Biochem. Cell. Arch. Vol. 17, No. 2, pp. 777-788, 2017 | www.connectjournals.com/bca | *ISSN 0972-5075*

PEPTIDE BASED VACCINE DESIGN FOR THERAPEUTIC INTERVENTION AGAINST HTLV-I : A COMPUTATIONAL APPROACH

Tasneem Bano[1], Salman Akhtar[2], Mohd Haris Siddiqui[2], Jamal M. Arif[1], Mohtashim Lohani[3], Usman Sayeed[2] and M. Kalim A. Khan[2]*

[1]Department of Biosciences, Faculty of Applied Sciences, Integral University, Lucknow - 226 026, India.
[2]Department of Bioengineering, Faculty of Engineering, Integral University, Lucknow - 226 026, India.
[3]Research and Scientific Studies Unit, College of Nursing and Allied Health Sciences, Jazan University, Jazan, KSA.
*e-mail: mkakhaniu@gmail.com

(Accepted 12 October 2017)

ABSTRACT : Human T-cell leukemia virus type 1 (HTLV-1) is a single-stranded diploid RNA genome (9,030–9,040 nucleotides long) belonging to the deltaretrovirus family that causes adult T- cell leukemia (ATL) and HTLV-1–associated myelopathy and tropical spastic paraparesis (HAM/TSP). As there is neither a satisfactory treatment nor a vaccine against this chronic ailment, hence the purpose of the study is to design *in silico* based peptide vaccine, to analyse the oncogenic protein (tax) of different strains of HTLV-1 and to determine the conserved region which is further subjected to predict all putative epitope that can provide therapeutic intervention against HTLV-1. A total of four protein sequences (tax) retrieved from UniProtKB database were aligned to determine the sequence conservancy and to screen the potential epitopes, netCTLpan server, IEDB analysis resource were used. Further, final epitope set were subjected for immunogenicity assessment and molecular docking analysis. Three T-cell epitopes predicted as a peptide vaccine represents high affinity to MHC class I ("RVIGSALQF", "SLLFGYPVY", "NLHLLFEEY") and high population coverage against the whole world. We also predicted B-cell epitope "TPPITHTTPNIPPSFLQAMR" using ABCPred online server. Further, toxicity assessment of resulted epitope set confers that all three T-cell epitope as well as predicted B-cell epitope is non-toxic and non mutant and would be a relevant representative of a large proportion of the human population. However, the results require validation through *in vivo* and *in vitro* study to prove the effectiveness of these predicted epitopes as a peptide vaccine.

Key words : HTLV-I, vaccine, Docking, MHC-I, immunoinformatics.

INTRODUCTION

HTLV-1 is the etiologic agent of adult T cell leukemia (ATL), a progressive lymphoma, and HTLV-1 associated myelopathy/tropical spastic paraparesis (HAM/TSP) (Gallo, 2005; Osame *et al*, 1986; Takatsuki, 2005). Human T cell leukemia virus type 1 was first isolated in 1980 in a T-cell line derived from a patient with cutaneous T-cell lymphoma (Poiesz *et al*, 1980) and was shown to be the etiological agent of adult T cell leukemia (ATL; Hinuma *et al*, 1981; Yoshida *et al*, 1982; Takatsuki, 2005). In 1985, HTLV-1-seropositive patients in French Martinique were diagnosed with a neurodegenerative disorder; termed tropical spastic paraparesis (TSP; Gessain *et al*, 1985). A similar clinical disorder was reported in Japanese patients and was named HTLV-1 associated myelopathy (HAM Osame, *et al,* 1986; Osame and Igata, 1989). The genetic variation among HTLV-1 strains is less than 8%, (Gessain *et al*, 1996; Feuer and Green, 2005). Globally, as many as 20 million people are infected with HTLV-1 (Edlich *et al*, 2000). The virus is endemic in south western Japan, Central Africa, the Caribbean Islands, and Australia (aborigines) as well as in some regions of South America, Melanesia, Middle East, and India. Thus, HTLV-1 prevalence shows a quite strange ethnic distribution (Hiroo, 2012). In these endemic areas, the seroprevalence rates range from 0.1 to 30% (Blattner *et al*, 1982; Saxinger *et al*, 1984; Yanagihara *et al*, 1990; Gotuzzo *et al*, 2000; Sonoda *et al*, 2011). After prolonged latency periods of 40–60 years, approximately 5% of HTLV-1-infected individuals (6.6% of males and 2.1% of females) will developATL (Edlich *et al*, 2000; Proietti *et al*, 2005; Sonoda *et al,* 2011).

The molecular mechanisms associated with driving HTLV-1 into a quiescent versus active replication mode are not clearly understood; however, it is well established that expression of the viral transactivator protein, Tax, is required for the efficient viral gene expression (Divya *et al,* 2015; Brady, 1987; Felber *et al,* 1985; Rosen *et al,* 1985). In ATL, the *tax* gene plays a central role in the proliferation and transformation of HTLV-1-infected cells *in vivo* (Denise *et al*, 2010; Marriott *et al*, 2005). As a result, *tax* (oncogenic protein) is considered a crucial component for the development of peptide based vaccine, providing therapeutic intervention, against HTLV-I.

With the advances in genomics, proteomics, and the understanding of pathogens, the field of viral vaccine preparation has been recently expanded by a most promising approach, known as epitope-based vaccine design (Sharmin *et al*, 2014). Epitope represents the negligible immunogenic region of a protein sequence, which specifically elicits accurate immune responses (Purcell *et al,* 2007). Various studies recently reported that the vaccination process based on epitope efficiently educes defensive immune responses against diverse pathogens (Stanekova *et al*, 2010; He *et al*, 2015; Sominskaya *et al*, 2010). Vaccine production that depends on biochemical experiments can be expensive, time consuming and not always work, although this vaccine formulation of attenuated or inactivated form of microorganism contains a few hundred of unnecessary proteins for the induction of immunity, that may cause allergenic or reactogenic responses (Lo *et al*, 2013; Li *et al*, 2014; Ahmed *et al*, 2017).

Therefore, *in silico* prediction of epitopes of appropriate protein residues would help in production of peptide vaccine with powerful immunogenic and minimal allergenic effect (Purcell *et al,* 2007; Reche *et al,* 2014; Ahmed *et al,* 2017). This is the first study conducted to predict major immunogenic epitopes against HTLV-I using computational approaches.

MATERIALS AND METHODS

Retrieval of Oncogenic Protein Sequence

From the complete proteome of HTLV-I, Oncogenic Protein (*Tax*) (Kannian *et al.,* 2010) sequence of different strains (*Japan ATK-1, Caribbea HS-35, Zaire EL, Melanesia Mel*) were retrieved using UniProt Knowledge Base (UniProtKB) database in FASTA format.The retrieved sequences: P03409|TAX_HTL1A, P14079|TAX_HTL1C, P0C213|TAX_HTL1F and P0C222|TAX_HTL1L, for different strains: *Japan ATK-1, Caribbea HS-35, Zaire EL and Melanesia Mel* respectively, were further subjected to immunogenicity assessment and epitope prediction analysis.

