Biotechnology Journal International
16(3): 1-23, 2016, Article no.BJI.29184
Previously known as British Biotechnology Journal
ISSN: 2231–2927, NLM ID: 101616695

SCIENCEDOMAIN *international*
www.sciencedomain.org

Insilico Prediction of T-cell Epitopes to Therapeutic Interferon -Beta (IFN-β) Protein

Swathi Krishna Reddy[1] and Venkata Bharat Kumar Pinnelli[2*]

[1]Department of Genetics, Vydehi Institute of Medical Sciences and Research Centre, #82, EPIP Area, Nallurhalli, Whitefield, Bangalore – 560066, Karnataka, India.
[2]Department of Biochemistry, Vydehi Institute of Medical Sciences and Research Centre, #82, EPIP Area, Nallurhalli, Whitefield, Bangalore – 560066, Karnataka, India.

Authors' contributions

This work was carried out in collaboration between both authors. Author SKR Took part in designing the study, wrote the protocol, performed analysis, performed statistical analysis, did literature searches and designed the first draft of the manuscript. Author VBKP designed the study, analyzed the data, performed literature searches and corrected drafts of manuscript. Both authors read and approved the final draft of the manuscript.

Article Information

DOI: 10.9734/BJI/2016/29184
Editor(s):
(1) Sukesh Voruganti, Department of Pharmaceutical Sciences, Texas Tech University Health Sciences Center, USA.
Reviewers:
(1) Arafat Rahman Oany, Mawlana Bhashani Science and Technology University, Bangladesh.
(2) Anonymous, Complutense University of Madrid, Spain.
(3) Sanjay Mishra, IFTM University, Moradabad, Uttar Pradesh, India.
Complete Peer review History: http://www.sciencedomain.org/review-history/16537

Original Research Article

Received 27th August 2016
Accepted 1st October 2016
Published 13th October 2016

ABSTRACT

Aims: Several studies have reported the existence for T helper cell epitopes with the persistence of unwanted immune reactions for several protein drugs. T-cell epitope is an amino acid or set of amino acids that are capable of being recognized form one or more T-cell receptors. There is also an indication that T helper cells are involved in the anti-drug antibodies development to therapeutic interferon beta-1a. Protein drugs containing Major histocompatibility complex class II T cell epitopes are likely to elicit anti-drug antibodies. Binding specificity between T-cell epitopes and major histocompatibility molecules are the most important determinant step in finding the T-cellular immune responses. The data obtained from the present study provides new insights into prediction of therapeutic Interferon beta T helper cells epitopes using T cell epitope prediction tools, mapping of clusters of predicted epitopes.

**Corresponding author: E-mail: pvbharatkumar@yahoo.co.in;*

Study Design: Insilico analysis by bioinformatics tools was to predict T-cell epitopes of Interferon beta-1a.
Methodology: Several Insilico prediction tools (immunoinformatics tools) including Proped, NetMHCIIpan3.0 and Immune Epitope Database Analysis Resource (IEDB-AR) are available to map the potential major histocompatibility class II T cell epitopes. After predicting potential T-cell epitopes, epitopes were mapped on interferon beta-1a using MIMOX2 server.
Results: The potential MHC class II immunogenic sequence of 50 amino acids "**TRGKLMSSLHLKRYYGRILHYLKAKEYSHCAWTIVRVEILRNFYFINRLTG"** With IFN-β-1a (position 111-161) were identified. This study can provide the understanding the relevance to T-cell activation for prediction and assessment of unwanted immune responses.
Conclusions: Insilico prediction by using the available tools helps in reducing the time and cost for the immunologists during the vaccine design. By predicting them we will come to know, which peptides play major role and synthesize them using invitro technologies.

Keywords: T-cell epitopes; IEDB; proped; MIMOX2; Interferon beta.

1. INTRODUCTION

Interferons are naturally occurring proteins which stimulate intracellular and intercellular responses that regulate resistance to viral infections enhance immune responses and modulate cell survival and death. There are three types of interferons Type I IFN (α,β,ε,κ,ω), Type II IFN, Type III IFN or IFN-λ. Interferons help in first line of defense, exhibits antiviral activities, exhibit antiviral responses and activate IFN-stimulated genes respectively [1].

Interferon beta is cylindrical in shape and belongs to family of long-chain helical cytokines. Interferon beta exhibits potent T-cell anti-proliferative properties. A glycosylation site exists at residue aspargine-80. It contains 166 amino acids residues which comprises of five α helices designated as helix A (amino acid residues 2-22), helix B (amino acid residues 51-71), helix C (amino acid 80-107), helix D(amino acid 118-136) and helix E (amino acid 139-162). Several studies suggest that the glycosylation of IFN-β plays a critical role in protein solubility and stability. Non-glycosylated forms of IFN-β are more susceptible to form aggregates and lead to immunogenicity. The structural-functional analysis helps in neutralizing antibody binding sites, identification of T-cell and B-cell antigenic determinants [2]. Functions involves, interferons acts as immunomodulators, antineoplastic, activate immune cells such as NK cells, macrophage and increases host defense by up regulating antigen resenting (AP) by virtue of increase the expression of MHC alleles. Applications of IFN-β involves in the diagnosis of rheumatoidal arthritis, Multiple sclerosis (MS), auto immunological diseases. Advances in the antigen processing and presentation has made the subunit vaccines an integral part of vaccine design. In the subunit vaccine, the vaccine candidates are used as immunogenic peptides/regions of protein instead of complete protein. Binding of this peptide to the Major histocompatibility molecules play vital role in the activation of autogenic specific T- cells and hence finding these peptides or epitopes is important [3].

Insilico prediction of T-cell epitopes can helpful in identification of peptides using computational programs which have in designing the vaccine development. Some of the methods include involves quantitative matrices method and the other is position specific binding profiles. Major histocompatibility class II molecules helps in identification of the foreign molecule on the membrane surface and degrade the molecule into peptides and present them to T-cell receptors for particular immunological responses [4]. An Immune epitope database (IEDB) is the one of the important server to find the immune related information including antigens, anti-antibody responses, T and B cell information. Information to this server is retrieved from TEPITOPE; NetMHCIIpan [5]. User can select the prediction for HLA-DP, DQ, DR molecules with suitable parameters and threshold. Proped is theserver which uses the panspecific method to identify the predicted binders. The Pan specific method helps in providing the predictions for those molecules which are previously uncharacterized Major histocompatibility molecules [6]. Pan specific methodology play an important role, as they are capable of giving predictions to those molecules, which have not been characterized experimentally [7]. *NetMHCIIpan-3.0* method is used to cluster the most prevalent HLA alleles of the European population. For HLA class I, clustering of molecules into super types was proposed by the

analysis carried out using experimental data [8] and extending by applying pan–specific class I prediction analysis. Whereas for MHC class II, the amount of experimental data remains too limited and hence they have been limited to HLA-DR molecules. The analysis performed here suggesting reduction of polymorphism of HLA class II molecules by definition of clusters based on similarities in predicted functional binding specificities. Such clustering builds a base for facilitating identification of T helper cell epitopes within different ethnic groups having a high value in the design of epitope-based vaccines. There are many methods identified for the prediction of peptide-major histocompatibility binding including artificial neural network, quantitative matrices method, and Markova model. For major histocompatibility class I, the peptides which bind to Major histocompatibility class I are of same length and hence they are easily characterized. Whereas binding of major histocompatibility class II binding is very different because of they are distributed all over the length of natural major histocompatibility binding peptides [9] and hence they are identified. Interferon beta triggers the synthesis of host cell proteins contributing to it's to its several immunomodulatory properties. It also increases the expression of HLA class I molecules and this might contribute to antiviral effects. It inhibits HLA class II molecules induced from IFN- γ and so inhibits the antigen presentation [10]. Application of predicting the T-cell epitopes intereferon-beta-1a include: Interferon beta -1a epitopes it helps the treatment of auto immune diseases, T-cell epitope prediction are used in the treatment of cancer, auto immunity, allergy and infectious diseases, There are many application of T-cell epitope prediction in case of disease diagnosis and vaccine design which include Mapping in Foot and mouth disease, Mapping in Type 1 Diabetes,Mapping in pneumonia, Mapping in HIV.

2. MATERIALS AND METHODOLOGY

Methodology helps in giving step by step procedure of the work done in the study. Below mentioned is the work flow involved in the study (Fig. 1).

2.1 Protein Sequence Retrieval

Interferon beta-1a protein FASTA sequence was retrieved from drug bank having the accession number DB00060 with the length of 166 amino acids (http://www.drugbank.ca/drugs/) this protein sequence is basis to perform different computational predictions of linear amino acid residues.

2.2 Potential T-Cell Epitope Prediction

T-cell epitope is an amino acid or set of amino acids capable of being recognized from one or more T-cell receptors. Cells recognize the linear peptides that bound to the MHC class II molecules. Understanding the relevance of T-cell activation in antidrug antibodies play a crucial role for predicting and assessing immunogenicity. To check whether interferon beta can elicit T-cell responses, we predicted potential T-cell epitopes using IEDB analysis resource. IEDB-AR follows allele specific method for prediction. For our prediction we fallowed IEDB recommended prediction method choosing the all the HLA-DR alleles from the panel. Later from the predicted T-cell epitopes only those epitopes having low percentile rank were chosen, since lower the percentile rank indicates the good binders [11].

In IEDB 15mer amino acids MHC class II T-cell epitope prediction was performed using the NetMHCIIpan method, consensus method [12], Average Relative Binding (ARB) matrix method and stabilization matrix alignment method (SMM).

Proped is a server to predict major histocompatibility class II binding peptide prediction which can predict major histocompatibility class II binding regions in an antigen sequence [13], method is based on the quantitative matrix method. For our prediction panspecific method is used for the identification of the predicted affinity using 51 alleles of Human leukocyte antigen-DR alleles. The Pan specific method covers all the binding information of the molecule from the different loci or species. Later 41 patterns (overlapping) with their regions are identified.

NETMHCII pan 3.0 is a server capable of predicting the T-cell epitopes. This server either use allele specific method or pan specific method based on the user requirement. We adopted allele specific method and we have done the prediction for 12 alleles of Human leukocyte antigens-DR [14]. By 15-mer amino acid analysis of Major histocompatibility class II T-cell epitope prediction was performed and the prediction is based on the affinity i.e. if the affinity is <500 nm then it indicates weak binder, if the affinity is <

50nm then it indicates strong binder. In the current study we adopted the method based on the principle of artificial neural networks and it is trained on 56,062 quantitative peptide binding data which covers 12HLA alleles as well.

2.3 Epitope Mapping of IFN-β

Once the T-cell peptides are identified by major histocompatibility class II binding prediction servers (IEDB-AR, Proped, NetMHCIIpan) they need to mapped and this is done through both in vivo and invitrotechniques. In this study we followed in vivo method of epitope mapping using MIMOX2 server which adopts mapping of cluster of epitopes based on the phage display method and which is done manually. In MIMOX mapping is done based on the epitope of an antibody of one or more user supplied mimitopes with suitable antigen structure and this helps in the computational immunovaccinology or computer aided vaccine design. Mimitopes are the peptides which mimic the structure of epitopes [15].

Epitope mapping through MIMOX server involves two steps, first step include obtain the potential MHC class II epitopes on to the MIMOX2 server. Second step set the parameters since it is online based server and fix the threshold and upload the monomer structure of interferon beta structure downloaded from protein databank for matching residues and place them into an stack of candidate residues with their respective positions. Epitope mapping is based on the Mimitopes i.e. mimitopes are the peptides mimicking the protein, lipids, carbohydrates that can be generated by phage display technology. Principle behind involves MIMOX2, MIMOX2 can map epitopes as an individual mimitopes or group of patterns on the selected antigenic structure and cluster of residues searching is done which represents the naïve epitopes and hence mapping of epitopes is based on input sequence and uploaded antigen structure. For mapping different modes are available first strict mode where the mimitope residues matches with the antigen residues, second conservative mode, identifies similar residues which are involved in the stack. MIMOX2 server helps in identification of neighbor candidates with their respective distances from one peptide to other. The distances between the neighbors are calculated by the threshold value (distance between all c-alpha and c-beta atoms).

2.4 Visualization of MHC Binding Motifs

Once after predicting the MHC binding motifs it is important to view those motifs (patterns). One such method involves visualizing the receptor binding motif is by using sequence2logos. Tom Schneider and Mike Stephens for the first time found seq2logos [16].

Graphical form of representing different amino acids or nucleic acids are called as seqlogos .This information tells us which amino acids are highly conserved in nature, this is interpreted by looking at the height of the all the twenty amino acids distributed in the graphical plot. The higher the amino acid in the column more conserved the amino acid. This seqlogos are generated using MHC motif viewer. We can found the seqlogos of Human leukocyte antigen-DP, DQ, DR molecules. To differentiate between amino acids colour coding is done based on the properties of individual amino acids, basic amino acids, hydrophobic amino acids are represented in red, blue and black colours respectively.

2.5 Comparative Studies

Comparative studies help in finding the pros and cons of the methodologies used in the study. From the literature we have identified few differences between experimental prediction of T-cell and Insilico prediction of T-cell epitopes. First Experimental prediction is also called as invitro methodology, which involves the longer methodology, with more number of allelic variants with high accuracy but predicted epitopes cannot view immediately. Insilico prediction uses computational methodology, which involves the short procedures i.e. number of steps of methodology is less, involves less number of allele variants with comparatively less accuracy but epitopes can be viewed after the predictions is done.

Table 1. Table show the list of servers and suitable methods followed in the study

Server	Method	URL
IEDB-AR	ANN	http://tools.immuneepitope.org/mhcii.
PROPED	QM	http://www.imtech.res.in/raghava/propred/
NetMHCIIpan-3.0	ANN	http://www.cbs.dtu.dk/services/NetMHCIIpan-3.0.

