J. Ponomarenko and J.D. Clark *et al*., 2015. The immune epitope database (IEDB) 3.0. Nucleic Acids Res., 43: D405-D412.
17. Bui, H.H., J. Sidney, W. Li, N. Fusseder and A. Sette, 2007. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinform., Vol. 8. 10.1186/1471-2105-8-361.
18. Thevenet, P., Y. Shen, J. Maupetit, F. Guyon, P. Derreumaux and P. Tuffery, 2012. PEP-FOLD: An updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucl. Acids Res., 40: W288-W293.
19. Maupetit, J., P. Derreumaux and P. Tuffery, 2009. PEP-FOLD: An online resource for de novo peptide structure prediction. Nucl. Acids Res., 37: W498-W503.
20. Maupetit, J., P. Derreumaux and P. Tuffery, 2010. A fast method for large-scale de novo peptide and miniprotein structure prediction. J. Comput. Chem., 31: 726-738.
21. Fieser, T.M., J.A. Tainer, H.M. Geysen, R.A. Houghten and R.A. Lerner, 1987. Influence of protein flexibility and peptide conformation on reactivity of monoclonal anti-peptide antibodies with a protein alpha-helix. Proc. Natl. Acad. Sci., 84: 8568 -8572.

22. Zhang, W., Y. Niu, Y. Xiong, M. Zhao, R. Yu and J. Liu, 2012. Computational prediction of conformational B-cell epitopes from antigen primary structures by ensemble learning. PloS One, Vol. 7, No. 8. 10.1371/journal.pone.0043575.
23. Saha, S. and G.P.S. Raghava, 2006. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins: Struc. Func. Bioinf., 65: 40-48.
24. Larsen, J.E.P., O. Lund and M. Nielsen, 2006. Improved method for predicting linear B-cell epitopes. Immunome Res., Vol. 2. 10.1186/1745-7580-2-2.
25. Saha, S. and G.P.S. Raghava, 2004. BcePred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-Chemical Properties. In: Artificial Immune Systems, ICARIS 2004, Nicosia, G., V. Cutello, P.J. Bentley and J. Timmis (Eds.)., Lecture Notes in Computer Science, Vol. 3239. Springer, Berlin, Heidelberg, pp: 197-204.
26. Karplus, P.A. and G.E. Schulz, 1985. Prediction of chain flexibility in proteins. Naturwissenschaften, 72: 212-213.
27. Kolaskar, A.S. and P.C. Tongaonkar, 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett., 276: 172-174.
28. Emini, E.A., J.V. Hughes, D. Perlow and J. Boger, 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol., 55: 836-839.
29. Parker, J.M.R., D. Guo and R.S. Hodges, 1986. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: Correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry, 25: 5425-5432.
30. Chou, P.Y. and G.D. Fasman, 1978. Prediction of the Secondary Structure of Proteins from their Amino Acid Sequence. In: Advances in Enzymology and Related Areas of Molecular Biology, Meister, A. (Ed.). John Wiley and Sons, New York, pp: 45-148.
31. Arnold, K., L. Bordoli, J. Kopp and T. Schwede, 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22: 195-201.
32. Colovos, C. and T.O. Yeates, 1993. Verification of protein structures: Patterns of nonbonded atomic interactions. Prot. Sci., 2: 1511-1519.
33. Laskowski, R.A., M.W. MacArthur, D.S. Moss and J.M. Thornton, 1993. PROCHECK: A program to check the stereochemical quality of protein structures. J. Applied Cryst., 26: 283-291.
34. Lovell, S.C., I.W. Davis, W.B. Arendall III, P.I.W. de Bakker and J.M. Word *et al*., 2003. Structure validation by Cα geometry: Φ, ψ and Cβ deviation. Proteins: Struct. Funct. Bioinform., 50: 437-450.
35. O'Hehir, R.E., S.R. Prickett and J.M. Rolland, 2016. T cell epitope peptide therapy for allergic diseases. Curr. Allergy Asthma Rep., Vol. 16, No. 2. 10.1007/s11882-015-0587-0.
36. Pastorello, E.A., M. Monza, V. Pravettoni, R. Longhi and P. Bonara *et al*., 2010. Characterization of the T-cell epitopes of the major peach allergen Pru p 3. Int. Arch. Allergy Immunol., 153: 1-12.
37. Ravkov, E.V., I.Y. Pavlov, T.B. Martins, G.J. Gleich, L.A. Wagner, H.R. Hill and J.C. Delgado, 2013. Identification and validation of shrimp-tropomyosin specific CD4 T cell epitopes. Hum. Immunol., 74: 1542-1549.
38. Fanuel, S., S. Tabesh, E. Sadroddiny and G.A. Kardar, 2017. Analysis of predicted B and T-cell epitopes in Der p 23, allergen from *Dermatophagoides pteronyssinus*. Bioinformation, 13: 307-312.
39. Oseroff, C., J. Sidney, V. Tripple, H. Grey and R. Wood *et al*., 2012. Analysis of T cell responses to the major allergens from German cockroach: Epitope specificity and relationship to IgE production. J. Immunol., 189: 679-688.
40. Hasan, M.A., M.A. Khan, A. Datta, M.H.H. Mazumder and M.U. Hossain, 2015. A comprehensive immunoinformatics and target site study revealed the corner-stone toward Chikungunya virus treatment. Mol. Immunol., 65: 189-204.