Immunogenicity Prediction of Oncogenic Protein

For the assessment of antigenecity of different strains of HTLV-I, retrieved Oncogenic Protein sequences were submitted to the VaxiJen v2.0 server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html). Using default parameter (threshold 0.4), amino acid sequence in the FASTA Format was uploaded, and virus was selected as target organism. This server works on Auto Cross Covariance (ACC) algorithm that predicts protective antigens, tumor antigens and subunit vaccines with the precision level of 70 to 89% for the discrimination between antigens and non-antigens (Doytchinova *et al.,* 2007).

Multiple Sequence Alignment

Using ClustalW (www.ebi.ac.uk/clustalw) multiple sequence alignment program, best match (conserved regions) for the selected sequences were calculated. ClustalW is a multiple sequence alignment program for DNA or Proteins. It calculates the best matches for the selected sequences & lines them up so that the identities, similarities & differences can be seen.

Epitope Prediction

Most Immunogenic protein sequence was used for the prediction of putative epitope set.

a) B-cell Epitope Prediction

For the prediction of B- cell epitopes, full length of most immunogenic protein sequence of HTLV-I was submitted to ABCPred online server (www.imtech.res.in/abcpred) & window length of 20 amino acids (recommended for humoral immunity) was selected. This server is based on information - processing algorithms inspired by the biological nervous system. The server ranks the epitopes as per their respective scores. The higher score of the peptides means a greater probability of becoming most suitable immunogenic epitope (Saha *et al,* 2006).

b) T-cell Epitope Prediction

Reliable predictions of immunogenic peptides are essential in rational vaccine design and can minimize the experimental effort needed to identify epitopes (Thomas S *et al.,* 2010). CTL epitope were identified using a pan-specific major histocompatibility complex (MHC) class I epitope predictor, NetCTLpan 1.1 online server (http://www.cbs.dtu.dk/services/NetCTLpan/). The method integrates predictions of proteasomal cleavage, transporter associated with antigen processing (TAP) transport efficiency, and MHC class I binding affinity into a MHC class I pathway likelihood score and is an improved and extended version of NetCTL (Thomas S *et al*, 2010). The total score was then calculated by summing up the values of TAP transport efficiency, proteasomal cleavage and MHC-I molecules binding affinity. For our present study, we set the threshold at 0.5, 0.89 and 0.94 for sensitivity and accuracy. The restriction for epitope identification was set to 12 MHC-I super types (Dash *et al*, 2017). Five highest score containing epitopes were selected for further dry laboratory experimentation.

MHC-I binding & Conservancy Analysis of Identified Epitope Set

MHC-I binding & Conservancy analysis of the identified epitopes were then calculated using tools from immune epitope database (IEDB) (Bui *et al,* 2007). These tools calculate the half maximal inhibitory concentration (IC50) value of epitope binding to human leukocyte antigen (HLA) molecules using the stabilized matrix base method (Tenzer *et al,* 2005).

Prior to the run, all the alleles were considered, and the length of the peptides was set at 9.0 (Dash *et al,* 2017) for both frequent and non-frequent allele. For further validation, we only selected the alleles with less than 200 nm binding affinity IC50 (Khan *et al,* 2015). We also implemented the SMM tool to predict these scores for every single selected peptide (Peters *et al,* 2005).

Antigenecity Assessment of Predicted Epitopes

Predicted epitopes (B-cell & T-cell) were evaluated for their antigenecity using VaxiJen v2.0 online server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html).

Virtual scanning of Toxic/Non-toxic peptides

The immunogenic epitopes were subjected to ToxinPred online server for the differentiation of toxic or nontoxic peptides (Gupta *et al*, 2013). The ToxinPred server is based on Support Vector Machine (SVM) and Quantitative Matrix based algorithm, and generates quantitative matrix on the basis of probability or frequency of amino acid at a particular position (Janahi *et al*, 2017).

Determination of Epitope Physio-Chemical Property

In order to make the candidate peptide dispensable, Peptide Property Calculator (https://www.genscript.com) online Server was used, that determine the best solvent for the resulting peptides.

HLA Distribution Analysis

The analysis of population coverage in the context of MHC polymorphism is typical as different HLA type expresses at considerably different frequencies in different ethnicities. Due to the MHC restriction of T cell response, the peptides with more different HLA binding specificities mean more population coverage in defined geographical regions where the peptide-based vaccine might be employed (Jiandong *et al*, 2015). The population coverage rate of individual epitope was calculated using the IEDB population coverage tool (http://tools.immuneepitope.org/tools/population/iedb_input) (Jiandong *et al*, 2015; Bui *et al*, 2006). Predicted epitopes and their corresponding binding HLA alleles were added, and before submission, population coverage area was selected.

3D- structure generation & validation of Potential Epitope and HLA Protein

The 3D- structure of all the identified five epitopes was predicted by PEP-FOLD web server (Thevenet *et al*, 2012). This server predicts five most provable structures for each sequence, the best of which, having the lowest energy model, was chosen for further analysis.

To validate the binding affinity of identified epitope for HLA molecule, we considered the homology modeling as there is no relevant structure available in the protein data bank. We selected homology modeling using the most popular offline tool Modeller 9.10 (*Fiser et al, 2003)*. The protein sequence of HLA - A*32:15 was retrieved from UniProt database (accession id: P10314). For 3D-structure generation, suitable template (PDB ID: 2XPG) was obtained using BLASTp program. Modeller 9.10 generated 3D-structures was further used for validation using PROCHECK online server (Laskowski *et al*, 1996).

Molecular Docking Analysis

Molecular docking analysis was performed using ZDOCK 3.0.2 online server, by considering HLA - A*32:15 molecule as a protein and identified epitopes as ligands. ZDOCK is a rigid-body protein–protein docking program that uses the Fast Fourier Transform algorithm to enable an efficient global docking search on a 3D grid, and utilizes a combination of shape complementarity, electrostatics and statistical potential terms for scoring (Brian *et al*, 2014). ZDOCK achieves high predictive accuracy on protein–protein docking benchmarks, with 470% success in the top 1000 predictions for rigid-body cases in the most recent benchmark version (Pierce *et al,* 2011), and consistent success (acceptable or better predictions for 22 of the last 35 submitted targets) in the international protein– protein docking experiment, Critical Assessment of Predicted Interactions (CAPRI) (Hwang *et al*, 2010; Vreven *et al*, 2013; Wiehe *et al*, 2005).