3. RESULTS AND DISCUSSION

3.1 Identification of Interferon-beta-1a T-cell Epitopes Using Insilico Approach

3.1.1 IEDB-AR server results

Insilico technologies are nothing but using the computational algorithms to identify variant sequences that exhibits desired functional properties [17]. The computer algorithms can also identify immunogenic T-cell epitopes. In our study Insilico predictions were used to identify MHC class II IFN-β-1a T-cell epitopes using IEDB-AR and Proped. The screenshots Fig. 1 represent the MHC class II binding of HLA-DR alleles for IFN-β-1a. However the assessment of all the available alleles in respective Insilico prediction tools was carried out systematically. In IEDB-AR low percentile rank indicates good binders, the analysis done for all the 696 alleles of HLA-DR alleles. The potential MHC class II immunogenic sequence of 50 amino acids, "**TRGKLMSSLHLKRYYGRILHYLKAKEYSHCA WTIVRVEILRNFYFINRLTG"** with IFN-β-1a (position 111-161) were identified.

Training data is used in the server consists ofquantitative peptide-MHC class II binding data. It comprises 52,062 affinity measurements covering 24 HLA-DR, 5HLA-DP, 6 HLA-DQ, and 2 murine H-2 molecules. Additionally, a set of 9860 binding affinity measurements covering 13 HLA-DR alleles introduced was used as an independent evaluation set [18]. This is based on the percentile rank which assess lower the percentile rank higher the binding.

3.2 Proped MHC Class II Binding Prediction Results for IFN-β-1a

In proped panspecific method is to determine binding patterns where it considers all the 51 alleles of HLA-DR alleles. The data in the blue colour amino acids indicate good binders whereas the red colour amino acids suggest the possible promiscuous amino acids. From proped 41 patterns were identified. The potential MHC class II immunogenic sequence of 50 amino acids "**TRGKLMSSLHLKRYYGRILHYLKAKEYSHCA WTIVRVEILRNFYFINRLT"** With IFN-β-1a (position 111-161) were identified.

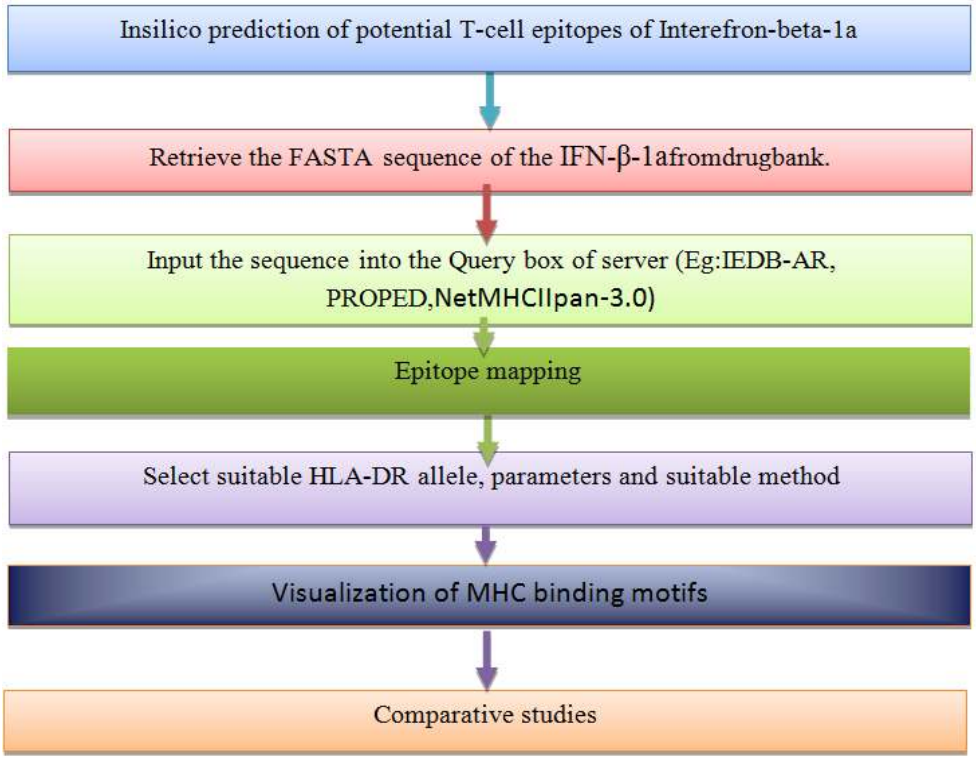


Fig. 1. Workflow involved in the study which give us the information about how insilico prediction, mapping, visualization has been performed

3.2.1 Proped server results

This sever works based on the algorithm i.e. quantitative matrix based predictions. The algorithm works as follows: First find all the possible "Nanomeric Peptides" from the interferon beta sequence. In the second step the position and side chain specific values, derived from virtual matrices are assigned to each residue within peptide frames. In Third step involves the linear combination i.e. summation of all the position and side chain specific values result in a numeric value/peptide frame score for each peptide. Finally all the calculated peptide scores are compared with user selected threshold value and the results are indicated graphical as well HTML format.

Table 2. Screenshots showing the IEDB-AR text results for Interferon β-1a proteinfor different HLA-DR alleles which also show the predicted epitope start and ending point and they adopted different methods like sturniolo and consensus method. The consensus method is based on the combinatorial library method

IEDB RESULTS

Figure 1:

Check to expanded the result:

Allele	#	Start	End	Peptide	Method used	Percentile rank
HLA-DRB1*01:01	1	57	71	LTIYEMLQNIFAIFR	Consensus (comb.lib./smm/nn)	2.51
HLA-DRB1*01:01	1	56	70	ALTIYEMLQNIFAIF	Consensus (comb.lib./smm/nn)	2.74
HLA-DRB1*01:01	1	55	69	AALTIYEMLQNIFAI	Consensus (comb.lib./smm/nn)	3.02

Download result

Allele	#	Start	End	Peptide	Method used	Percentile rank
HLA-DRB1*03:01	1	113	127	RGKLMSSLHLKRYYG	Consensus (smm/nn/sturniolo)	0.81
HLA-DRB1*03:01	1	112	126	TRGKLMSSLHLKRYY	Consensus (smm/nn/sturniolo)	0.84
HLA-DRB1*03:01	1	34	48	DRMNFDIPEEIKQLQ	Consensus (smm/nn/sturniolo)	1.34

Allele	#	Start	End	Peptide	Method used	Percentile rank
HLA-DRB1*04:01	1	64	78	QNIFAIFRQDSSSTG	Consensus (smm/nn/sturniolo)	0.54
HLA-DRB1*04:01	1	65	79	NIFAIFRQDSSSTGW	Consensus (smm/nn/sturniolo)	0.54
HLA-DRB1*04:01	1	66	80	IFAIFRQDSSSTGWN	Consensus (smm/nn/sturniolo)	0.54

Allele	#	Start	End	Peptide	Method used	Percentile rank
HLA-DRB1*07:01	1	140	154	HCAWTIVRVEILRNF	Consensus (comb.lib./smm/nn)	1.84
HLA-DRB1*07:01	1	141	155	CAWTIVRVEILRNFY	Consensus (comb.lib./smm/nn)	1.84
HLA-DRB1*07:01	1	137	151	EYSHCAWTIVRVEIL	Consensus (comb.lib./smm/nn)	1.86

Allele	#	Start	End	Peptide	Method used	Percentile rank
HLA-DRB1*08:01	1	123	137	KRYYGRILHYLKAKE	sturniolo	0.10
HLA-DRB1*08:01	1	124	138	RYYGRILHYLKAKEY	sturniolo	0.10
HLA-DRB1*08:01	1	125	139	YYGRILHYLKAKEYS	sturniolo	0.10

INPUT & PARAMETER INFORMATION

Antigen Name	test
Scanned on	Tue Mar 24 16:08:34 2015
Length of input sequence	166 amino acids
Number of nanomers from input sequence	158
Number of nanomers with obligatory P1 anchor residue	63
Threshold setting	3
Number of alleles in query	51


Predicted epitopes
Predicted epitopes

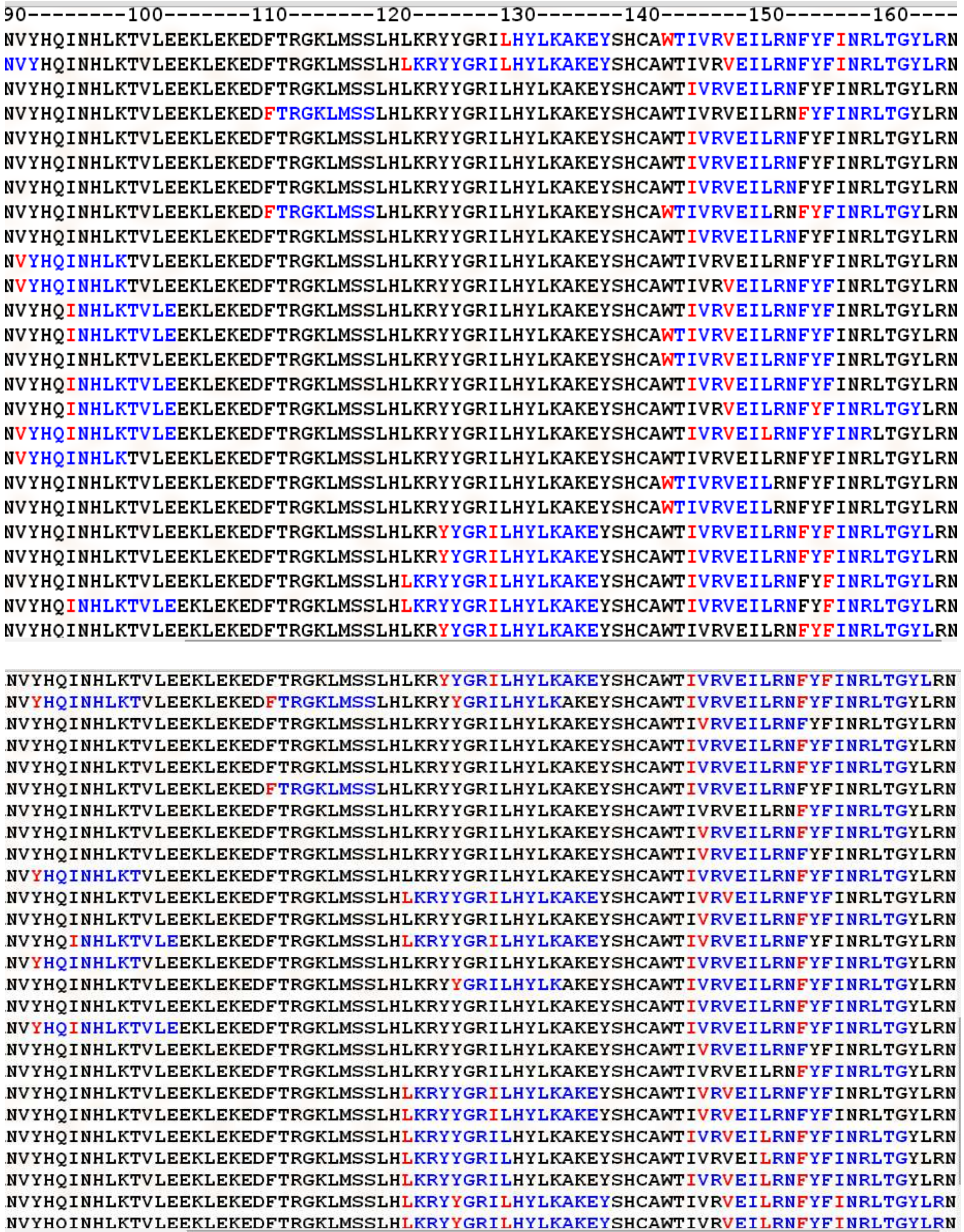


Fig. 2. Screenshot of the peptide MHC class II binding prediction for interferon-β-1a protein that demonstrate the pan-specific methods to provide accurate predictions of MHC molecules for 51 alleles.The blue colour amino acids indicate good binders whereas the red colour amino acids suggest the possible promiscuous amino acids

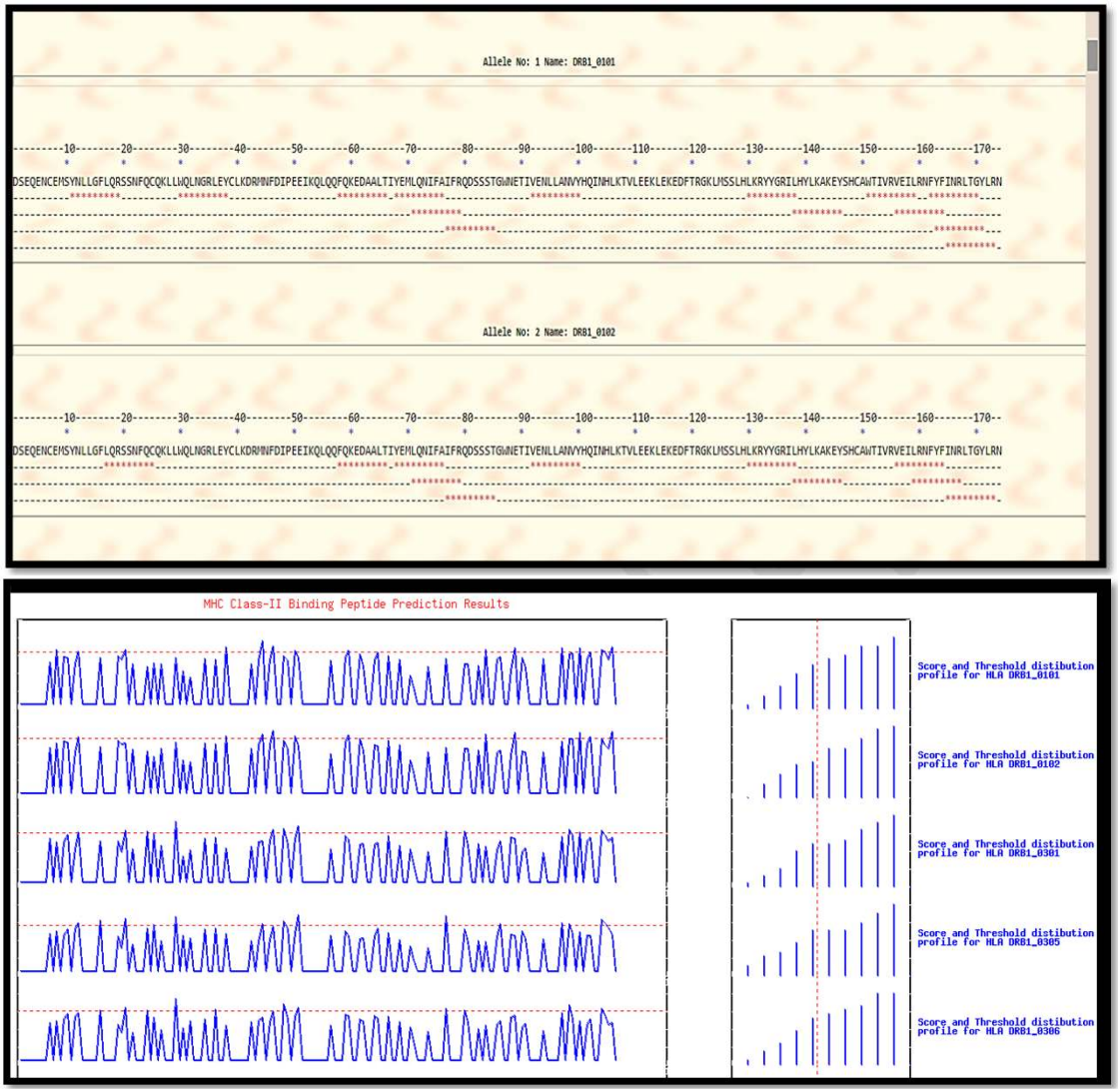


Fig. 3. Screenshots from the proped server showing graphical and text output forms for different interpretation

3.3 Results of NETMHCII Pan 3.0

NETMHCII pan 3.0 is the first pan specific method to determine the binding patterns. It is applied to predict binding for uncharacterized MHC molecules. It consists of 12 alleles. For 12 HLA-DR alleles the binding predictions are calculated and based on the pan specific method and the predicted binders (affinity) are differentiated into strong binders (affinity<50 nm) and weak binders (affinity<500 nm). 15 mer amino acids prediction were chosen. To increase the performance of the NetMHCIIpan-3.0 method demonstrates its promising ability to improve when more data available for molecules from other loci/species hence this server is chosen for the prediction.

