Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2014, 6(11):500-506

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Structural insights of CD_4+ Gag epitopes and HLA-DRB1*10 allele complexes

Jemmy Christy. H[1*] and Alex Anand D.[2]

[1]*Department of Bioinformatics, Sathyabama University, Chennai*
[2]*Department of Biomedical Engineering, Sathyabama University, Chennai*

ABSTRACT

*An efficient vaccine which would stimulate both humoral and cell mediated immune response is the ideal solution for HIV/AIDS problem. CD4+ T cells play a significant role in induction and maintenance of CD8+ T cell and antibody-producing B cell responses their by aid a tremendous role HIV disease control. Our study based on immunoinformatics approach focus on the prediction of HLA-DRB1*10 allele specific epitopes capable of triggering immunogenic activity. IEDB method was adopted to identify Gag epitopes vaccine candidates* ***MYSPISILDIKQGPK-P1, RMYSPISILDIKQGP-P2, YSPISILDIKQGPKE-P3, VPVGEIY KRWIILGL-P4, PVGEIYKRWIILGLN-P5*** *restricted to HLA-DRB1*10 allele would aid significant CD4+ T cell immune response against HIV infection and population coverage among south Indian population were assessed. Three dimensional structures of epitopes P1-P5 modeled using I Tasser, and their insights of binding affinity towards HLA binding groove analyzed by ClusPro based docking studies resulted with conventional hydrogen bonding. Thus the interaction between the screened P1-P5 epitopes and DRB1*10 alleles exhibiting stability and would aid immune response. Current data would provide the insights for design and development of novel GAG based vaccine candidate against HIV infection*

Keywords: HLA-DRB1*10, Gag epitope , Vaccine, CD4+ T cell response, Human Leukocyte Antigen, Binding affinitity,$IC_{50.}$

INTRODUCTION

Acquired immunodeficiency syndrome (AIDS) a major health concern is caused by *Human immunodeficiency virus* (HIV), a retrovirus that belongs to the Lentiviridae family. The Human Immunodeficiency Virus infection implements progressive decline in immunological response and leading to Acquires Immunodeficiency Syndrome (AIDS) claiming morbidity and mortality. Currently it is estimated that in worldwide 75 million people living with HIV-1 [1].Recent report of 20th International AIDS conference revealed that nearly 5 million people living with HIV in Asia and the Pacific [2].Human immunodeficiency virus-1 (HIV-1) has infected more than 75 million people and caused nearly 36 million deaths worldwide[3]. Nearly 20.89 lakh people are said to be HIV carriers in India based on National AIDS control organization recent report (2013-14) [4].HAART treatment has had a profound impact on the AIDS epidemic, it should be understood that the HAART treatment is not a cure for HIV and carries its own drawbacks [5].Current research focusing on subunit vaccine especially peptide based vaccine which would aid induction of specific cellular immune response thus reduce viral load [6]. Nine subtypes (A–D, F-H, J and K) and 61 circulating recombinant forms (CRFs) existing in M group a most prominent HIV-1 group in the world [7].C clade of M group is given global priority due to its prevalence.CD_4+ role in the immune induction is not only limited to antibody production also capable of generating CD8+ T cell responses [8-11]. Earlier studies on DRB1* 15, DRB1

*07, DRB1*1301, DRB1*1302, DRB1*1101-05 allele [12-14] restricted epitope mapping for GAG and ENV protein of HIV-1 "C" clade isolates if India provided the structural insights of peptide-MHC complexes So prediction of novel epitopes restricted to MHC class II allele facilitate the development of novel Gag based vaccine construct, The identification of MHC class II restricted peptide epitopes is an important goal in immunological research. Since DRB1*10 allele frequency is more among the south Indian population [15], thus our study restricted to prediction of Gag epitopes specific for DRB1*10 allele.

EXPERIMENTAL SECTION

HIV-1 "C" clade Gag sequence assessment for conservancy:
Conserved fragment of Gag protein sequences were retrieved based our earlier studies[13,14] and conservancy score for each amino acid the Gag sequence considered for further computational studies to predict promiscuous epitopes.