RESULTS

Immunogenicity Prediction & Sequence Similarity Analysis

As a result of the search of oncogenic protein (*tax*) of different strains of HTLV-I in UniProtKB Database, 4 hits were generated in total. Among them the three sequences: P03409|TAX_HTL1A, P14079|TAX_HTL1C and P0C213|TAX_HTL1F, for *Japan ATK-1, Caribbea HS-35 and Zaire EL* respectively, were found to be immunogenic using VaxiJen server. From multiple

sequence alignment analysis, it is clearly seen that these three sequences were having a close relationship 96% - 99% (Fig. 1a). This result also confers the possibilities of mutation in oncogenic protein of immunogenic strains of HTLV-I (Fig. 1b).

Screening of Most Immunogenic Protein

The capability of each protein sequence to initiate an immune response is indicated by the overall score produced by the specific protein sequence using vaxijen server (Khan *et al,* 2015). The protein sequence having the UniProtKB id: P14079 was found to be most immunogenic protein having a maximum total prediction score of 0.4596 at threshold 0.4. This sequence was selected for further dry laboratory experiments.

Prediction of Potent B-cell & T-cell Epitopes

ABCpred predicts Humoral immunity associated epitopes in descending order of their score, depicting the top to be the best binding affinity. We obtained B-cell epitope "TPPITHTTPNIPPSFLQAMR" at position 91 with binding score 0.997.

On the basis of high combinatorial score, five best T-cell epitopes were selected from the database generated by the NetCTLpan server 1.1. Selected epitopes are represented in Table 1. In combination with several methods such as proteasomal cleavage/transporter associated with antigen processing (TAP)/MHC-I combined predictor, MHC-I processing of the NetCTL server calculates an overall score for each peptide's intrinsic potential from a protein for the designing of T-cell epitope. Peptides with a higher score represent higher processing capabilities (Dash *et al*, 2017).

MHC-I binding & Conservancy analysis

The selected T-cell epitopes were subjected to MHC-I binding and processing prediction using the stabilized matrix base method implemented in IEDB analysis resource tool. In this study, we opted for the selection of the MHC-I molecules with coupled IC50 value less than 200 nm (IC50<200), for which the selected epitopes showed higher affinity. MHC-I processing efficiency tool of IEDB generates an overall score for every epitope accordingly their proteasomal cleavage efficiency, TAP transport efficiency and MHC-I binding efficiency combined. The combined score represented the potentiality of the peptides for presentation, the higher the score, the better they are, and this is the most crucial step for initiating a successful immune response (Khan *et al*, 2015). The scores obtained from IEDB MHC-I binding analysis and processing tools are summarized in Table 2. Identification of epitopes by HLA molecules determines the intensity of immune response. So, a peptide recognized by the highest number of HLA alleles possesses the highest potentiality to initiate a strong immune response. As shown in Table 2, among the 5 epitopes selected, two epitopes "GAFLTNVPY" & "SLLFGYPVY" have interacted with higher number (13) of HLA alleles and other two epitopes "RVIGSALQF" & "NLHLLFEEY" interacted with 11 HLA alleles while one epitope "KYSPFRNGY" interacted with 10 HLA alleles.

Furthermore, better conservancy of epitope is expected for more successful immunization. Epitope conservancy analysis (Table 2) exposed that the epitope "GAFLTNVPY" & "RVIGSALQF" to be 100% conserved whereas epitope "KYSPFRNGY" & "SLLFGYPVY" scored 75% conservancy while "NLHLLFEEY" represents 50% conservancy.

Immunogenicity Assessment of Putative Epitopes

Peptides with strong antigenecity are always better candidates to be B- and T-cell epitopes than those with weak antigenecity. Hence, the assessment of immunogenicity of the identified peptides was performed using the VaxiJen V2.0 server (http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html). This tool predicts the antigenecity of a peptide on the basis of amino acid properties and their positions in the peptide (Janahi *et al*, 2017). The results in Table 3 demonstrates that the identified B-cell epitope "TPPITHTTPNIPPSFLQAMR" is immunogenic in nature with score of 0.7962 and among all the identified T-cell epitopes, two epitopes "KYSPFRNGY" & "GAFLTNVPY" are non immunogenic while three epitopes are "RVIGSALQF", "SLLFGYPVY" & "NLHLLFEEY" are immunogenic.

Toxicity & Solvency Analysis of Final Epitope

The final epitope set (B- and T-cell epitopes) was further tested for the toxicity analysis and medium solvency. Final immunogenic epitopes were found to be non-toxic and were classified as acidic, basic or neutral (Table 3). Acidic peptides were initially dissolved in water, if not, then < 50 µl NH_4OH was added and diluted to 1 ml with deionized water. For basic peptides, 10% and higher solutions of acetic acid was tried; in case not dissolved, then < 50 µl TFA was added and diluted up to 1 ml with deionized water to solubilize the peptide. Neutral peptides required organic solvents like acetonitrile, methanol, or isopropanol (Janahi *et al*, 2017).

Population Coverage

The identified optimum MHC-I binders for each immunogenic epitopes were considered for the cumulative percentage of population coverage analysis. As shown

(a)

SeqA	Name	Len(aa)	SeqB	Name	Len(aa)	Score
1	sp|P0C213|TAX_HTL1F	353	2	sp|P03409|TAX_HTL1A	353	97.0
1	sp|P0C213|TAX_HTL1F	353	3	sp|P14079|TAX_HTL1C	353	96.0
2	sp|P03409|TAX_HTL1A	353	3	sp|P14079|TAX_HTL1C	353	99.0