3.4 Predicted Results

After obtaining the results from the different servers, predicting the results from the server are important to conclude the study.

H LA-A*01:01

Allele: DRB1*01:01

pos	Allele	peptide	Identity	Pos	Core	1-log50k(aff)	Affinity(nM)	%Rank	BindingLevel
0	DRB1*01:01	MSYNLLGFLQRSSNF	Sequence	1	SYNLLGFLQ	0.594	80.55	32.00	<=WB
1	DRB1*01:01	SYNLLGFLQRSSNFQ	Sequence	3	LLGFLQRSS	0.615	64.39	32.00	<=WB
2	DRB1*01:01	YNLLGFLQRSSNFQC	Sequence	5	FLQRSSNFQ	0.606	71.04	32.00	<=WB
3	DRB1*01:01	NLLGFLQRSSNFQCQ	Sequence	4	FLQRSSNFQ	0.625	57.63	32.00	<=WB
4	DRB1*01:01	LLGFLQRSSNFQCQK	Sequence	3	FLQRSSNFQ	0.643	47.50	16.00	<=SB
5	DRB1*01:01	LGFLQRSSNFQCQKL	Sequence	2	FLQRSSNFQ	0.599	76.75	32.00	<=WB
6	DRB1*01:01	GFLQRSSNFQCQKLL	Sequence	1	FLQRSSNFQ	0.536	151.48	50.00	<=WB
7	DRB1*01:01	FLQRSSNFQCQKLLW	Sequence	0	FLQRSSNFQ	0.462	339.13	50.00	<=WB
8	DRB1*01:01	LQRSSNFQCQKLLWQ	Sequence	6	FQCQKLLWQ	0.425	504.33	50.00	
9	DRB1*01:01	QRSSNFQCQKLLWQL	Sequence	5	FQCQKLLWQ	0.462	337.64	50.00	<=WB
10	DRB1*01:01	RSSNFQCQKLLWQLN	Sequence	4	FQCQKLLWQ	0.474	297.53	50.00	<=WB
11	DRB1*01:01	SSNFQCQKLLWQLNG	Sequence	3	FQCQKLLWQ	0.469	311.81	50.00	<=WB
12	DRB1*01:01	SNFQCQKLLWQLNGR	Sequence	2	FQCQKLLWQ	0.479	279.22	50.00	<=WB
13	DRB1*01:01	NFQCQKLLWQLNGRL	Sequence	6	LLWQLNGRL	0.663	38.20	15.00	<=SB
14	DRB1*01:01	FQCQKLLWQLNGRLE	Sequence	5	LLWQLNGRL	0.693	27.67	10.00	<=SB
15	DRB1*01:01	QCQKLLWQLNGRLEY	Sequence	4	LLWQLNGRL	0.740	16.73	5.00	<=SB
16	DRB1*01:01	CQKLLWQLNGRLEYC	Sequence	3	LLWQLNGRL	0.756	14.05	4.00	<=SB
17	DRB1*01:01	QKLLWQLNGRLEYCL	Sequence	2	LLWQLNGRL	0.741	16.47	5.00	<=SB
18	DRB1*01:01	KLLWQLNGRLEYCLK	Sequence	1	LLWQLNGRL	0.703	24.82	9.00	<=SB
19	DRB1*01:01	LLWQLNGRLEYCLKD	Sequence	0	LLWQLNGRL	0.640	49.24	32.00	<=SB
20	DRB1*01:01	LWQLNGRLEYCLKDR	Sequence	0	LWQLNGRLE	0.455	363.57	50.00	<=WB
21	DRB1*01:01	WQLNGRLEYCLKDRM	Sequence	2	LNGRLEYCL	0.341	1245.40	50.00	
22	DRB1*01:01	QLNGRLEYCLKDRMN	Sequence	1	LNGRLEYCL	0.316	1643.55	50.00	
23	DRB1*01:01	LNGRLEYCLKDRMNF	Sequence	4	LEYCLKDRM	0.339	1271.85	50.00	
24	DRB1*01:01	NGRLEYCLKDRMNFD	Sequence	3	LEYCLKDRM	0.325	1491.08	50.00	
25	DRB1*01:01	GRLEYCLKDRMNFDI	Sequence	3	EYCLKDRMN	0.351	1124.72	50.00	
26	DRB1*01:01	RLEYCLKDRMNFDIP	Sequence	3	YCLKDRMNF	0.349	1141.90	50.00	
27	DRB1*01:01	LEYCLKDRMNFDIPE	Sequence	2	YCLKDRMNF	0.320	1563.46	50.00	
28	DRB1*01:01	EYCLKDRMNFDIPEE	Sequence	1	YCLKDRMNF	0.281	2401.05	50.00	
29	DRB1*01:01	YCLKDRMNFDIPEEI	Sequence	0	YCLKDRMNF	0.280	2414.55	50.00	
30	DRB1*01:01	CLKDRMNFDIPEEIK	Sequence	5	MNFDIPEEI	0.263	2912.01	50.00	
31	DRB1*01:01	LKDRMNFDIPEEIKQ	Sequence	4	MNFDIPEEI	0.290	2167.66	50.00	
32	DRB1*01:01	KDRMNFDIPEEIKQL	Sequence	3	MNFDIPEEI	0.334	1343.06	50.00	
33	DRB1*01:01	DRMNFDIPEEIKQLQ	Sequence	2	MNFDIPEEI	0.319	1580.32	50.00	
34	DRB1*01:01	RMNFDIPEEIKQLQQ	Sequence	1	MNFDIPEEI	0.323	1517.27	50.00	
35	DRB1*01:01	MNFDIPEEIKQLQQF	Sequence	0	MNFDIPEEI	0.299	1963.02	50.00	
36	DRB1*01:01	NFDIPEEIKQLQQFQ	Sequence	2	DIPEEIKQL	0.267	2777.68	50.00	
37	DRB1*01:01	FDIPEEIKQLQQFQK	Sequence	6	IKQLQQFQK	0.315	1663.64	50.00	
38	DRB1*01:01	DIPEEIKQLQQFQKE	Sequence	5	IKQLQQFQK	0.320	1562.51	50.00	
39	DRB1*01:01	IPEEIKQLQQFQKED	Sequence	4	IKQLQQFQK	0.316	1631.84	50.00	
40	DRB1*01:01	PEEIKQLQQFQKEDA	Sequence	3	IKQLQQFQK	0.331	1386.98	50.00	
41	DRB1*01:01	EEIKQLQQFQKEDAA	Sequence	2	IKQLQQFQK	0.318	1606.59	50.00	
42	DRB1*01:01	EIKQLQQFQKEDAAL	Sequence	6	QFQKEDAAL	0.389	746.90	50.00	
43	DRB1*01:01	IKQLQQFQKEDAALT	Sequence	6	FQKEDAALT	0.605	71.87	32.00	<=WB
44	DRB1*01:01	KQLQQFQKEDAALTI	Sequence	5	FQKEDAALT	0.690	28.60	10.00	<=SB
45	DRB1*01:01	QLQQFQKEDAALTIY	Sequence	4	FQKEDAALT	0.752	14.63	5.00	<=SB
46	DRB1*01:01	LQQFQKEDAALTIYE	Sequence	3	FQKEDAALT	0.762	13.07	4.00	<=SB
47	DRB1*01:01	QQFQKEDAALTIYEM	Sequence	2	FQKEDAALT	0.709	23.43	8.00	<=SB
48	DRB1*01:01	QFQKEDAALTIYEML	Sequence	1	FQKEDAALT	0.662	38.66	15.00	<=SB
49	DRB1*01:01	FQKEDAALTIYEMLQ	Sequence	0	FQKEDAALT	0.551	128.98	32.00	<=WB
50	DRB1*01:01	QKEDAALTIYEMLQN	Sequence	3	DAALTIYEM	0.433	462.68	50.00	<=WB
51	DRB1*01:01	KEDAALTIYEMLQNI	Sequence	2	DAALTIYEM	0.447	397.22	50.00	<=WB
52	DRB1*01:01	EDAALTIYEMLQNIF	Sequence	6	IYEMLQNIF	0.518	184.41	50.00	<=WB
53	DRB1*01:01	DAALTIYEMLQNIFA	Sequence	6	YEMLQNIFA	0.722	20.32	7.00	<=SB
54	DRB1*01:01	AALTIYEMLQNIFAI	Sequence	5	YEMLQNIFA	0.763	12.94	4.00	<=SB
55	DRB1*01:01	ALTIYEMLQNIFAIF	Sequence	4	YEMLQNIFA	0.790	9.74	2.00	<=SB
56	DRB1*01:01	LTIYEMLQNIFAIFR	Sequence	3	YEMLQNIFA	0.799	8.83	2.00	<=SB
57	DRB1*01:01	TIYEMLQNIFAIFRQ	Sequence	2	YEMLQNIFA	0.778	11.03	3.00	<=SB
58	DRB1*01:01	IYEMLQNIFAIFRQD	Sequence	1	YEMLQNIFA	0.749	15.06	5.00	<=SB
59	DRB1*01:01	YEMLQNIFAIFRQDS	Sequence	0	YEMLQNIFA	0.691	28.32	10.00	<=SB
60	DRB1*01:01	EMLQNIFAIFRQDSS	Sequence	2	LQNIFAIFR	0.505	212.50	50.00	<=WB
61	DRB1*01:01	MLQNIFAIFRQDSSS	Sequence	1	LQNIFAIFR	0.493	240.20	50.00	<=WB
62	DRB1*01:01	LQNIFAIFRQDSSST	Sequence	6	IFRQDSSST	0.530	162.27	50.00	<=WB
63	DRB1*01:01	QNIFAIFRQDSSSTG	Sequence	5	IFRQDSSST	0.560	117.44	32.00	<=WB
64	DRB1*01:01	NIFAIFRQDSSSTGW	Sequence	4	IFRQDSSST	0.609	68.83	32.00	<=WB