CD $_4$ + epitope prediction using IEDB recommended method:
Based on IC_{50} value peptide binding affinity of MHC molecules were calculated in IEDB prediction server's IEDB recommended method [16-20], a novel approach that capture distinct features of MHC class II peptide interactions. IEDB Recommended uses the Consensus approach, combining NN-align, SMM-align, and CombLib if any corresponding predictor is available for the molecule, otherwise NetMHCIIpan is used .Thus reduce the number of experiments required for identifying helper T cell epitopes and play an important role in rational vaccine design.

DRB1*10 allele sequence retrieval frequency analysis:
Protein sequence of DRB1*10 allele retrieved from the IMGT HLA sequence database [21], IMGT provides a centralized resource for clinical or scientific requirement in the HLA system [22]. Allele Frequency of DRB1*10 allele was retrieved literature survey of Balakrishnan et al and the frequency ranging from 11.53 % (sourashtrans) 25 % among the gounders [15], thus DRB1*10 allele restricted epitope prediction and their MHC interaction would be help in determining the Gag based vaccine construct.

Modeling of DRB1 *01 HLA allele and CD $_4$ + Epitopes, evaluation and binding groove analysis:
To explore MHC-peptide complex interaction pattern three dimensional structure of DRB1*10 allele structure is needed,since experimentally resolved structure for DRB1*10 is not available in public protein structure repositories, structure was modeled using I-Tasser [23]. I Tasser an online platform for protein structure and function predictions and 3D models are built based on multiple threading alignments by LOMETS and iterative templatefragemnt assembly simulations;function insights are derived by matching the 3D models with BioLip protein function database [24].Finally TM-Score and sequence identity in the structurally aligned regions were used to evaluate structural similarity between target and template models. Modeled 3D structure of DRB1*10 was evaluated by Mol probity version 4.02 [25] analysis the model's stereo chemical quality and identify the amino acid with unusual backbone confirmation's scores ,further PROSA evaluation [26] was used to compare the target and template models to ensure the modeled structure quality. DRB1*10 binding grooves were identified using Accelrys Discovery Studio 2.0 Binding site analysis module [27] and allele structure was their energies were minimized to the closest local minima using generalized CHARMM forcefield [28] as implemented in Accelrys Discovery Studio 2.0.

Modeling of Epitopes using threading method:
I Tasser [23] predicts the GAG epitopes structure specific for DRB1*10 allele based on sequence-to-structure. First generates 3D atomic models from multiple threading alignments and iterative structural assembly simulations. Accuracy of the prediction assessed based on confidence score of the modeling.

HLA-DRB1*07 allele and Gag Epitope affinity analysis:
Peptide and DRB1*10 allele complexes were generated using ClusPro server [29] that automatically computing the docking of two protein structures supplied by the user (or as PDB IDs). The result set is a ranked list of putative complexes, ordered by clustering properties. It is based on PIPER program's FFT (Fast Fourier Trans-form) [30] based rigid docking includes two stages , generation of low energy docked complexes based on pairwise interaction potential as first stage and clustering of docked complexes and low energy clusters assessment using SDU(Semi-Definite programming based Underestimation) program which predicts clusters stability using medium range optimization algorithm as second stage and finally stable clusters are further refined using Monte-Carlo simulation,top ranked models of peptide and DRB1*10 allele complex interaction visualized using DS Visulalizer 4.0 [31] .

RESULTS AND DISCUSSION

Conserved fragment of Gag protein sequences of Indian origin and their consensus scores were evaluated for further promiscuous epitope identification.