(b)

```
CLUSTAL 2.0.10 multiple sequence alignment


sp|P03409|TAX_HTL1A      MAHFPGFGQSLLFGYPVYVFGDCVQGDWCPISGGLCSARLHRHALLATCPEHQITWDPID 60
sp|P14079|TAX_HTL1C      MAHFPGFGQSLLFGYPVYVFGDCVQGDWCPISGGLCSARLHRHALLATCPEHQITWDPID 60
sp|P0C213|TAX_HTL1F      MAHFPGFGQSLLFGYPVYVFGDCVQGDWCPISGGLCSARLHRHALLATCPEHQITWDPID 60
                         ************************************************************

sp|P03409|TAX_HTL1A      GRVIGSALQFLIPRLPSFPTQRTSKTLKVLTPPITHTTPNIPPSFLQAMRKYSPFRNGYM 120
sp|P14079|TAX_HTL1C      GRVIGSALQFLIPRLPSFPTQRTSKTLKVLTPPITHTTPNIPPSFLQAMRKYSPFRNGYM 120
sp|P0C213|TAX_HTL1F      GRVIGSALQFLIPRLPSFPTQRTSKTLKVLTPPTTHTTPNIPPSFLQAMRKYSPFRNGYM 120
                         ********************************* ************************

sp|P03409|TAX_HTL1A      EPTLGQHLPTLSFPDPGLRPQNLYTLWGGSVVCMYLYQLSPPITWPLLPHVIFCHPGQLG 180
sp|P14079|TAX_HTL1C      EPTLGQHLPTLSFPDPGLRPQNLYTLWGGSVVCMYLYQLSPPITWPLLPHVIFCHPGQLG 180
sp|P0C213|TAX_HTL1F      EPTLGQHLPTLSFPDPGLRPQNLYTLWGGSVVCMYLYQLSPPITWPLLPHVIFDHPDQLG 180
                         ************************************************** **.***

sp|P03409|TAX_HTL1A      AFLTNVPYKRIEELLYKISLTTGALIILPEDCLPTTLFQPARAPVTLTAWQNGLLPFHST 240
sp|P14079|TAX_HTL1C      AFLTNVPYKRIEKLLYKISLTTGALIILPEDCLPTTLFQPARAPVTLTAWQNGLLPFHST 240
sp|P0C213|TAX_HTL1F      AFLTNVPYKRMEELLYKISLTTGALIILPEDCLPTTLFQPARAPVTLTAWQSGLLPFHST 240
                         **********:*:****************************************.******** 

sp|P03409|TAX_HTL1A      LTTPGLIWTFTDGTPMISGPCPKDGQPSLVLQSSSFIFHKFQTKAYHPSFLLSHGLIQYS 300
sp|P14079|TAX_HTL1C      LTTPGLIWTFTDGTPMISGPCPKDGQPSLVLQSSSFIFHKFQTKAYHPSFLLSHGLIQYS 300
sp|P0C213|TAX_HTL1F      LTTPGLIWAFTDGTPMISGPCPKDGRPSLVLQSSSFIFHKFQTKAYHPSFLLSHGLIQYS 300
                         ********:***************:************************************

sp|P03409|TAX_HTL1A      SFHSLHLLFEEYTNIPISLLFNEKEADDNDHEPQISPGGLEPPSEKHFRETEV 353
sp|P14079|TAX_HTL1C      SFHNLHLLFEEYTNIPISLLFNEKEADDNDHEPQISPGGLEPLSEKHFRETEV 353
sp|P0C213|TAX_HTL1F      SFHNLHLLFEEYTNIPISLLFNEKEADDTDHEPQVSPGGLEPPSEKHFRETEV 353
                         ***.**************************.*****:******** **********
```

Fig. 1 : Multiple Sequence Alignment of different immunogenic strains of HTLV-I.
(a) Sequence similarity percentage among immunogenic strains. **(b)** Amino acid sequence alignment of *tax* protein (353 length) of three strain of HTLV-I. Alignment sequence characters are indicated as follows: asterisk (*) indicates identical amino acids, colon (**:**) indicates similar amino acids, and period (**.**) indicates nearly similar ones.

Table 1 : Five most potent T-cell epitopes, according to the overall score by the NetCTLpan 1.1 server.

Number	Epitopes	Overall Score (nM)
1	KYSPFRNGY	**3.74428**
2	GAFLTNVPY	**3.67548**
3	RVIGSALQF	**3.62894**
4	SLLFGYPVY	**3.60945**
5	NLHLLFEEY	**3.55031**

in Table 4, the population coverage for South Africa was found to be 66.72%; in West and North Africa, it was 57.89% & 55.32%, respectively and for Central Africa it was observed to be 55.32%. The population coverage for East Asian region was recorded at 66.44%. For Europe, the population coverage was found to be 60.85%; and in North America, it was recorded 60.56%.

Structure Validation & Docking Analysis

The hypothetical structure of HLA - A*32:15 protein generated through homology technique was validated through PROCHECK online server (Laskowski *et al*, 1996). As shown in Fig. 2, Ramachandran Plot generated by PROCHECK server showed that about 98.6% of residues of the protein are located in the favoured regions, as against 0% in the outlier region and 1.4% in the generously allowed region.

Protein-peptide docking analysis revealed that the proposed epitopes bound in the cleft of the HLA - A*32:15 (Fig. 3), where the highest binding affinity was 1584 (Table 5), observed for the epitope “SLLFGYPVY”. Discovery Studio 3.1 was used to analyse the intermolecular Hydrogen Bond and Hydrogen bond distance & to design

Table 2 : Five most potent T-cell epitopes with interacting MHC-I alleles, total score and epitope conservancy analysis result.

No.	Epitope	Interacting MHC-I allele with an affinity <200 nM	TOTAL score (Proeasome score, TAP Score, MHC score, Processing Score)	MHC-I binding (IC 50<200nm)	Epitope conservancy (%)
1	**KYSPFRNGY**	HLA-A*32:07	1.9	6.2	
		HLA-A*68:23	1.32	23.16	
		HLA-A*30:02	0.86	67.52	75.00
		HLA-A*32:15	0.43	182.92	
		HLA-B*27:20	2.02	4.68	
		HLA-B*40:13	1.27	26.24	
		HLA-C*12:03	1.41	19.11	
		HLA-C*07:02	1.31	23.83	
		HLA-C*14:02	1.19	31.31	
		HLA-C*03:03	0.5	155.1	
		10			
2	**GAFLTNVPY**	HLA-A*29:02	0.21	176.51	
		HLA-A*68:23	1.25	15.99	
		HLA-A*32:07	1.02	27.23	100.00
		HLA-A*32:15	0.28	149.37	
		HLA-B*27:20	1.23	16.58	
		HLA-B*15:03	0.87	38.64	
		HLA-B*15:17	0.77	47.94	
		HLA-B*35:01	0.63	67.08	
		HLA-B*15:02	0.17	190.65	
		HLA-B*40:13	1.17	19.36	
		HLA-C*12:03	1.87	3.87	
		HLA-C*03:03	1.09	23.1	
		HLA-C*14:02	0.34	128.71	
		13			
3	**RVIGSALQF**	HLA-A*32:07	1.4	16.67	
		HLA-A*68:23	1.4	16.74	
		HLA-A*32:15	1.14	30.08	100.00
		HLA-A*32:01	0.39	170.98	
		HLA-B*15:17	1.8	6.54	
		HLA-B*15:03	1.04	38.11	
		HLA-B*27:20	0.58	108.77	
		HLA-B*15:01	0.52	127.06	
		HLA-B*58:01	0.73	77.32	
		HLA-C*03:03	0.97	44.32	
		HLA-C*12:03	0.65	93.59	
		11			
4	**SLLFGYPVY**	HLA-A*29:02	0.95	66.03	
		HLA-A*30:02	0.78	97.38	
		HLA-A*32:07	1.82	8.91	75.00
		HLA-A*68:23	1.54	16.74	
		HLA-A*32:15	1.01	57.58	
		HLA-B*27:20	1.73	10.83	
		HLA-B*15:03	0.93	67.93	
		HLA-B*15:02	0.84	84.18	
		HLA-B*15:01	0.72	110.15	
		HLA-B*40:13	0.53	170.58	
		HLA-C*03:03	1.24	33.62	
		HLA-C*14:02	1.07	49.05	
		HLA-C*12:03	0.99	60.29	
		13			

Table 2 continued...