63	DRB1*01:01	QNIFAIFRQDSSSTG	Sequence	5	IFRQDSSST	0.560	117.44	32.00	<=WB
64	DRB1*01:01	NIFAIFRQDSSSTGW	Sequence	4	IFRQDSSST	0.609	68.83	32.00	<=WB
65	DRB1*01:01	IFAIFRQDSSSTGWN	Sequence	3	IFRQDSSST	0.644	46.90	16.00	<=SB
66	DRB1*01:01	FAIFRQDSSSTGWNE	Sequence	2	IFRQDSSST	0.598	77.32	32.00	<=WB
67	DRB1*01:01	AIFRQDSSSTGWNET	Sequence	1	IFRQDSSST	0.532	158.42	50.00	<=WB
68	DRB1*01:01	IFRQDSSSTGWNETI	Sequence	0	IFRQDSSST	0.461	342.54	50.00	<=WB
69	DRB1*01:01	FRQDSSSTGWNETIV	Sequence	0	FRQDSSSTG	0.272	2627.82	50.00	
70	DRB1*01:01	RQDSSSTGWNETIVE	Sequence	4	SSTGWNETI	0.166	8286.52	50.00	
71	DRB1*01:01	QDSSSTGWNETIVEN	Sequence	3	SSTGWNETI	0.172	7790.57	50.00	
72	DRB1*01:01	DSSSTGWNETIVENL	Sequence	6	WNETIVENL	0.239	3752.48	50.00	
73	DRB1*01:01	SSSTGWNETIVENLL	Sequence	5	WNETIVENL	0.308	1787.29	50.00	
74	DRB1*01:01	SSTGWNETIVENLLA	Sequence	4	WNETIVENL	0.366	955.50	50.00	
75	DRB1*01:01	STGWNETIVENLLAN	Sequence	3	WNETIVENL	0.402	643.97	50.00	
76	DRB1*01:01	TGWNETIVENLLANV	Sequence	6	IVENLLANV	0.461	342.59	50.00	<=WB
77	DRB1*01:01	GWNETIVENLLANVY	Sequence	5	IVENLLANV	0.525	170.57	50.00	<=WB
78	DRB1*01:01	WNETIVENLLANVYH	Sequence	5	VENLLANVY	0.559	118.11	32.00	<=WB
79	DRB1*01:01	NETIVENLLANVYHQ	Sequence	3	IVENLLANV	0.600	76.09	32.00	<=WB
80	DRB1*01:01	ETIVENLLANVYHQI	Sequence	3	VENLLANVY	0.628	56.01	32.00	<=WB
81	DRB1*01:01	TIVENLLANVYHQIN	Sequence	2	VENLLANVY	0.611	67.28	32.00	<=WB
82	DRB1*01:01	IVENLLANVYHQINH	Sequence	4	LLANVYHQI	0.593	81.54	32.00	<=WB
83	DRB1*01:01	VENLLANVYHQINHL	Sequence	3	LLANVYHQI	0.585	88.78	32.00	<=WB
84	DRB1*01:01	ENLLANVYHQINHLK	Sequence	5	NVYHQINHL	0.609	68.64	32.00	<=WB
85	DRB1*01:01	NLLANVYHQINHLKT	Sequence	5	VYHQINHLK	0.666	37.19	15.00	<=SB
86	DRB1*01:01	LLANVYHQINHLKTV	Sequence	4	VYHQINHLK	0.721	20.36	7.00	<=SB
87	DRB1*01:01	LANVYHQINHLKTVL	Sequence	3	VYHQINHLK	0.769	12.15	4.00	<=SB
88	DRB1*01:01	ANVYHQINHLKTVLE	Sequence	3	YHQINHLKT	0.764	12.81	4.00	<=SB
89	DRB1*01:01	NVYHQINHLKTVLEE	Sequence	2	YHQINHLKT	0.739	16.80	6.00	<=SB
90	DRB1*01:01	VYHQINHLKTVLEEK	Sequence	4	INHLKTVLE	0.731	18.41	6.00	<=SB
91	DRB1*01:01	YHQINHLKTVLEEKL	Sequence	3	INHLKTVLE	0.716	21.49	7.00	<=SB
92	DRB1*01:01	HQINHLKTVLEEKLE	Sequence	2	INHLKTVLE	0.659	39.88	15.00	<=SB
93	DRB1*01:01	QINHLKTVLEEKLEK	Sequence	1	INHLKTVLE	0.612	66.79	32.00	<=WB
94	DRB1*01:01	INHLKTVLEEKLEKE	Sequence	0	INHLKTVLE	0.495	235.21	50.00	<=WB
95	DRB1*01:01	NHLKTVLEEKLEKED	Sequence	2	LKTVLEEKL	0.306	1821.24	50.00	
96	DRB1*01:01	HLKTVLEEKLEKEDF	Sequence	1	LKTVLEEKL	0.274	2591.45	50.00	
97	DRB1*01:01	LKTVLEEKLEKEDFT	Sequence	0	LKTVLEEKL	0.243	3622.12	50.00	
98	DRB1*01:01	KTVLEEKLEKEDFTR	Sequence	2	VLEEKLEKE	0.137	11344.57	50.00	
99	DRB1*01:01	TVLEEKLEKEDFTRG	Sequence	1	VLEEKLEKE	0.136	11517.73	50.00	
101	DRB1*01:01	LEEKLEKEDFTRGKL	Sequence	3	KLEKEDFTR	0.173	7712.33	50.00	
102	DRB1*01:01	EEKLEKEDFTRGKLM	Sequence	6	EDFTRGKLM	0.267	2774.92	50.00	
103	DRB1*01:01	EKLEKEDFTRGKLMS	Sequence	5	EDFTRGKLM	0.312	1707.46	50.00	
104	DRB1*01:01	KLEKEDFTRGKLMSS	Sequence	6	FTRGKLMSS	0.464	330.79	50.00	<=WB
105	DRB1*01:01	LEKEDFTRGKLMSSL	Sequence	5	FTRGKLMSS	0.551	129.13	32.00	<=WB
106	DRB1*01:01	EKEDFTRGKLMSSLH	Sequence	4	FTRGKLMSS	0.595	79.97	32.00	<=WB
107	DRB1*01:01	KEDFTRGKLMSSLHL	Sequence	6	GKLMSSLHL	0.712	22.63	8.00	<=SB
108	DRB1*01:01	EDFTRGKLMSSLHLK	Sequence	5	GKLMSSLHL	0.726	19.39	7.00	<=SB
109	DRB1*01:01	DFTRGKLMSSLHLKR	Sequence	4	GKLMSSLHL	0.758	13.68	4.00	<=SB
110	DRB1*01:01	FTRGKLMSSLHLKRY	Sequence	3	GKLMSSLHL	0.780	10.81	3.00	<=SB
111	DRB1*01:01	TRGKLMSSLHLKRYY	Sequence	2	GKLMSSLHL	0.759	13.57	4.00	<=SB
112	DRB1*01:01	RGKLMSSLHLKRYYG	Sequence	1	GKLMSSLHL	0.730	18.57	6.00	<=SB
113	DRB1*01:01	GKLMSSLHLKRYYGR	Sequence	0	GKLMSSLHL	0.668	36.14	15.00	<=SB
114	DRB1*01:01	KLMSSLHLKRYYGRI	Sequence	1	LMSSLHLKR	0.542	141.80	50.00	<=WB
115	DRB1*01:01	LMSSLHLKRYYGRIL	Sequence	6	LKRYYGRIL	0.550	130.02	50.00	<=WB
116	DRB1*01:01	MSSLHLKRYYGRILH	Sequence	5	LKRYYGRIL	0.550	130.33	50.00	<=WB
117	DRB1*01:01	SSLHLKRYYGRILHY	Sequence	4	LKRYYGRIL	0.584	89.86	32.00	<=WB
118	DRB1*01:01	SLHLKRYYGRILHYL	Sequence	3	LKRYYGRIL	0.642	48.34	32.00	<=SB
119	DRB1*01:01	LHLKRYYGRILHYLK	Sequence	6	YGRILHYLK	0.694	27.45	10.00	<=SB
120	DRB1*01:01	HLKRYYGRILHYLKA	Sequence	5	YGRILHYLK	0.713	22.23	8.00	<=SB
121	DRB1*01:01	LKRYYGRILHYLKAK	Sequence	4	YGRILHYLK	0.732	18.14	6.00	<=SB
122	DRB1*01:01	KRYYGRILHYLKAKE	Sequence	3	YGRILHYLK	0.734	17.81	6.00	<=SB
123	DRB1*01:01	RYYGRILHYLKAKEY	Sequence	2	YGRILHYLK	0.734	17.84	6.00	<=SB
124	DRB1*01:01	YYGRILHYLKAKEYS	Sequence	5	LHYLKAKEY	0.726	19.48	7.00	<=SB
125	DRB1*01:01	YGRILHYLKAKEYSH	Sequence	4	LHYLKAKEY	0.739	16.90	6.00	<=SB
126	DRB1*01:01	GRILHYLKAKEYSHC	Sequence	3	LHYLKAKEY	0.738	17.04	6.00	<=SB
127	DRB1*01:01	RILHYLKAKEYSHCA	Sequence	2	LHYLKAKEY	0.710	23.05	8.00	<=SB
128	DRB1*01:01	ILHYLKAKEYSHCAW	Sequence	1	LHYLKAKEY	0.674	33.91	15.00	<=SB
129	DRB1*01:01	LHYLKAKEYSHCAWT	Sequence	0	LHYLKAKEY	0.559	118.40	32.00	<=WB
130	DRB1*01:01	HYLKAKEYSHCAWTI	Sequence	1	YLKAKEYSH	0.478	284.21	50.00	<=WB
131	DRB1*01:01	YLKAKEYSHCAWTIV	Sequence	6	YSHCAWTIV	0.497	229.98	50.00	<=WB
132	DRB1*01:01	LKAKEYSHCAWTIVR	Sequence	5	YSHCAWTIV	0.509	203.03	50.00	<=WB
133	DRB1*01:01	KAKEYSHCAWTIVRV	Sequence	4	YSHCAWTIV	0.568	107.18	32.00	<=WB
134	DRB1*01:01	AKEYSHCAWTIVRVE	Sequence	3	YSHCAWTIV	0.563	112.79	32.00	<=WB
135	DRB1*01:01	KEYSHCAWTIVRVEI	Sequence	2	YSHCAWTIV	0.539	146.64	50.00	<=WB
136	DRB1*01:01	EYSHCAWTIVRVEIL	Sequence	6	WTIVRVEIL	0.582	91.62	32.00	<=WB
137	DRB1*01:01	YSHCAWTIVRVEILR	Sequence	5	WTIVRVEIL	0.589	85.38	32.00	<=WB
138	DRB1*01:01	SHCAWTIVRVEILRN	Sequence	4	WTIVRVEIL	0.600	75.93	32.00	<=WB
139	DRB1*01:01	HCAWTIVRVEILRNF	Sequence	3	WTIVRVEIL	0.623	59.03	32.00	<=WB
140	DRB1*01:01	CAWTIVRVEILRNFY	Sequence	2	WTIVRVEIL	0.622	59.98	32.00	<=WB
141	DRB1*01:01	AWTIVRVEILRNFYF	Sequence	1	WTIVRVEIL	0.637	50.87	32.00	<=WB
142	DRB1*01:01	WTIVRVEILRNFYFI	Sequence	5	VEILRNFYF	0.641	48.80	32.00	<=SB
143	DRB1*01:01	TIVRVEILRNFYFIN	Sequence	4	VEILRNFYF	0.624	58.77	32.00	<=WB
144	DRB1*01:01	IVRVEILRNFYFINR	Sequence	3	VEILRNFYF	0.622	59.65	32.00	<=WB
145	DRB1*01:01	VRVEILRNFYFINRL	Sequence	2	VEILRNFYF	0.589	85.25	32.00	<=WB
146	DRB1*01:01	RVEILRNFYFINRLT	Sequence	1	VEILRNFYF	0.576	98.24	32.00	<=WB
147	DRB1*01:01	VEILRNFYFINRLTG	Sequence	5	NFYFINRLT	0.550	129.84	50.00	<=WB
148	DRB1*01:01	EILRNFYFINRLTGY	Sequence	4	NFYFINRLT	0.587	87.35	32.00	<=WB
149	DRB1*01:01	ILRNFYFINRLTGYL	Sequence	6	FINRLTGYL	0.672	34.92	15.00	<=SB
150	DRB1*01:01	LRNFYFINRLTGYLR	Sequence	5	FINRLTGYL	0.734	17.79	6.00	<=SB
151	DRB1*01:01	RNFYFINRLTGYLRN	Sequence	4	FINRLTGYL	0.752	14.68	5.00	<=SB