Table 2 continued...

5	**NLHLLFEEY**	HLA-A*32:07	1.34	24.66	
		HLA-A*68:23	1.32	25.87	
		HLA-A*32:15	1.04	48.57	50.00
		HLA-A*29:02	0.59	138.6	
		HLA-A*80:01	0.53	158.86	
		HLA-B*27:20	1.84	7.77	
		HLA-B*15:02	0.69	110.72	
		HLA-B*40:13	0.57	146.19	
		HLA-C*14:02	1.17	36.19	
		HLA-C*12:03	0.99	55.36	
		HLA-C*03:03	0.79	87.42	
		11			

Table 3 : Prediction of most immunogenic, non-mutant and non-toxic B-cell and T-cell epitopes.

No.	Epitope	Immunogenicity			Toxicity	Attribute
		Threshold	VaxiJen Score	Antigen/Non-Antigen		
	B-cell Epitope					
	TPPITHTTPNIPPSFLQAMR	0.4	**0.7962**	Probable **ANTIGEN**	Non-Toxic	Basic
	T-cell Epitope					
1	KYSPFRNGY	0.4	**0.3940**	**NON-ANTIGEN**	Non-Toxic	Basic
2	GAFLTNVPY	0.4	**0.1769**	**NON-ANTIGEN**	Non-Toxic	Neutral
3	**RVIGSALQF**	0.4	**0.4657**	Probable **ANTIGEN**	Non-Toxic	Basic
4	**SLLFGYPVY**	0.4	**0.5040**	Probable **ANTIGEN**	Non-Toxic	Neutral
5	**NLHLLFEEY**	0.4	**0.7284**	Probable **ANTIGEN**	Non-Toxic	Acidic

Table 4 : Population Coverage Analysis for the Potential Epitope Set against HTLV-I

Population/area	Coverage (%)[a]	Average Hit[b]	PC90[c]
East Asia	66.44%	2.34	0.30
Northeast Asia	66.09%	2.11	0.29
South Asia	52.97%	1.51	0.21
Southeast Asia	60.72%	1.55	0.25
Southwest Asia	44.95%	1.22	0.18
Europe	60.85%	2.01	0.26
East Africa	53.78%	1.52	0.22
West Africa	57.89%	1.43	0.24
Central Africa	55.32%	1.57	0.22
North Africa	55.32%	1.70	0.22
South Africa	66.72%	2.36	0.30
West Indies	44.77%	0.95	0.18
North America	60.56%	1.77	0.25
Central America	3.55%	0.05	0.10
South America	48.72%	1.21	0.20
Oceania	40.07	1.23	0.17
Average (Standard deviation)	52.42% (**14.81%**)	1.53 (**0.55**)	0.22 (**0.05**)

Note: [a] Projected population coverage.

[b] Average number of epitope hits/HLA combinations recognized by the population.

[c] Minimum number of epitope hits/HLA combinations recognized by 90% of the population.

Table 5 : zdock score of Protein-Peptide docking analysis.

Receptor	Peptide	Z-Dock Score
HLA-A*32:15	KYSPFRNGY	1488
HLA-A*32:15	GAFLTNVPY	1380
HLA-A*32:15	RVIGSALQF	1235
HLA-A*32:15	SLLFGYPVY	1584
HLA-A*32:15	NLHLLFEEY	1225

the docked HLA - A*32:15-epitope complexes, as shown in Fig. 3.

DISCUSSION

In the present study, primary focus is on the screening of putative epitope from oncogenic protein (*tax*) of HTLV-I, providing therapeutic intervention using *in silico* approach. The central role of tax in the pathogenesis of ATL (Adult T-cell Lymphoma) makes it an appropriate immunogenic target for the development of vaccine. Computational based screening of candidate epitope is efficient and cost effective in the context of HLA molecules, therefore the study incorporated immunoinformatics, molecular modelling and molecular interaction tools to predict putative epitopes.

Initially, oncogenic protein sequence (*tax*) from different variants of HTLV-I has been submitted for their immunogenicity assessment and further, only immunogenic sequences were subjected to perform sequence alignment

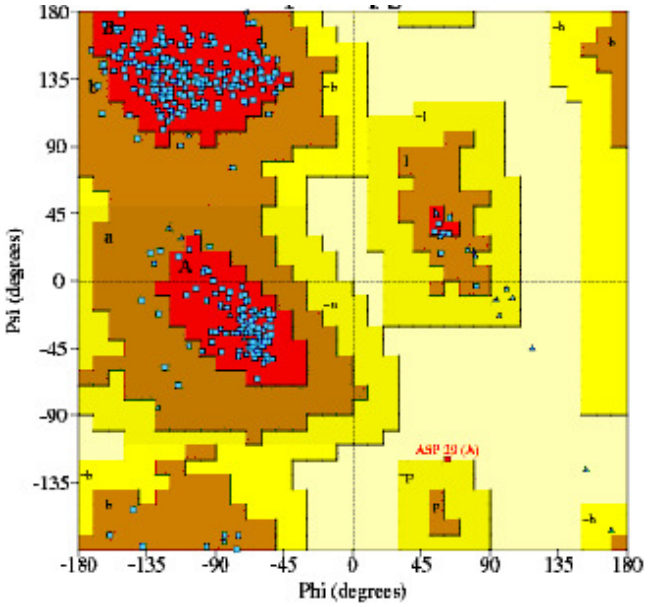
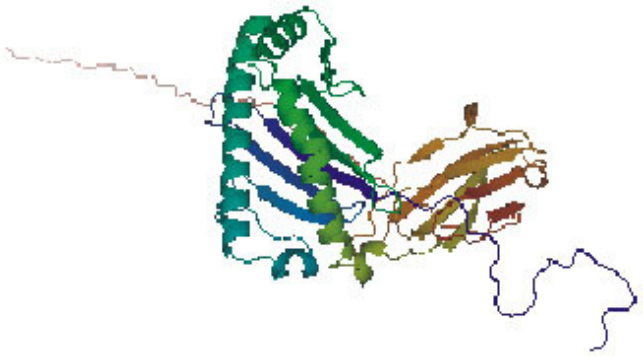


Fig. 2 : (a) Ramachandran plot generated by PROCHECK online server for the validation of hypothetical structure of HLA-A*32:15 protein generated through homology technique.



Three – dimensional structure of the final model of HLA - A*32:15 designed through PyMol.