Allele: DRB1*03:01

pos	Allele	peptide	Identity	Pos	Core	1-logSOk(aff)	Affinity(nM)	%Rank	BindingLevel
0	DRB1*03:01	MSYNLLGFLQRSSNF	Sequence	2	YNLLGFLQR	0.274	2590.72	50.00	
1	DRB1*03:01	SYNLLGFLQRSSNFQ	Sequence	6	FLQRSSNFQ	0.298	1985.58	50.00	
2	DRB1*03:01	YNLLGFLQRSSNFQC	Sequence	5	FLQRSSNFQ	0.304	1867.81	50.00	
3	DRB1*03:01	NLLGFLQRSSNFQCQ	Sequence	4	FLQRSSNFQ	0.301	1933.87	50.00	
4	DRB1*03:01	LLGFLQRSSNFQCQK	Sequence	3	FLQRSSNFQ	0.315	1656.31	50.00	
5	DRB1*03:01	LGFLQRSSNFQCQKL	Sequence	2	FLQRSSNFQ	0.292	2123.12	50.00	
6	DRB1*03:01	GFLQRSSNFQCQKLL	Sequence	1	FLQRSSNFQ	0.256	3145.89	50.00	
7	DRB1*03:01	FLQRSSNFQCQKLLW	Sequence	0	FLQRSSNFQ	0.240	3709.57	50.00	
8	DRB1*03:01	LQRSSNFQCQKLLWQ	Sequence	4	SNFQCQKLL	0.220	4612.96	50.00	
9	DRB1*03:01	QRSSNFQCQKLLWQL	Sequence	3	SNFQCQKLL	0.232	4082.83	50.00	
10	DRB1*03:01	RSSNFQCQKLLWQLN	Sequence	2	SNFQCQKLL	0.230	4136.50	50.00	
11	DRB1*03:01	SSNFQCQKLLWQLNG	Sequence	3	FQCQKLLWQ	0.214	4913.26	50.00	
12	DRB1*03:01	SNFQCQKLLWQLNGR	Sequence	6	KLLWQLNGR	0.271	2662.10	50.00	
13	DRB1*03:01	NFQCQKLLWQLNGRL	Sequence	6	LLWQLNGRL	0.431	473.60	15.00	<=WB
14	DRB1*03:01	FQCQKLLWQLNGRLE	Sequence	5	LLWQLNGRL	0.450	385.36	15.00	<=WB
15	DRB1*03:01	QCQKLLWQLNGRLEY	Sequence	4	LLWQLNGRL	0.531	160.03	5.00	<=WB
16	DRB1*03:01	CQKLLWQLNGRLEYC	Sequence	3	LLWQLNGRL	0.537	149.44	4.00	<=WB
17	DRB1*03:01	QKLLWQLNGRLEYCL	Sequence	4	WQLNGRLEY	0.543	140.27	4.00	<=WB
18	DRB1*03:01	KLLWQLNGRLEYCLK	Sequence	3	WQLNGRLEY	0.542	141.63	4.00	<=WB
19	DRB1*03:01	LLWQLNGRLEYCLKD	Sequence	2	WQLNGRLEY	0.441	425.22	15.00	<=WB
20	DRB1*03:01	LWQLNGRLEYCLKDR	Sequence	1	WQLNGRLEY	0.370	910.72	32.00	
21	DRB1*03:01	WQLNGRLEYCLKDRM	Sequence	0	WQLNGRLEY	0.374	877.66	32.00	
22	DRB1*03:01	QLNGRLEYCLKDRMN	Sequence	5	LEYCLKDRM	0.333	1356.04	50.00	
23	DRB1*03:01	LNGRLEYCLKDRMNF	Sequence	4	LEYCLKDRM	0.425	501.39	15.00	
24	DRB1*03:01	NGRLEYCLKDRMNFD	Sequence	5	YCLKDRMNF	0.433	463.97	15.00	<=WB
25	DRB1*03:01	GRLEYCLKDRMNFDI	Sequence	4	YCLKDRMNF	0.446	402.60	15.00	<=WB
26	DRB1*03:01	RLEYCLKDRMNFDIP	Sequence	3	YCLKDRMNF	0.447	396.14	15.00	<=WB
27	DRB1*03:01	LEYCLKDRMNFDIPE	Sequence	2	YCLKDRMNF	0.394	701.97	32.00	
28	DRB1*03:01	EYCLKDRMNFDIPEE	Sequence	1	YCLKDRMNF	0.339	1279.51	32.00	
29	DRB1*03:01	YCLKDRMNFDIPEEI	Sequence	0	YCLKDRMNF	0.381	813.70	32.00	
30	DRB1*03:01	CLKDRMNFDIPEEIK	Sequence	5	MNFDIPEEI	0.364	972.82	32.00	
31	DRB1*03:01	LKDRMNFDIPEEIKQ	Sequence	4	MNFDIPEEI	0.398	677.73	32.00	
32	DRB1*03:01	KDRMNFDIPEEIKQL	Sequence	3	MNFDIPEEI	0.430	475.33	15.00	<=WB
33	DRB1*03:01	DRMNFDIPEEIKQLQ	Sequence	2	MNFDIPEEI	0.393	709.23	32.00	
34	DRB1*03:01	RMNFDIPEEIKQLQQ	Sequence	1	MNFDIPEEI	0.376	855.36	32.00	
35	DRB1*03:01	MNFDIPEEIKQLQQF	Sequence	0	MNFDIPEEI	0.321	1555.65	50.00	
36	DRB1*03:01	NFDIPEEIKQLQQFQ	Sequence	3	IPEEIKQLQ	0.221	4579.24	50.00	
37	DRB1*03:01	FDIPEEIKQLQQFQK	Sequence	6	IKQLQQFQK	0.243	3616.71	50.00	
38	DRB1*03:01	DIPEEIKQLQQFQKE	Sequence	5	IKQLQQFQK	0.237	3856.83	50.00	
39	DRB1*03:01	IPEEIKQLQQFQKED	Sequence	4	IKQLQQFQK	0.231	4096.46	50.00	
40	DRB1*03:01	PEEIKQLQQFQKEDA	Sequence	3	IKQLQQFQK	0.231	4127.73	50.00	
41	DRB1*03:01	EEIKQLQQFQKEDAA	Sequence	2	IKQLQQFQK	0.222	4541.59	50.00	
42	DRB1*03:01	EIKQLQQFQKEDAAL	Sequence	4	LQQFQKEDA	0.245	3526.75	50.00	
43	DRB1*03:01	IKQLQQFQKEDAALT	Sequence	6	FQKEDAALT	0.292	2114.52	50.00	
44	DRB1*03:01	KQLQQFQKEDAALTI	Sequence	5	FQKEDAALT	0.331	1391.67	50.00	
45	DRB1*03:01	QLQQFQKEDAALTIY	Sequence	4	FQKEDAALT	0.380	821.80	32.00	
46	DRB1*03:01	LQQFQKEDAALTIYE	Sequence	3	FQKEDAALT	0.390	736.35	32.00	
47	DRB1*03:01	QQFQKEDAALTIYEM	Sequence	2	FQKEDAALT	0.361	1008.13	32.00	
48	DRB1*03:01	QFQKEDAALTIYEML	Sequence	2	QKEDAALTI	0.321	1546.65	50.00	
49	DRB1*03:01	FQKEDAALTIYEMLQ	Sequence	0	FQKEDAALT	0.251	3291.83	50.00	
50	DRB1*03:01	QKEDAALTIYEMLQN	Sequence	0	QKEDAALTI	0.203	5549.96	50.00	
51	DRB1*03:01	KEDAALTIYEMLQNI	Sequence	6	TIYEMLQNI	0.213	4981.29	50.00	
52	DRB1*03:01	EDAALTIYEMLQNIF	Sequence	5	TIYEMLQNI	0.238	3812.68	50.00	
53	DRB1*03:01	DAALTIYEMLQNIFA	Sequence	4	TIYEMLQNI	0.253	3226.77	50.00	
54	DRB1*03:01	AALTIYEMLQNIFAI	Sequence	3	TIYEMLQNI	0.272	2638.56	50.00	
55	DRB1*03:01	ALTIYEMLQNIFAIF	Sequence	2	TIYEMLQNI	0.284	2319.42	50.00	
56	DRB1*03:01	LTIYEMLQNIFAIFR	Sequence	6	LQNIFAIFR	0.305	1836.81	50.00	
57	DRB1*03:01	TIYEMLQNIFAIFRQ	Sequence	5	LQNIFAIFR	0.309	1775.19	50.00	
58	DRB1*03:01	IYEMLQNIFAIFRQD	Sequence	4	LQNIFAIFR	0.303	1894.19	50.00	
59	DRB1*03:01	YEMLQNIFAIFRQDS	Sequence	6	IFAIFRQDS	0.339	1274.38	32.00	
60	DRB1*03:01	EMLQNIFAIFRQDSS	Sequence	5	IFAIFRQDS	0.344	1208.85	32.00	
61	DRB1*03:01	MLQNIFAIFRQDSSS	Sequence	4	IFAIFRQDS	0.363	982.63	32.00	
62	DRB1*03:01	LQNIFAIFRQDSSST	Sequence	3	IFAIFRQDS	0.401	652.89	32.00	
63	DRB1*03:01	QNIFAIFRQDSSSTG	Sequence	2	IFAIFRQDS	0.413	574.61	32.00	
64	DRB1*03:01	NIFAIFRQDSSSTGW	Sequence	5	FRQDSSSTG	0.428	485.19	15.00	<=WB
65	DRB1*03:01	IFAIFRQDSSSTGWN	Sequence	4	FRQDSSSTG	0.448	393.65	15.00	<=WB
66	DRB1*03:01	FAIFRQDSSSTGWNE	Sequence	3	FRQDSSSTG	0.429	482.66	15.00	<=WB
67	DRB1*03:01	AIFRQDSSSTGWNET	Sequence	2	FRQDSSSTG	0.380	822.27	32.00	
68	DRB1*03:01	IFRQDSSSTGWNETI	Sequence	1	FRQDSSSTG	0.322	1528.21	50.00	

67	DRB1*03:01	AIFRQDSSSTGWNET	Sequence	2	FRQDSSSTG	0.380	822.27	32.00	
68	DRB1*03:01	IFRQDSSSTGWNETI	Sequence	1	FRQDSSSTG	0.322	1528.21	50.00	
69	DRB1*03:01	FRQDSSSTGWNETIV	Sequence	0	FRQDSSSTG	0.171	7843.78	50.00	
70	DRB1*03:01	RQDSSSTGWNETIVE	Sequence	6	TGWNETIVE	0.074	22392.35	50.00	
71	DRB1*03:01	QDSSSTGWNETIVEN	Sequence	5	TGWNETIVE	0.081	20801.97	50.00	
72	DRB1*03:01	DSSSTGWNETIVENL	Sequence	6	WNETIVENL	0.116	14328.83	50.00	
73	DRB1*03:01	SSSTGWNETIVENLL	Sequence	5	WNETIVENL	0.147	10208.10	50.00	
74	DRB1*03:01	SSTGWNETIVENLLA	Sequence	4	WNETIVENL	0.168	8131.10	50.00	
75	DRB1*03:01	STGWNETIVENLLAN	Sequence	6	TIVENLLAN	0.213	5002.36	50.00	
76	DRB1*03:01	TGWNETIVENLLANV	Sequence	6	IVENLLANV	0.295	2053.33	50.00	
77	DRB1*03:01	GWNETIVENLLANVY	Sequence	5	IVENLLANV	0.336	1321.05	32.00	
78	DRB1*03:01	WNETIVENLLANVYH	Sequence	4	IVENLLANV	0.364	972.22	32.00	
79	DRB1*03:01	NETIVENLLANVYHQ	Sequence	3	IVENLLANV	0.377	842.56	32.00	
80	DRB1*03:01	ETIVENLLANVYHQI	Sequence	6	LLANVYHQI	0.390	736.29	32.00	
81	DRB1*03:01	TIVENLLANVYHQIN	Sequence	5	LLANVYHQI	0.387	755.77	32.00	
82	DRB1*03:01	IVENLLANVYHQINH	Sequence	4	LLANVYHQI	0.394	703.74	32.00	
83	DRB1*03:01	VENLLANVYHQINHL	Sequence	3	LLANVYHQI	0.411	585.00	32.00	
84	DRB1*03:01	ENLLANVYHQINHLK	Sequence	2	LLANVYHQI	0.427	493.20	15.00	<=WB
85	DRB1*03:01	NLLANVYHQINHLKT	Sequence	1	LLANVYHQI	0.430	478.00	15.00	<=WB
86	DRB1*03:01	LLANVYHQINHLKTV	Sequence	4	VYHQINHLK	0.450	384.35	15.00	<=WB
87	DRB1*03:01	LANVYHQINHLKTVL	Sequence	3	VYHQINHLK	0.468	314.68	10.00	<=WB
88	DRB1*03:01	ANVYHQINHLKTVLE	Sequence	2	VYHQINHLK	0.442	416.59	15.00	<=WB
89	DRB1*03:01	NVYHQINHLKTVLEE	Sequence	1	VYHQINHLK	0.412	579.75	32.00	
90	DRB1*03:01	VYHQINHLKTVLEEK	Sequence	0	VYHQINHLK	0.392	720.74	32.00	
91	DRB1*03:01	YHQINHLKTVLEEKL	Sequence	3	INHLKTVLE	0.356	1057.52	32.00	
92	DRB1*03:01	HQINHLKTVLEEKLE	Sequence	2	INHLKTVLE	0.314	1676.81	50.00	
93	DRB1*03:01	QINHLKTVLEEKLEK	Sequence	1	INHLKTVLE	0.313	1695.35	50.00	
94	DRB1*03:01	INHLKTVLEEKLEKE	Sequence	3	LKTVLEEKL	0.286	2256.81	50.00	
95	DRB1*03:01	NHLKTVLEEKLEKED	Sequence	4	TVLEEKLEK	0.254	3199.24	50.00	
96	DRB1*03:01	HLKTVLEEKLEKEDF	Sequence	3	TVLEEKLEK	0.273	2614.41	50.00	
97	DRB1*03:01	LKTVLEEKLEKEDFT	Sequence	2	TVLEEKLEK	0.253	3239.68	50.00	
98	DRB1*03:01	KTVLEEKLEKEDFTR	Sequence	1	TVLEEKLEK	0.212	5060.39	50.00	
99	DRB1*03:01	TVLEEKLEKEDFTRG	Sequence	0	TVLEEKLEK	0.189	6457.45	50.00	
100	DRB1*03:01	VLEEKLEKEDFTRGK	Sequence	0	VLEEKLEKE	0.176	7469.64	50.00	
101	DRB1*03:01	LEEKLEKEDFTRGKL	Sequence	3	KLEKEDFTR	0.189	6492.48	50.00	
102	DRB1*03:01	EEKLEKEDFTRGKLM	Sequence	3	LEKEDFTRG	0.206	5397.82	50.00	
103	DRB1*03:01	EKLEKEDFTRGKLMS	Sequence	2	LEKEDFTRG	0.210	5181.50	50.00	
104	DRB1*03:01	KLEKEDFTRGKLMSS	Sequence	6	FTRGKLMSS	0.223	4487.04	50.00	
105	DRB1*03:01	LEKEDFTRGKLMSSL	Sequence	5	FTRGKLMSS	0.222	4521.35	50.00	
106	DRB1*03:01	EKEDFTRGKLMSSLH	Sequence	4	FTRGKLMSS	0.231	4089.33	50.00	
107	DRB1*03:01	KEDFTRGKLMSSLHL	Sequence	3	FTRGKLMSS	0.263	2895.33	50.00	
108	DRB1*03:01	EDFTRGKLMSSLHLK	Sequence	6	KLMSSLHLK	0.312	1714.83	50.00	
109	DRB1*03:01	DFTRGKLMSSLHLKR	Sequence	6	LMSSLHLKR	0.630	54.81	0.80	<=WB
110	DRB1*03:01	FTRGKLMSSLHLKRY	Sequence	5	LMSSLHLKR	0.672	34.78	0.30	<=SB
111	DRB1*03:01	TRGKLMSSLHLKRYY	Sequence	4	LMSSLHLKR	0.705	24.23	0.12	<=SB
112	DRB1*03:01	RGKLMSSLHLKRYYG	Sequence	3	LMSSLHLKR	0.695	27.17	0.15	<=SB
113	DRB1*03:01	GKLMSSLHLKRYYGR	Sequence	2	LMSSLHLKR	0.642	48.13	0.80	<=SB
114	DRB1*03:01	KLMSSLHLKRYYGRI	Sequence	1	LMSSLHLKR	0.607	70.46	1.50	<=WB
115	DRB1*03:01	LMSSLHLKRYYGRIL	Sequence	0	LMSSLHLKR	0.528	164.30	5.00	<=WB
116	DRB1*03:01	MSSLHLKRYYGRILH	Sequence	3	LHLKRYYGR	0.389	744.32	32.00	
117	DRB1*03:01	SSLHLKRYYGRILHY	Sequence	5	KRYYGRILH	0.391	729.64	32.00	
118	DRB1*03:01	SLHLKRYYGRILHYL	Sequence	4	KRYYGRILH	0.411	588.76	32.00	
119	DRB1*03:01	LHLKRYYGRILHYLK	Sequence	3	KRYYGRILH	0.435	453.27	15.00	<=WB
120	DRB1*03:01	HLKRYYGRILHYLKA	Sequence	2	KRYYGRILH	0.414	564.34	32.00	
121	DRB1*03:01	LKRYYGRILHYLKAK	Sequence	4	YGRILHYLK	0.420	532.80	16.00	
122	DRB1*03:01	KRYYGRILHYLKAKE	Sequence	6	ILHYLKAKE	0.443	412.97	15.00	<=WB
123	DRB1*03:01	RYYGRILHYLKAKEY	Sequence	5	ILHYLKAKE	0.438	435.19	15.00	<=WB
124	DRB1*03:01	YYGRILHYLKAKEYS	Sequence	4	ILHYLKAKE	0.456	361.55	15.00	<=WB
125	DRB1*03:01	YGRILHYLKAKEYSH	Sequence	3	ILHYLKAKE	0.484	265.86	8.00	<=WB
126	DRB1*03:01	GRILHYLKAKEYSHC	Sequence	2	ILHYLKAKE	0.456	361.22	15.00	<=WB
127	DRB1*03:01	RILHYLKAKEYSHCA	Sequence	1	ILHYLKAKE	0.431	472.85	15.00	<=WB
128	DRB1*03:01	ILHYLKAKEYSHCAW	Sequence	0	ILHYLKAKE	0.377	841.65	32.00	
129	DRB1*03:01	LHYLKAKEYSHCAWT	Sequence	2	YLKAKEYSH	0.250	3338.24	50.00	
130	DRB1*03:01	HYLKAKEYSHCAWTI	Sequence	1	YLKAKEYSH	0.198	5885.48	50.00	
131	DRB1*03:01	YLKAKEYSHCAWTIV	Sequence	0	YLKAKEYSH	0.183	6935.97	50.00	
132	DRB1*03:01	LKAKEYSHCAWTIVR	Sequence	6	SHCAWTIVR	0.178	7320.11	50.00	
133	DRB1*03:01	KAKEYSHCAWTIVRV	Sequence	5	SHCAWTIVR	0.202	5596.09	50.00	
134	DRB1*03:01	AKEYSHCAWTIVRVE	Sequence	4	SHCAWTIVR	0.212	5038.75	50.00	
135	DRB1*03:01	KEYSHCAWTIVRVEI	Sequence	3	SHCAWTIVR	0.230	4167.13	50.00	
136	DRB1*03:01	EYSHCAWTIVRVEIL	Sequence	6	WTIVRVEIL	0.246	3474.86	50.00	
105	DRB1*03:01	LEKEDFTRGKLMSSL	Sequence	5	FTRGKLMSS	0.222	4521.35	50.00	
106	DRB1*03:01	EKEDFTRGKLMSSLH	Sequence	4	FTRGKLMSS	0.231	4089.33	50.00	
107	DRB1*03:01	KEDFTRGKLMSSLHL	Sequence	3	FTRGKLMSS	0.263	2895.33	50.00	
108	DRB1*03:01	EDFTRGKLMSSLHLK	Sequence	6	KLMSSLHLK	0.312	1714.83	50.00	
109	DRB1*03:01	DFTRGKLMSSLHLKR	Sequence	6	LMSSLHLKR	0.630	54.81	0.80	<=WB
110	DRB1*03:01	FTRGKLMSSLHLKRY	Sequence	5	LMSSLHLKR	0.672	34.78	0.30	<=SB
111	DRB1*03:01	TRGKLMSSLHLKRYY	Sequence	4	LMSSLHLKR	0.705	24.23	0.12	<=SB
112	DRB1*03:01	RGKLMSSLHLKRYYG	Sequence	3	LMSSLHLKR	0.695	27.17	0.15	<=SB
113	DRB1*03:01	GKLMSSLHLKRYYGR	Sequence	2	LMSSLHLKR	0.642	48.13	0.80	<=SB
114	DRB1*03:01	KLMSSLHLKRYYGRI	Sequence	1	LMSSLHLKR	0.607	70.46	1.50	<=WB
115	DRB1*03:01	LMSSLHLKRYYGRIL	Sequence	0	LMSSLHLKR	0.528	164.30	5.00	<=WB
116	DRB1*03:01	MSSLHLKRYYGRILH	Sequence	3	LHLKRYYGR	0.389	744.32	32.00	
117	DRB1*03:01	SSLHLKRYYGRILHY	Sequence	5	KRYYGRILH	0.391	729.64	32.00	
118	DRB1*03:01	SLHLKRYYGRILHYL	Sequence	4	KRYYGRILH	0.411	588.76	32.00	
119	DRB1*03:01	LHLKRYYGRILHYLK	Sequence	3	KRYYGRILH	0.435	453.27	15.00	<=WB
120	DRB1*03:01	HLKRYYGRILHYLKA	Sequence	2	KRYYGRILH	0.414	564.34	32.00	
121	DRB1*03:01	LKRYYGRILHYLKAK	Sequence	4	YGRILHYLK	0.420	532.80	16.00	
122	DRB1*03:01	KRYYGRILHYLKAKE	Sequence	6	ILHYLKAKE	0.443	412.97	15.00	<=WB
123	DRB1*03:01	RYYGRILHYLKAKEY	Sequence	5	ILHYLKAKE	0.438	435.19	15.00	<=WB
124	DRB1*03:01	YYGRILHYLKAKEYS	Sequence	4	ILHYLKAKE	0.456	361.55	15.00	<=WB
125	DRB1*03:01	YGRILHYLKAKEYSH	Sequence	3	ILHYLKAKE	0.484	265.86	8.00	<=WB
126	DRB1*03:01	GRILHYLKAKEYSHC	Sequence	2	ILHYLKAKE	0.456	361.22	15.00	<=WB
127	DRB1*03:01	RILHYLKAKEYSHCA	Sequence	1	ILHYLKAKE	0.431	472.85	15.00	<=WB
128	DRB1*03:01	ILHYLKAKEYSHCAW	Sequence	0	ILHYLKAKE	0.377	841.65	32.00	
129	DRB1*03:01	LHYLKAKEYSHCAWT	Sequence	2	YLKAKEYSH	0.250	3338.24	50.00	
130	DRB1*03:01	HYLKAKEYSHCAWTI	Sequence	1	YLKAKEYSH	0.198	5885.48	50.00	
131	DRB1*03:01	YLKAKEYSHCAWTIV	Sequence	0	YLKAKEYSH	0.183	6935.97	50.00	
132	DRB1*03:01	LKAKEYSHCAWTIVR	Sequence	6	SHCAWTIVR	0.178	7320.11	50.00	
133	DRB1*03:01	KAKEYSHCAWTIVRV	Sequence	5	SHCAWTIVR	0.202	5596.09	50.00	
134	DRB1*03:01	AKEYSHCAWTIVRVE	Sequence	4	SHCAWTIVR	0.212	5038.75	50.00	
135	DRB1*03:01	KEYSHCAWTIVRVEI	Sequence	3	SHCAWTIVR	0.230	4167.13	50.00	
136	DRB1*03:01	EYSHCAWTIVRVEIL	Sequence	6	WTIVRVEIL	0.246	3474.86	50.00	

137	DRB1*03:01	YSHCAWTIVRVEILR	Sequence	6	TIVRVEILR	0.315	1658.03	50.00	
138	DRB1*03:01	SHCAWTIVRVEILRN	Sequence	6	IVRVEILRN	0.394	704.33	32.00	
139	DRB1*03:01	HCAWTIVRVEILRNF	Sequence	5	IVRVEILRN	0.466	324.20	10.00	<=WB
140	DRB1*03:01	CAWTIVRVEILRNFY	Sequence	4	IVRVEILRN	0.513	193.49	6.00	<=WB
141	DRB1*03:01	AWTIVRVEILRNFYF	Sequence	3	IVRVEILRN	0.539	147.25	4.00	<=WB
142	DRB1*03:01	WTIVRVEILRNFYFI	Sequence	3	VRVEILRNF	0.541	143.46	4.00	<=WB
143	DRB1*03:01	TIVRVEILRNFYFIN	Sequence	2	VRVEILRNF	0.517	185.62	6.00	<=WB
144	DRB1*03:01	IVRVEILRNFYFINR	Sequence	1	VRVEILRNF	0.473	300.79	9.00	<=WB
145	DRB1*03:01	VRVEILRNFYFINRL	Sequence	0	VRVEILRNF	0.417	546.05	16.00	
146	DRB1*03:01	RVEILRNFYFINRLT	Sequence	3	ILRNFYFIN	0.368	931.33	32.00	
147	DRB1*03:01	VEILRNFYFINRLTG	Sequence	6	FYFINRLTG	0.330	1414.36	50.00	
148	DRB1*03:01	EILRNFYFINRLTGY	Sequence	5	FYFINRLTG	0.356	1059.50	32.00	
149	DRB1*03:01	ILRNFYFINRLTGYL	Sequence	4	FYFINRLTG	0.394	700.50	32.00	
150	DRB1*03:01	LRNFYFINRLTGYLR	Sequence	6	INRLTGYLR	0.488	255.76	8.00	<=WB
151	DRB1*03:01	RNFYFINRLTGYLRN	Sequence	5	INRLTGYLR	0.497	230.05	7.00	<=WB

Fig. 4.Text output after imputing the sequence into the NETMHCIIpan-3.0 of particular alleles DRB1*01:01, DRB3*01:01. Here 15mer peptide predictions were chosen. If the affinity < 500nm then it is weak binder, if the affinity is < 50nm then it is strong binder.*NetMHCII* method refers to the extended SMM align method including direct encoding of peptide flanking residues and penalties for longer peptides and short amino terminal peptide flanking residues

3.5 IEDB-AR Predicted Results

Predicted results are same as results obtained from the server because IEDB-AR server takes 9mer amino acids analysis to predict the T-cell epitopes and the predicted results are based on the percentile ranks, lower the percentile rank higher the binding and hence the predicted results are same as Table 4.

3.6 Proped Predicted Results

Proped is another server to predict the epitopes of T cells. In this study we adopted pan specific method used 51 human leukocyte antigen-DR alleles for the prediction of T-cell epitopes for interferon-beta -1a.It is predicted that 40 patterns are identified. The potential MHCclass II immunogenic sequence of 50 amino acids "TRGKLMSSLHLKRYYGRILHYLKAKEYSHCAW TIVRVEILRNFYFINRLTG" with IFN-β-1a (position 111-161) was identified.

3.7 Confirming the Results of the IEDB and PROPED by Performing Comparative Analysis

Comparative analysis done for the results obtained from the IEDB-AR and proped.After the comparative analysis 40 different overlapping of peptides were identified. The potential MHC class II immunogenic sequence of 50 amino acids"**TRGKLMSSLHLKRYYGRILHYLKAKEYS HCAWTIVRVEILRNFYFINRLTG"** with IFN-β-1a (position 111-161) were identified. Since the overlapping region is found the same it is confirmed that this is the potential T-cell epitope region.

3.8 Epitope Mapping

T-cell epitope mapping is performed for the predicted epitopes identified from IEDB-AR, proped, NetMHCIIpan-3.0 using MIMOX2 server using mimitope. Mimitopes are often peptides which mimics the structure of epitope. MIMOX2 server has an interface for both input and output. The output is shown is shown below.

3.9 Visualization of MHC Binding Motifs

Once MHC binding βpatterns are identified visualization helps us to define and represent the specific regions of the structure and to classify them to which group they belong to: Acidic, Basic, Hydrophilic and hydrophobicity.

T –cell epitope predication plays vital role in vaccine designing, vaccines are mostly based on the B cell immunity but at present it has been encouraged as the host can generate a strong immune response by CD8+ T cell against the infected cell [19]. It is also found that with time, due to antigenic drift, any foreign particle can escape the antibody memory response; however the T cell immune response often provides long-lasting immunity [20]. Therefore, in the present study T cell epitopes were designed using *in silico* computational approaches. In order to predict a peptide that can be used in vaccine development to prevent viral entry or its interaction with host cell, this study focused on interferon Beta -1a protein. In this study multiple prediction methods were applied to determine a potential T-cell epitope considering several criteria like percentile rank, binding affinity and differentiating predicted epitopes into strong and

weak binders. Binding studies show that HLAs are the most polymorphic human genes known [21] and each HLA allele recognizes a restricted set of peptide [22].For instance, vaccine candidates have a tendency to have more binding peptides with low conservation scores and/or lower total conservation scores and/or average stronger binding affinities than non-vaccine candidates [23]. Due to lack of a consensus mapping protocol with immunoinformatics tools, a combined prediction method was applied according to the hypothesis of Trost et al [24] who proposed that greater prediction accuracy can be achieved by combining the predictions from several algorithms rather than relying on just one. For T cell epitope prediction, plenty of algorithms are freely available and in this study we employed IEDB,PROPED,NETMHC analysis tool which is possibly the most wide-ranging database offering several B cell and T cell epitope-related analysis and prediction tools as well as provides both intrinsic biochemical and extrinsic context dependent information about them [25]. The ultimate goal of epitope prediction is to aid the design of molecules that can mimic the structure and function of a genuine epitope and replace it in vaccine design [26] we have developed the model that can be successfully applied as a generic protocol for easy Insilco identification of HLA-DR binding peptides. By examining the output it was predicted that "TRGKLMSSLHLKRYYGRILHYLKAKEYSHCAW TIVRVEILRNFYFINRLTG" these four epitopes would be the best epitope candidates. Thus it is believed that these suggested that T cell epitopes will definitely reduce time, cost and labor during *in vivo* and *in vitro* studies to be carried out for developing a vaccine against Interferon –beta 1a Design and development of vaccine against T cell epitope is much more promising due to the evoke of long lasting immune response and antigenic drift where antigen can easily escape the antibody memory response. The above findings are the result of analyzing the deposited data on various immune databases. The results suggest that, these epitopes may play a highly informative role in antidote production against interferon beta -1a that can trigger an effective immune response in vivo. Along with in silico study, both in vivo and in vitro experiments are required to prove the effectiveness of mounting an immune response.