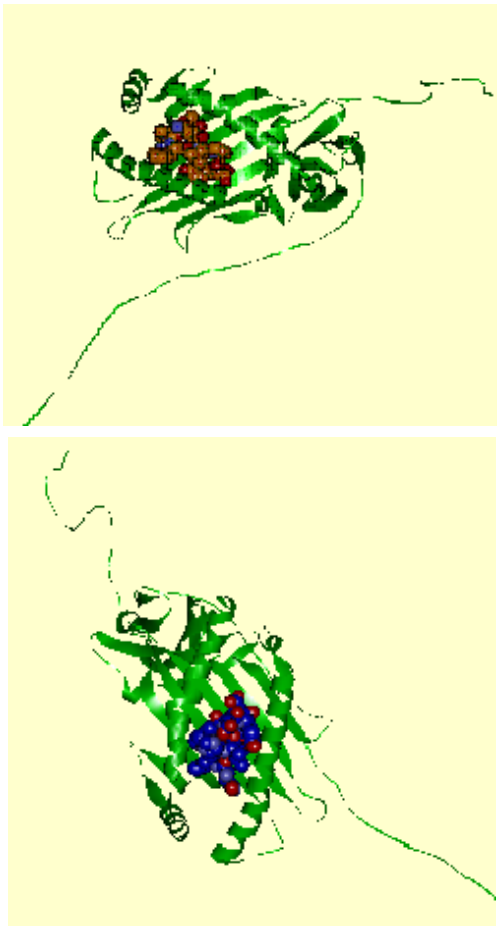


Fig. 3 : Visualization of docked complexes obtained from molecular docking (Zdock) analysis.
The complexes are representing the interaction of HLA - A*32:15 with (a) KYSPFRNGY, (b) GAFLTNVPY, (c) RVIGSALQF, (d) SLLFGYPVY, (e) NLHLLFEEY.

analysis. Finding revealed that there is 96%-99% conservation for the amino acid sequences of different strains within the species. From these sequences, most antigenic protein sequence was determined by VaxiJen server. Based on auto cross covariance (ACC), the VaxiJen server transform the protein sequence into uniform vectors of physicochemical properties of proteins. With 91% sensitive, 82% accuracy and 72 specificity, the l00-CV (leave one–out cross validation) was used to identify antigenecity of protein for viral species (Verma *et al*, 2015). The resultant most antigenic protein (VaxiJen score ≥0.4) was then subjected for onward immunoinformatics study.

Not all possible peptides originating from cell proteins will presented by MHC class I. In fact, it is estimated that only one out of 2,000 potential peptides will be immunodominant (Yewdell *et al*, 1999). One of the first steps involved in MHC class I antigen presentation is the degradation of intracellular proteins, including proteins from the cytoplasm and nucleus, by the proteasome (Larsen *et al*, 2007). These peptides may be trimmed at the N-terminal end by cytosolic exopeptidases (Levy *et al*, 2002). A subset of the peptides is transported by transporter associated with antigen processing (TAP) complex into the endoplasmic reticulum (ER), where further N terminal trimming occurs (Schatz *et al*, 2008). Inside the ER, a peptide may bind to an MHC class I molecule and the peptide–MHC complex will be transported to the cell surface, where it subsequently may be recognized by CTLs. These successive steps from protein to ligand presented on the cell surface are limiting the number of possible epitopes (Thomas *et al*, 2010). The predictive performance of the NetCTLpan method is validated on large and MHC diverse data sets derived from the SYFPEITHI (Rammensee *et al*, 1999) and Los Alamos HIV databases (http://www.hiv.lanl.gov/) and its performance has been compared to other state-of-the-art CTL epitope prediction methods. Hence, we predicted five potent 9-mer epitopes from NetCTLpan 1.1 server. Using the threshold of 0.5, the NetCTL 1.2 server predicts maximum number of epitopes without compromising the specificity or sensitivity levels, covering all 12 MHC class I supertypes (Dash *et al*, 2017). The five most potent epitopes with scores are represented in Table1. The most restricting step in antigen presentation is peptide binding to MHC class I molecule (Yewdell *et al*, 1999). For MHC-I binding prediction, peptides with IC50 values <50 nM are considered high affinity, <500 nM for intermediate affinity, and <5000 nM for low affinity. Therefore, we selected maximum alleles (Table 2) having binding affinity <200 nM (Mostafa *et al*, 2016). It is advocated that T-cell epitope binding to specific multiple HLA super types are termed as promiscuous in vaccine design, since they effectively increase the coverage of higher proportions of human populations (Wilson *et al,* 2003; Vani *et al*, 2006). According to the results, the epitope "SLLFGYPVY" bind to the highest number (13) of alleles (Table 2).

Further we checked the immunogenicity of predicted epitope using vaxiJen online server and result shows that only three epitope "RVIGSALQF", "SLLFGYPVY", "NLHLLFEEY" were found to be immunogenic, while two "KYSPFRNGY", "GAFLTNVPY" are non-immunogenic, therefore, can't be used as vaccine candidate (Table 3).

Now, obtained immunogenic epitopes "RVIGSALQF", "SLLFGYPVY", "NLHLLFEEY" were used for conservancy analysis. According to the results, epitope "RVIGSALQF" represents 100% conservancy while "SLLFGYPVY" and "NLHLLFEEY" are 75% and 50% conserved (Table 2).

Protein–protein interactions are responsible for wide range of key biological processes such as cell signaling, enzyme inhibition and immune recognition. A high-resolution structure of a protein complex provides the atomic details of the interaction, and enables rational design of therapeutic molecules that inhibit or improve binding (Brian *et al,* 2014). Hence, we further validated each immunogenic epitope by molecular docking studies with HLA - A*32:15 protein, as it was common in the results from MHC-I binding interaction analysis. Prior docking experiment, 3D- structure of epitope & HLA molecule was generated through PEP-FOLD server and homology modelling technique (modeller v9.10) respectively. Molecular docking was performed online tool zdock 3.0.2. The predictions are sorted by zdock score, Table 5, representing that the epitope "SLLFGYPVY" has highest binding affinity; 1584.

An important consideration in vaccine development is that the HLA distribution varies according to the diverse ethnic groups & geographic regions around the world. Hence, during the development of an effective design, wide range of population coverage must be considered. The results from population coverage analysis, the predicted epitopes showed wide range of population coverage in different regions of the world (Table 4), indicating that the epitopes will specifically bind with the prevalent HLA molecules in the target population, where the vaccine will be employed.

Prediction of B-cell epitope through ABCPred online

server generated epitope "TPPITHTTPNIPPSFLQAMR" at position 91 with binding score 0.997. Obtained B-cell epitope were found to be immunogenic having vaxiJen score 0.7962, as shown in Table 3.

Further toxicity assessment of final epitope set declares that all three CTL epitope as well as predicted B-cell epitope is non-toxic also (Table 3) and would be a relevant representative of a large proportion of the human population.

CONCLUSION

The development of effective therapeutics against HTLV-I causing Lymphoma in human being remains a high priority till today. The application of *in silico* approaches are helpful to speed up the *in vitro, in vivo* and *ex vivo* studies by cutting down the large epitope dataset into a smaller one that needs to be validated by the experimental analysis and prompt the future vaccine development. Keeping these facts in mind, we focused to provide therapeutic intervention against HTLV-I, by screening the most potent epitopes, which could be a breakthrough for the development of new therapeutic and diagnostic tools. The data obtained from present study might prove a milestone in the development of therapeutic vaccine; however, the potential of the screened epitopes must be subjected to *in vitro* and *in vivo* analysis for the assessment of the current study.

ACKNOWLEDGEMENT

The authors graceful to Integral University, Lucknow, U.P., India for providing the necessary facilities for the study.The manuscript has been approved by competent authority(Assigned Manuscript Communication Number-IU/R&D/2017-MCN000195.

REFERENCES

Ahmed H A, Khoubieb A A, Mojahid S I, Waleed H H, Mohammed S M, Marwan M B and Mohamed A S (2017) Multi Epitope Peptide Vaccine Prediction against Sudan Ebola Virus Using Immuno-Informatics Approaches. *Advanced Techniques in Biology & Medicine*. 5(1) doi: 10.4172/2379-1764.1000203

Blattner W A, Kalyanaraman V S, Robert G M, Lister T A, Galton D A, Sarin P S, Crawford M H, Catovsky D, Greaves M and Gallo R C (1982) The human type-C retrovirus, HTLV, in Blacks from the Caribbean region, and relationship to adult T-cell leukemia/lymphoma. *Int. J. Cancer* **30**, 257–264.

Brady J, Jeang K T, Duvall J and Khoury G (1987). Identification of p40x-responsive regulatory sequences within the human T-cell leukemia virus type I long terminal repeat. *J. Virol*. **61**, 2175–81.

Brian G P, Kevin W, Howook H, Bong H K, Thom V and Zhiping W (2014) ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. *Structural bioinformatics*. 30 (12): 1771–1773.

Bui H H, Sidney J, Dinh K, Southwood S, Newman M J and Sette A (2006) Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*. 7:153. PMID: 16545123.

Bui H H, Sidney J, Li W, Fusseder N and Sette A (2007). Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics*. 8:361.doi:10.1186/1471-2105-8-361.

Dash R, Junaid M, Forhad M, Chowdhury A, Islam A and Hosen S M Z (2017) *In silico*-based vaccine design against Ebola virus glycoprotein. *Advances and Applications in Bioinformatics and Chemistry*. 10:11–28.

Denise U, Goncalves¸ Fernando A P, Joao G R R, Marcelo G A, Sonia RP, Antonio C, Guedes and Anna Barbara F C P (2010) Epidemiology, Treatment, and Prevention of Human T-Cell Leukemia Virus Type 1-Associated Diseases. *Clin. Microbiol. Rev.* **23**(3), 577–589 doi:10.1128/CMR.00063-09.

Divya S, Mohit S, Alfonso L, Pooja J and Zafar K K (2015) Myocyte enhancer factor (MEF)-2 plays essential roles in T-cell transformation associated with HTLV-1 infection by stabilizing complex between Tax and CREB. *Retrovirology* **12**, 23

Doytchinova I A and Flower D R (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* **8**, 4. [PMID: 17207271].

Edlich R F, Arnette J A and Williams F M (2000) Global epidemic of human T-cell lymphotropic virus type-I (HTLV-I). *J Emerg Med*. **18**, 109–19.

Felber B K, Paskalis H, Kleinman-Ewing C, Wong-Staal F and Pavlakis G N (1985) The pX protein of HTLV-I is a transcriptional activator of its long terminal repeats. *Science* **229**, 675-679.

Feuer G and Green P L (2005) Comparative biology of human T- cell lympho tropic virus type-1 (HTLV-1) and HTLV-2. *Oncogene* **24**, 5996–6004.

Fiser A and Sali A (2003) Modeller: generation and refinement of homology-based protein structure models. Methods in Enzymology. **374**, 461–91. doi:10.1016/S0076-6879(03)74020-8. PMID 14696385.

Gallo R C (2005) The discovery of the first human retrovirus: HTLV-1 and HTLV-2. *Retrovirology* **2**, 17.

Gessain A, Barin F, Vernant J C, Gout O, Maurs L, Calender A and de The G (1985) Antibodies to human T-lymphotropic virus type-I in patients with tropical spastic paraparesis. *Lancet* **2**, 407–410.

Gessain A, Mahieux R and De TG (1996) Genetic variability and molecular epidemiology of human and simian T-cell leukemia/lymphoma virus type-I. J. Acquir. Immune Defic. Syndr. Hum. *Retrovirol.* **13**, S132– S145.

Gotuzzo E, Arango C, deq Ueiroz C A and Isturiz R E (2000) Human T-cell lymphotropic virus-I in Latin America. Infect. Dis. Clin. North Am. 14:211–239.

Gupta S, Kapoor P, Chaudhary K, Gautam A and Kumar R (2013). In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE*. 8:e73957.

He L, Cheng Y and Kong L (2015) Approaching rational epitope vaccine design for hepatitis C virus with meta-server and multivalent scaffolding. *Sci. Rep*. 5.

Hinuma Y, Nagata K, Hanaoka M, Nakai M, Matsumoto T, Kinoshita KI, Shirakawa S and Miyoshi I (1981) Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the

antigenin human sera. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6476–6480.

Hiroo H (2012) Cellular factors involved in HTLV-1entry and pathogenicity. *Frontiers in Microbiol.* **3**, 222

Hwang H, Thom V, Brian P, Jui-Hung H, and Zhiping W (2010) Performance of ZDOCK and ZRANK in CAPRI rounds 13-19. *Proteins* **78**, 3104–3110.

Janahi EM, Dhasmana A, Srivastava V and Haque S (2017) *In silico* CD4+ , CD8+ T- cell and B-cell immunity asociated immunogenic epitope predictionabd HLA distribution analysis of ZIKA virus. *EXCLI Journal* **16**, 63-72.

Jiandong S, Jing Z, Sijin L, Jing S, Yumei T, Meini W, Jianfan L, Yanhan L, Ningzhu H, Haixuan W and Yunzhang H (2015) Epitope-Based Vaccine Target Screening against Highly Pathogenic MERS-CoV: An In Silico Approach Applied to Emerging Infectious Diseases. *PLoS ONE* **10**(12), e0144475. doi:10.1371.

Kannian P and Patrick LG (2010) Human T Lymphotropic Virus Type 1 (HTLV-1): Molecular Biology and Oncogenesis. *Viruses.* **2**, 2037-2077. doi:10.3390/v2092037

Khan M A, Hossain M U, Rakib-Uz-Zaman S M and Morshed M N (2015) Epitope-based peptide vaccine design and target site depiction against Ebola viruses: an immunoinformatics study. *Exp. Immunol.* doi: 10.1111/sji.12302

Larsen M V, Lundegaard C, Lamberth K, Buus S, Lund O and Nielsen M (2007) Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics* **8**, 424

Laskowski R A, Rullmannn J A, MacArthur M W, Kaptein R and Thornton J M (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**(4), 477–486.