Table 3. Predicted patterns from the proped results with their positions of varying length.40 different patterns were identified with their varying their length and this is used for comparison analysis with the other epitopes from the other servers

Sl. no	Predicted binders	Binding region
1	NLLGFLQR	4-11
2	LQRSSNFQ	9-16
3	QRSSNFQC	10-17
4	LGFLQRSS	6-13
5	QCQKLL	16-21
6	QCQKLLWQ	16-23
7	QLNGRLEY	23-30
8	QLNGRLE	23-29
9	CKLDR	31-35
10	KDRMNF	33-38
11	KDRMNFDI	33-40
12	NFDIPEEI	37-44
13	QQFQKEDA	48-55
14	KQLQQFQK	45-52
15	QKEDAALT	51-58
16	EMLQNIFA	61-68
17	TIYEMLQN	58-65
18	QNIFAIFR	64-71
19	FAIFRQDS	67-74
20	RQDSSTG	71-77
21	NETIVENL	80-87

Sl. no	Predicted binders	Binding region
22	ENLLANVY	85-92
23	YHQINHLK	91-98
24	NHLKTVLE	96-103
25	HQINHLKT	94-101
26	TRGKLMSS	111-118
27	KRYYGR	122-127
28	KRYYGRIL	122-129
29	GRILHYLK	127-134
30	LHYLKAKKE	129-136
31	HYLKAKEY	130-137
32	VRVEILRN	145-152
33	RVEILRN	146-152
34	TIVRVEIL	143-150
35	EILRNFYF	148-155
36	FINRLTG	155-161
37	YFINRLTG	154-161
38	NRLTGYLR	158-165
39	INRLTGY	157-163
40	EFAI	67-70

Table 4. Patterns obtained after comparative analysis of the patterns from the servers IEDB-AR and proped to confirm the patterns

Sl. no	Allele	Overlapping	Overalpping regions of IEDB and PROPED
1	HLA-DRB1*01:01	TIYEMLQN	58-65
		TIYEMLQN	58-65
		TIYEMLQN	58-65
2	HLA-DRB1*03:01	RGKLMSS	112-118
		TRGKLMSS	111-118
		DRMNFDI	34-40
3	HLA-DRB1*04:01	QNIFAIFR	64-71
		NIFAIFR	65-71
		FAIFRQDS	67-71
4	HLA-DRB1*07:01	TIVRVEIL	143-150
		TIVRVEIL	143-150
		TIVRVEIL	143-150
5	HLA-DRB1*08:01	KRYYGRIL	122-129
		RYYGRIL	123-129
		GRILYLK	127-134
6	HLA-DRB1*09:01	TRGKLMSS	111-118
		TRGKLMSS	111-118
		NLLANVY	86-97
7	HLA-DRB1*10:01	TIYEMLQN	58-65
		TIYEMLQN	58-65
		YFINRLTG	154-161
		YFINRLTG	154-161
8	HLA-DRB1*11:01	YFINRLTG	154-161
		YFINRLTG	154-161
		YFINRLTG	154-161
9	HLA-DRB1*12:01	YFINRLTG	154-161
		YFINRLTG	154-161

Sl. no	Allele	Overlapping	Overalpping regions of IEDB and PROPED
		YFINRLTG	154-161
10	HLA-DRB1*13:01	NLLGFLQR	4-11
		NLLGFLQR	4-11
		NLLGFLQR	4-11
11	HLA-DRB1*14:01	TRGKLMSS	111-118
		RGKLMSS	111-118
		YFINRLTG	154-161
12	HLA-DRB1*15:01	KRYYGRIL	122-129
		KRYYGRIL	122-129
		KRYYGRIL	122-129
13	HLA-DRB1*16:01	KRYYGRIL	122-129
		KRYYGRIL	122-129
		KRYYGRIL	122-129

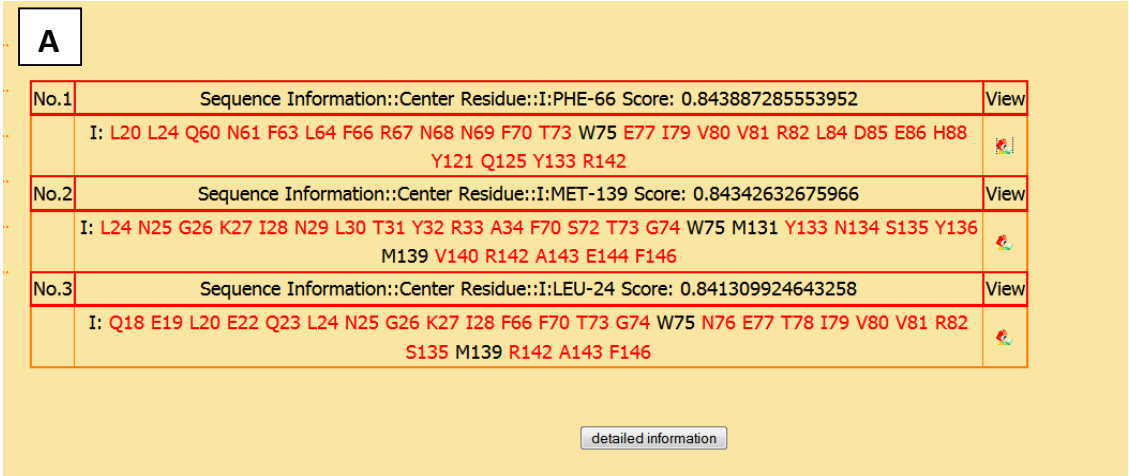
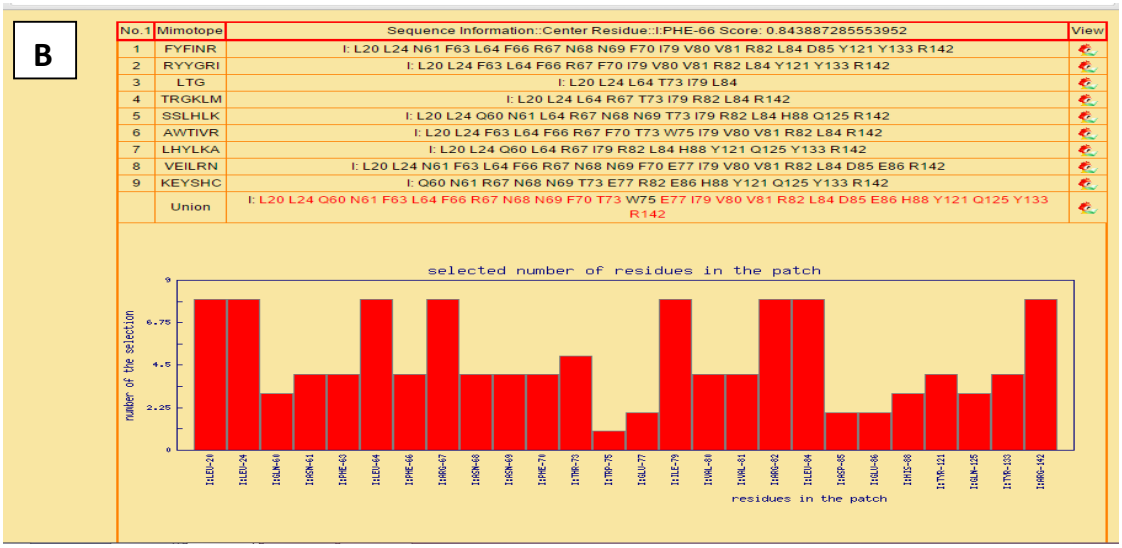


Fig. 5. Screenshots of the Partial output predicted mimitopes from the MIMOX2 server which have highest score. Screenshots representing which patch have an high possibility to be an epitope and indicated by red color. We can view the 3D structure of all the mimitopes



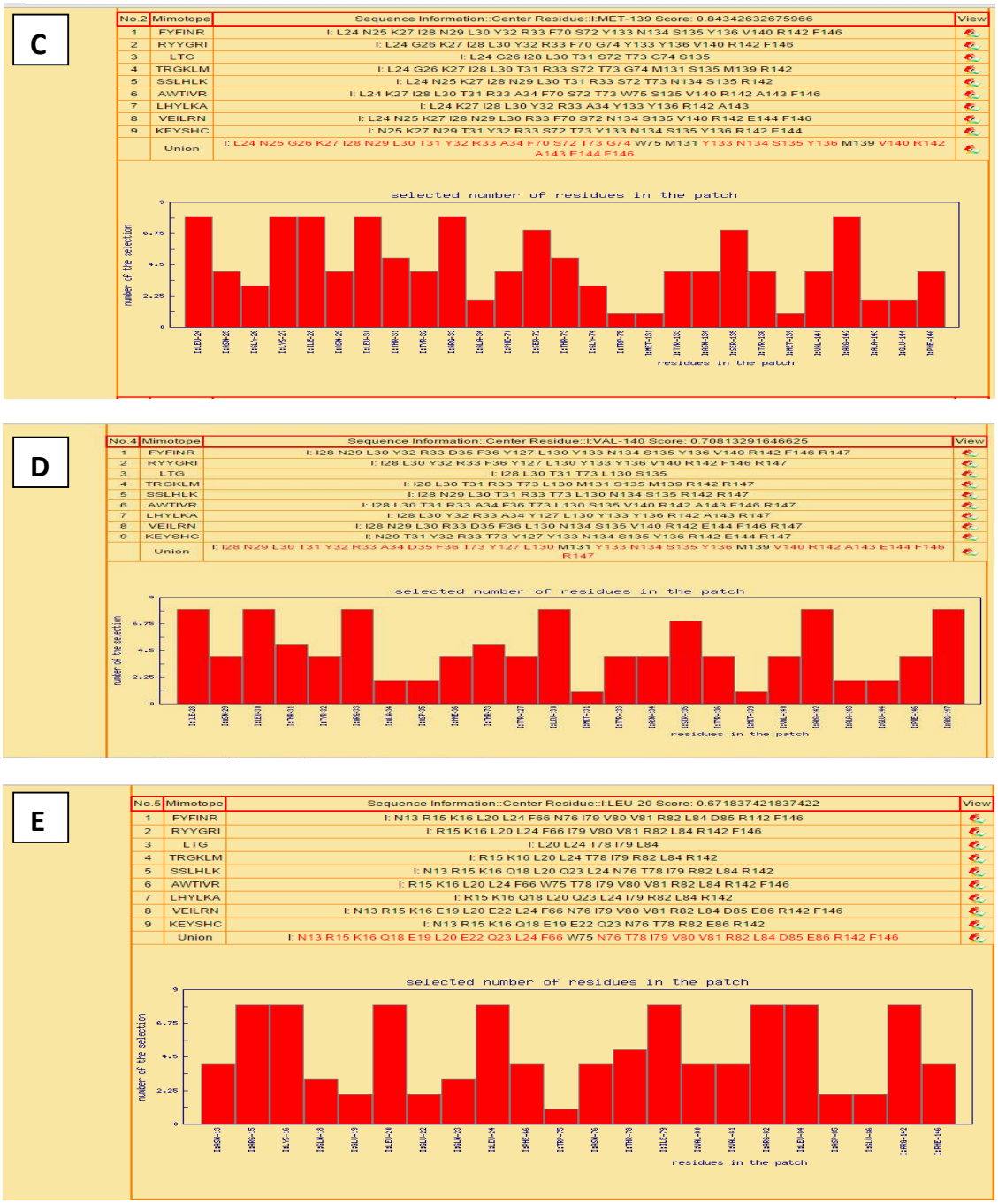


Fig. 6. Screenshots B,C,D,E which represents the detailed information of mimitopes which mimic the epitopes of PHE-66, MET-139, VAL-140, LEU-20 and amino acids which are similar to each mimitope with their corresponding 3D structure. The union represents all the amino acids which are appeared more than twice and denoted in red. The histogram drawn between the number of residues in the patch in the X-axis versus with the number of amino acid residues in the Y-axis which show the occurrence of individual aminoacids

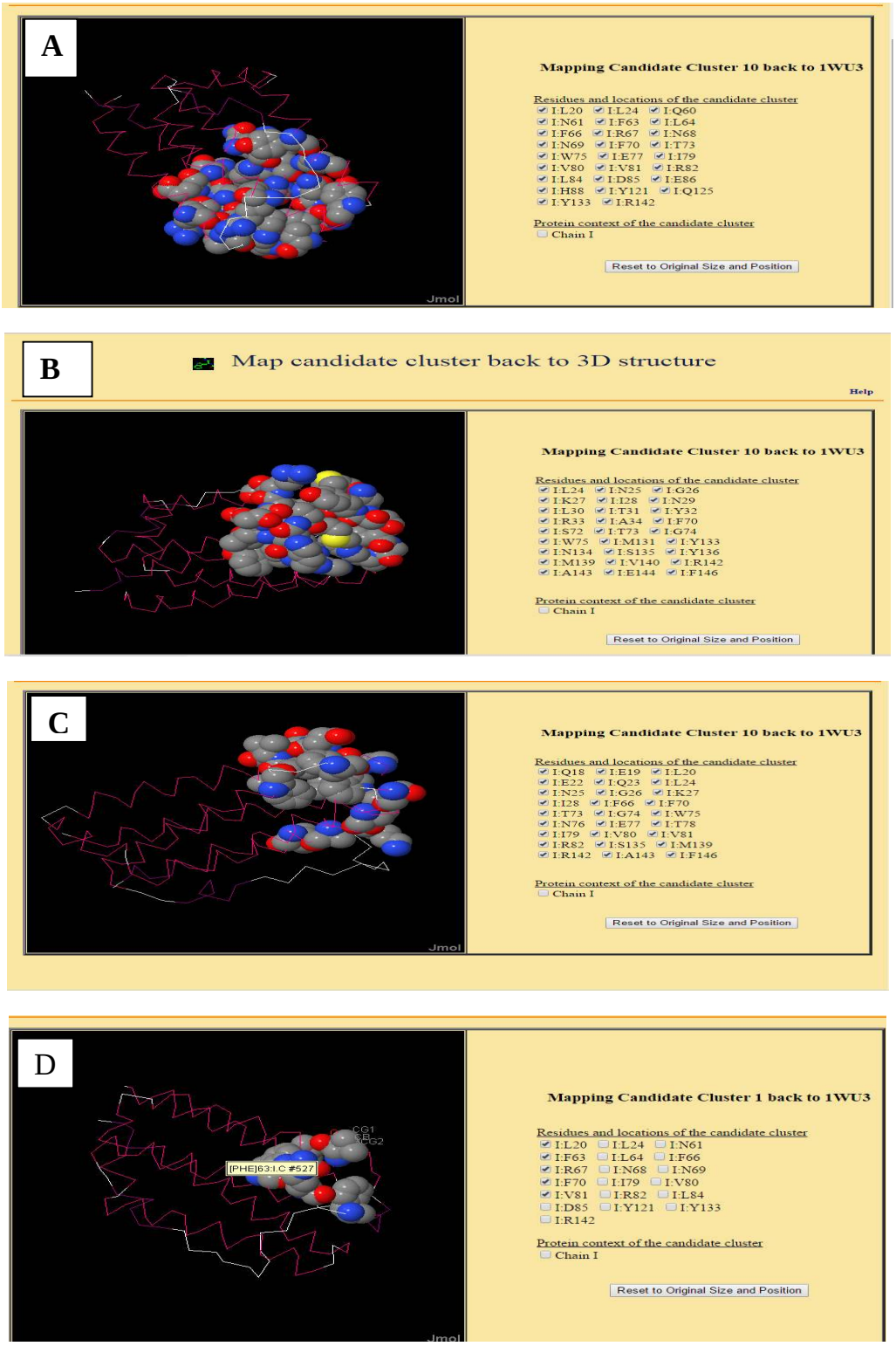
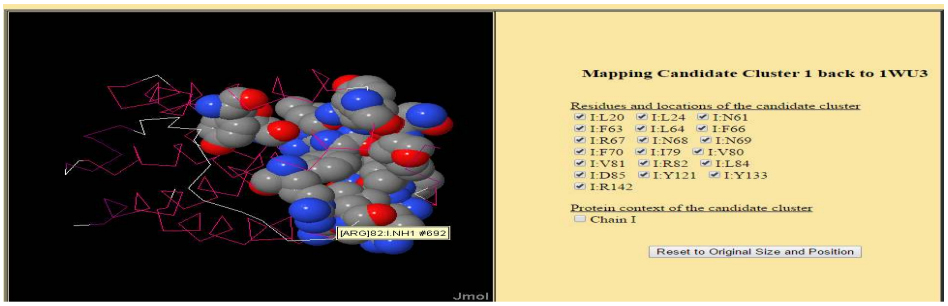
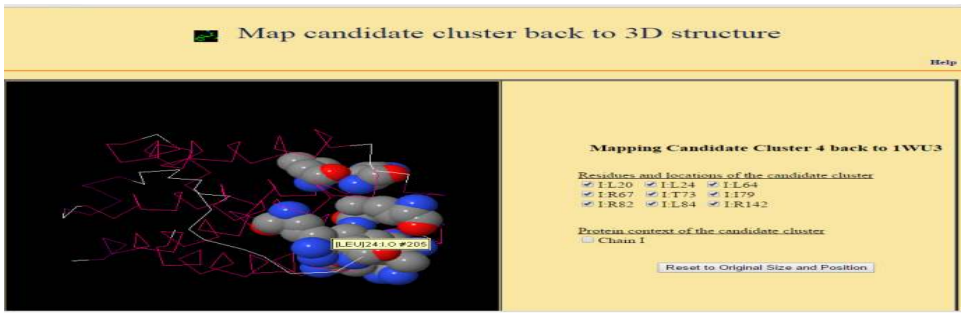


Fig. 7. A,B,C,D Mapping done with all the predicted mimitopes for A) PHE-66 B) MET-139 C) VAL-140 D) LEU-20

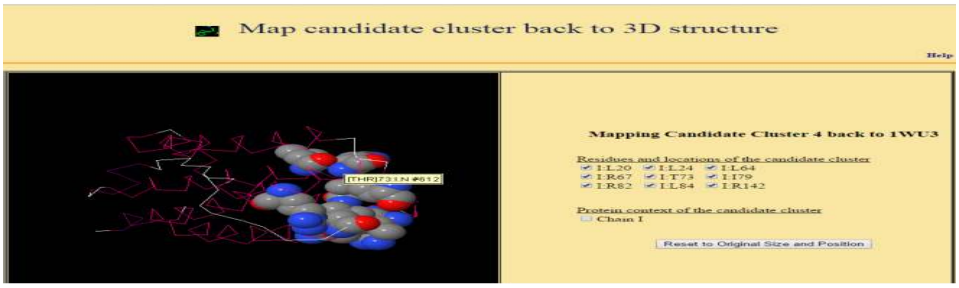
Mapping of FYFINR peptides



Mapping of RYYGRI



Mapping of TRGKLM



Union of all the peptides PHE-66

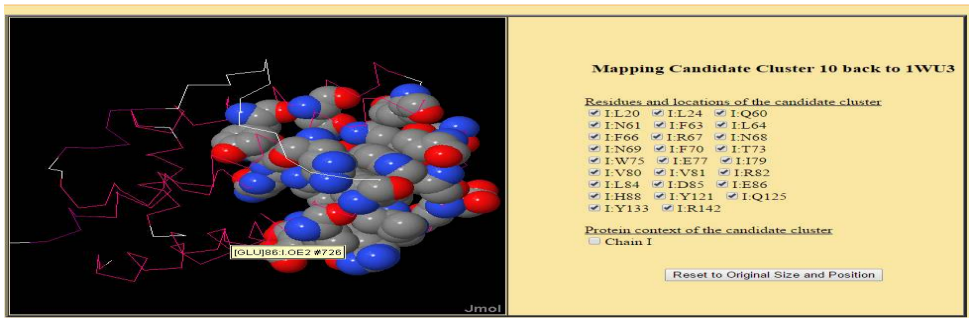


Fig. 8. Mapping candidates for the FYFINR, RYYGRI, TRGKLM, and unions set of all ll the peptides of PHE-66, 3D view of epitopes have been mapped using cartoon representation for 1WU3 (interferon-beta)molecule

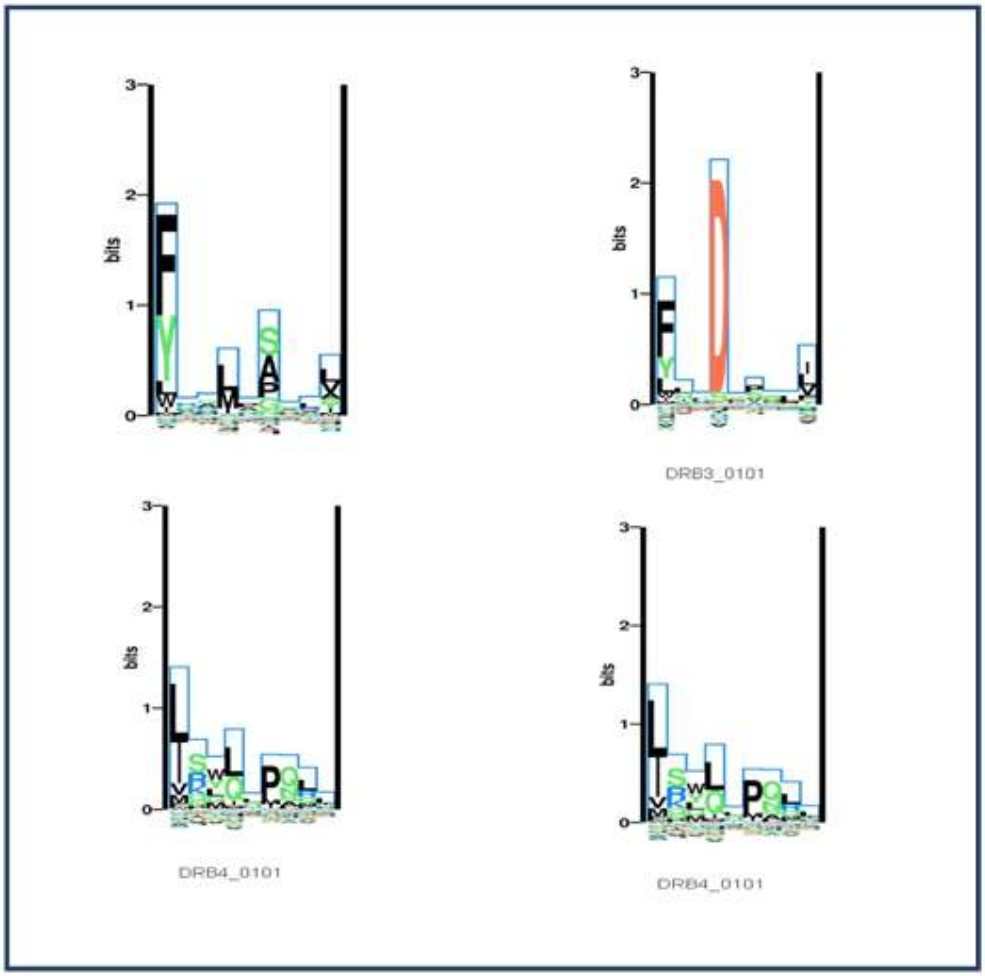


Fig. 9. Seq logos of HLA-DRB1*01:01, HLA-DRB3*01:01, HLA-DRB4*01:01, HLA-DRB5*01:01 are generated using MHC binding motif viewer. kullback –leibler representation has been fallowed which tells us that x-axis indicates the number of amino acids and Y-axis indicates the conserved nature of each amino acids using colour codings. In HLA-DRB1*01:01, HLA-DRB3*01:01, HLA-DRB4*01:01, (peptide1, peptide4, peptide6, peptide9) show the hydrophobic amino acids, (p3, p7) are basic amino acids,(p2,p7)are neutral amino acids and in HLA-DRB3*0101 (p5) represents the acidic amino acids

4. CONCLUSION

The human major histocompatibility genomic molecule region (HLA) is polymorphic comprising several thousand alleles, many encoding the entire distinct molecule. The potentially unique specificities remain experimentally uncharacterized for the majority of HLA molecules thus predicting the T cell responses in assessing the immunogenicity of the protein therapeutics play vital role. Insilico predictions are done using the prediction tools like IEDB-AR, proped, and NETMHCII pan-3.0. We identified an immunogenic sequence of 50 amino acids within IFN-β molecule (position 111-166). After the prediction of epitope, mapping is performed using MIMOX2server. Insilco predictions are advantageous in the pharmaceutical company for designing the vaccines and also helpful in the future invitro assessments. In conclusion NetMHCIIpan-3.0 method is an important step forward in boosting MHC classII binding predictions covering a large number of molecules from different species and therefore reduces experimental costs for the immunologists working within the field of epitope -based vaccine design. Next the seq2 logos were generated. Comparative studies are performed between

IEDB-AR and proped to confirm the patterns and then the influence of different HLA-DR alleles are studied.

The Future scope of predicting the T-cell epitopes is that once we predict the potential T-cell epitopes we can know which set of peptides has major role and synthesize them using in vitro technologies which will save the time and helps in the vaccine development.

ACKNOWLEDGEMENT

Authors would like to express their special thanks and great appreciation to Mr. C V Sudarshan Reddy, for providing assistance in data analysis which helped to finish this project within the time frame.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Govindappa K. Immunopharmacological consequences of immune responses to therapeutic interferon beta (Doctoral dissertation, University of Liverpool); 2013.
2. Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. Bioinformatics. 2001;17(12):1236-7.
3. Myhr K, Mellgren SI. Corticosteroids in the treatment of multiple sclerosis. Acta Neurologica Scandinavica. 2009;120(189): 73-80.
4. Zhang Hao, Claus Lundegard, Morten Nielsen. Pan specific MHC Class I predictions: A benchmark of HLA class I pan-specific prediction methods. Bioinformatics. 2009;25(1):83-89.
5. Nielsen Morten, Ole Lund, NN-align. An artificial neural network-based alignment algorithm for MH class II peptide binding prediction. BMC Bioinformatics. 2009;1: 296.
6. Peters Bjoem. The immune epitope database and analysis resource: from vision to blueprint. Plos Biology. 2005;3(3):91
7. Zhang Hao, Claus Lundegaard, Morten Nielsen. Pan-Specific MHC class I predictors: A benchmark of HLA class I pan-specific prediction methods. Bioinformatics. 2009;25(1):83-89.
8. Karosiene Edita. NetMHCIIpan-3.0, a common pan-specific MHC class I prediction method including the all three human MHC class II isotypes, HLA-DR, and HLA-DQ. Immunogenetics. 2013;10: 711-724.
9. Nielsen M, Lund O, Buus S, Lundegaard C. MHC class II epitope predictive algorithms. Immunology. 2010;130(3):319-28.
10. Lea martin, Russell Spears. Computer-mediated communication, deindividuation and group decision-making. International Journal of Man-machine Studies. 1991; 34(2):283-301.
11. Bergman-Leitner, Elke. Computational and experimental validation of B and T-cell epitopes of the in vivo immune response to a novel malarial antigen. Plos one. 2013;8(8):1610.
12. Singh H, Raghava GP. ProPred: Prediction of HLA-DR binding sites. Bioinformatics. 2001;17(12):1236-7.
13. Karosiene, Edita. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. Immunogenetics. 2013;65(10):711-724.
14. Richard A, Janis Kuby. Basics of Immunology. 1997;1.
15. Huang J, Gutteridge A, Honda W, Kanehisa. MIMOX: A web tool for phage display based epitope mapping. BMC Bioinformatics. 2006;7(1):451.
16. Oany AR, Emran AA, Jyoti TP. Design of an epitope-based peptide vaccine against spike protein of human coronavirus: An in silico approach. Drug Design, Development and Therapy. 2014;8:1139.
17. Wadood A, Ahmed N, Shah L, Ahmad A, Hassan H, Shams S. In-silico drug design: An approach which revolutionarised the drug discovery process. OA Drug Design & Delivery. 2013;1(1):4.
18. Karosiene E, Rasmussen M, Blicher T, Lund O, Buus S, Nielsen M. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. Immunogenetics. 2013;65(10):711-24.
19. Williams TM. Human leukocyte antigen gene polymorphism and the histo-compatibility laboratory. The Journal of Molecular Diagnostics. 2001;3(3):98-104.
20. Tong JC, Tan TW, Ranganathan S. Methods and protocols for prediction of immunogenic epitopes. Briefings in Bioinformatics. 2007;8(2):96-108.

21. Goodswen SJ, Kennedy PJ, Ellis JT. Enhancing in silico protein-based vaccine discovery for eukaryotic pathogens using predicted peptide-MHC binding and peptide conservation scores. PloS One. 2014;9(12):e115745.
22. Trost B, Bickis M, and Kusalik A. Strength in numbers: Achieving greater accuracy in MHC-I binding prediction by combining the results from multiple prediction tools. Immunome Research. 2007;3(1):1.
23. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9(1):1.
24. Gomara MJ, Haro I. Synthetic peptides for the immunodiagnosis of human diseases. Current Medicinal Chemistry. 2007;14(5): 531-46.
25. Yasmin T. Prediction of B and T cell epitope-based peptide vaccines from highly conserved regions in Enterovirus D68 capsid protein VP1: A computational approach.
26. Trainor NB, Crill WD, Roberson JA, Chang GJ. Mutation analysis of the fusion domain region of St. Louis encephalitis virus envelope protein. Virology. 2007;360(2): 398-406.

© 2016 Reddy and Pinnelli; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/16537