Levy F, Burri L, Morel S, Peitrequin A, Lévy N, Bachi A, Hellman U, Van Den Eynde BJ and Servis C (2002) The final N-terminal trimming of a subaminoterminal proline-containing HLA class Irestricted antigenic peptide in the cytosol is mediated by two peptidases. *J. Immunol.* **169**, 4161–4171.

Li W, Joshi M D, Singhania S, Ramsey K H and Murthy A K (2014) Peptide vaccine: Progress and challenges. *Vaccines* **2**, 515-536.

Lo YT, Pai TW, Wu WK and Chang HT (2013) Prediction of conformational epitopes with the use of a knowledge-based energy functions and geometrically related neighbouring residue characteristics. *BMC Bioinformatics* **14**, S3.

Marriott S and Semmes O J (2005) Impact of HTLV-1 Tax on cell cycle progression and the cellular DNA damage repair response. *Oncogene* **24**, 5986–5995.

Mostafa S M and Islam A B (2016) An *in silico* approach predicted potential therapeutics that can confer protection from maximum pathogenic Hantaviruses. *Future Virology* **11**(6), 411–428.

Osame M and Igata A (1989) The history of discovery and clinic-epidemiology of HTLV-I-associated myelopathy (HAM). Jpn.J.Med. 28:412–414. Osame M, Usuku K, Izumo S, Ijichi N, Amitani H, Igata A, Matsumoto M and Tara M (1986). HTLV-I associated myelopathy a new clinical entity. *Lancet* **1**, 1031–1032.

Osame M, Usuku K, Izumo S, Ijichi N, Amitani H and Igata A (1986) HTLV-I associated myelopathy, a new clinical entity. *Lancet.* **1**,1031–2.

Peters B and Sette A (2005) Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics* **6**, 132.

Pierce B G, Hourai Y and Weng Z (2011) Accelerating Protein Docking in ZDOCK Using an Advanced 3D Convolution Library. *PLoS ONE* **6**(9): e24657. doi:10.1371/journal.pone.0024657

Poiesz B J, Ruscetti F W, Gazdar A F, Bunn P A, Minna J D and Gallo R C (1980) Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7415–7419.

Proietti F A, Carneiro-Proietti A B, Catalan-Soares B C and Murphy E L (2005) Global epidemiology of HTLV-I infection and associated diseases. *Oncogene* **24**, 6058–6068.

Purcell A W, McCluskey J and Rossjohn J (2007) More than one reason to rethink the use of peptides in vaccine design. *Nat. Rev. Drug Discov.* **6**(5), 404–414.

Rammensee H, Bachmann J, Emmerich NP, Bachor O and Stevanoviæ S (1999) SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **50**, 213–219

Reche P A, Darren R F, Masha F H and Yoshihiko H (2014). Peptide-based immunotherapeutics and vaccines. *J. Immunol Res.* **25**, 6784.

Rosen C A, Sodroski J G and Haseltine W A (1985) Location of cis-acting regulatory sequences in the human T-cell leukemia virus type I long terminal repeat. *Proc Natl Acad Sci* U S A. **82**, 6502-6506.

Saha S and Raghava G P S (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* **65**, 40-48.

Saxinger W, Blattner WA, Levine PH, Clark J, Biggar R, Hoh M, Moghissi J, Jacobs P, Wilson L, Jacobson R, Crookes R and Strong M, Angari AA, Dean AG, Nkrumah FK, Mourali N and Gallo R C (1984) Human T-cell leukemia virus (HTLV-I) antibodies in Africa. *Science* **225**,1473–1476.

Schatz MM, Peters B, Akkad N, Ullrich N, Martinez AN, Carroll O, Bulik S, Rammensee H, van Endert P, Holzhütter H, Tenzer S and Schild H (2008) Characterizing the N-terminal processing motif of MHC class I ligands. *J. Immunol.* **180**, 3210–3217.

Sharmin R and Islam A B (2014) A highly conserved WDYPKCDRA epitope in the RNA directed RNA polymerase of human coronaviruses can be used as epitope-based universal vaccine design. *BMC Bioinformatics* **15**(161), 1471–2105.

Sominskaya I, Skrastina D and Dislers A (2010) Construction and immunological evaluation of multivalent hepatitis B virus (HBV) core virus-like particles carrying HBV and HCV epitopes. *Clin. Vaccine Immunol.* **17**(6),1027–1033.

Sonoda S, Li HC and Tajima K (2011) Ethno epidemiology of HTLV-1 relateddiseases: ethnic determinants of HTLV-1 susceptibility and its world wide dispersal. *Cancer Sci.* **102**, 295–301.

Stanekova Z and Vareckova E (2010) Conserved epitopes of influenza A virus inducing protective immunity and their prospects for universal vaccine development. *Virology* **7**(1), 1.

Takatsuki K (2005) Discovery of adult T-cell leukemia. *Retrovirology* **2**, 16.

Tenzer S, Peters B and Bulik S (2005) Modeling the MHC class I pathway by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. *Cell Mol. Life Sci.* **62**(9), 1025–1037.

Thevenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P and Tuffery P (2012) PEPFOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Res*. **40**, 11.

Thomas S, Mette VL, Claus L and Morten N (2010) NetCTLpan: pan-specific MHC class I pathway epitope predictions. *Immunogenetics* **62**, 357–368. DOI 10.1007/s00251-010-0441-4.

Vani J, Shaila M, Chandra N and Nayak R (2006) A combined immuno-informatics and structure-based modeling approach for prediction of T cell epitopes of secretory proteins of Mycobacterium tuberculosis. *Microbes Infect.* **8**(3), 738–746.

Verma S K, Yadav S and Kumar A (2015) *In silico* prediction of B- and T- cell epitope on Lassa virus proteins for peptide based subunit vaccine design. *Advanced Biomedical Res*. **4**, 201.

Vreven T, Brian G P, Howook H and Zhiping W (2013) Performance of ZDOCK in CAPRI rounds 20-26. *Proteins* **81**, 2175–2182.

Wiehe K, Brian P, Julian M, Wei WT, Robert A, Rong C and Zhiping W (2005) ZDOCK and RDOCK performance in CAPRI rounds 3, 4, and 5. *Proteins* **60**, 207–213.

Wilson C C, McKinney D and Anders M (2003) Development of a DNA vaccine designed to induce cytotoxic T lymphocyte responses to multiple conserved epitopes in HIV-1. *J. Immunol*. **171**(10), 5611–5623.

Yanagihara R, Jenkins C L, Alexander S S, Mora CA and Garruto R M (1990) Human T-lym- photropic virus type-I infection in Papua New Guinea: high prevalence among the Hagahai confirmed by western analysis. *J. Infect.Dis*. **162**, 649–654.

Yewdell J W and Bennink J R (1999) Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Ann. Rev. Immunol*. **17**,51–88

Yoshida M, Miyoshi I and Hinuma Y (1982) Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2031–2035.