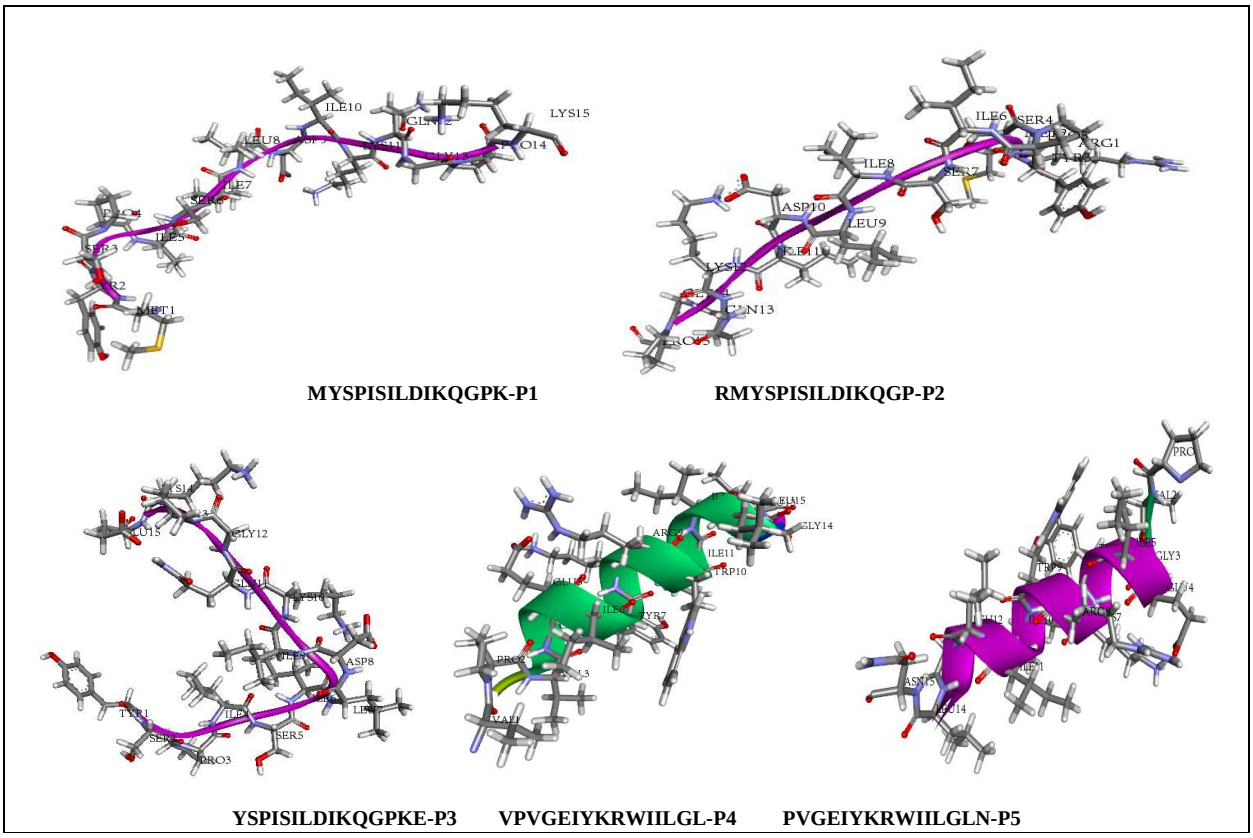
CD $_4$ + promiscuous epitope prediction and scoring evaluation:

IEDB epitope prediction restricted to DRB1*10 allele(frequency is high among south Indians) for Gag conserved fragment results were given in units of IC_{50}nM for combinatorial library and SMM_align, a lower number indicates higher affinity.Most known epitopes have high or intermediate affinity. Sturniolo is method resulted with raw score. Higher score indicates higher affinity.For each peptide and finally a percentile rank for each of the three methods (combinatorial library, SMM_align and Sturniolo) is generated,small numbered percentile rank indicates high affinity. Based on percentile rank top listed epitopes were selected and designated with identifier like P1-P6 according to their score **(Table 1)**.

Table 1: CD $_4$ + promiscuous epitope and prediction scores

Epitopes	Percentile Rank	Netmhc IIscore	smmalign score	Sturniolo score
MYSPISILDIKQGPK-P1	0.53	8.4	6	33
RMYSPISILDIKQGP-P2	0.53	7.9	6	31
YSPISILDIKQGPKE-P3	0.53	9.1	7	41
VPVGEIYKRWIILGL-P4	0.77	10.1	7	47
PVGEIYKRWIILGLN-P5	0.77	10.4	7	52

Figure 1: Three dimensional model of CD_4 + epitopes



3D modeling of DRB1 *01 HLA allele and CD_4 + Epitopes, evaluation and binding groove analysis:

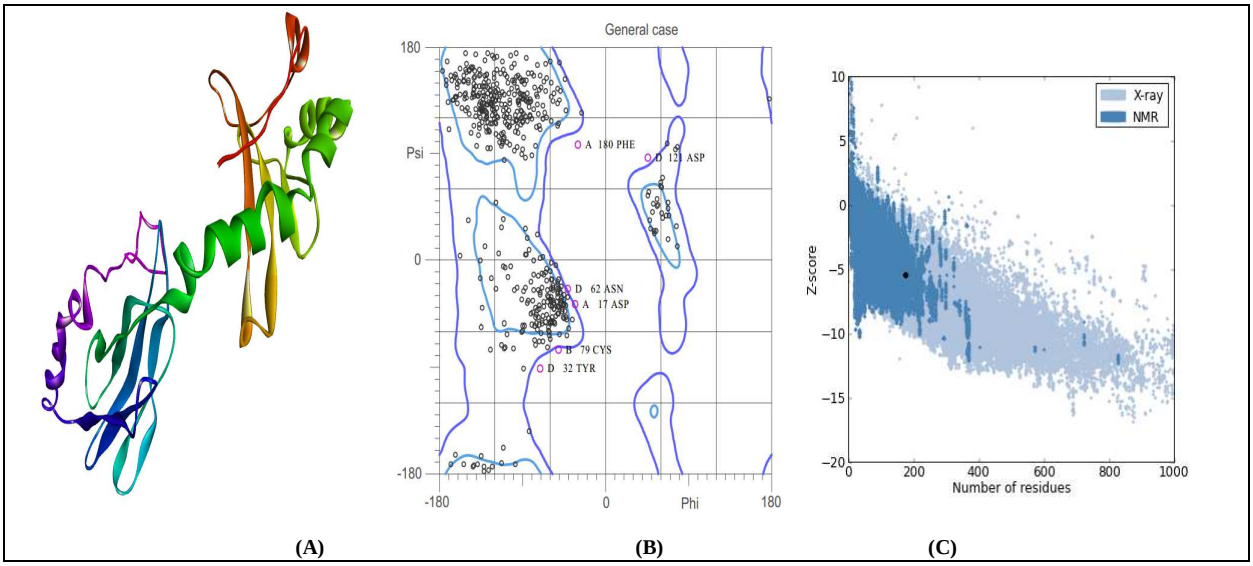
To infer the MHC-peptide affinity cannot be interpreted only on primary sequence level, to explore structural level interaction 3D structure of both peptides and DRB1*10 allele is needed and it was predicted using I-Tasser

(Figure1).CD_4 + Epitopes and DRB1*10 allele structures were modeled using I Tasser and their score were listed in the **Table-2** ,Top listed model was selected based on C score and TM score and further model was evaluated using Molprobity version 4.02, it assess the stereo chemical quality of the structure based on Ramachandran plot **(Figure-2B)** 89.7% of residues are in the favorable region. Z-score determined by NMR and X-ray generated (signify overall model quality) was -5.43 for the model DRB1*10 allele and this score was generated by PROSA web tool. **(Figure-2C)** and finally binding pocket analysis done using Accelrys DS 2.0 software and visualization of modeled DRB1*010 carried out using Accelrys DS 4.0 software (**Figure 2A**) (Accelrys Inc., San Diego, CA, USA)

Table 2: 3D structure of CD_4 + promiscuous epitopes and DRB1*10 prediction scores

EPITOPE	C SCORE	TM SCORE	RMSD
MYSPISILDIKQGPK-P1	1.31	0.52	2.1 A^0
RMYSPISILDIKQGP-P2	1.34	0.55	2.1 A^0
YSPISILDIKQGPKE-P3	1.43	0.54	2.4 A^0
VPVGEIYKRWIILGL-P4	0.28	0.75	0.5 A^0
PVGEIYKRWIILGLN-P5	0.53	0.65	1.3 A^0
DRB1*10	1.52	0.53	1.7 A^0

Figure 2: Three dimensional model of DRB1*10 allele and model evaluation

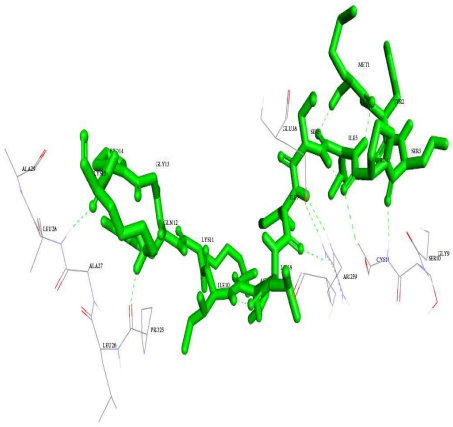


(A) **(B)** **(C)**

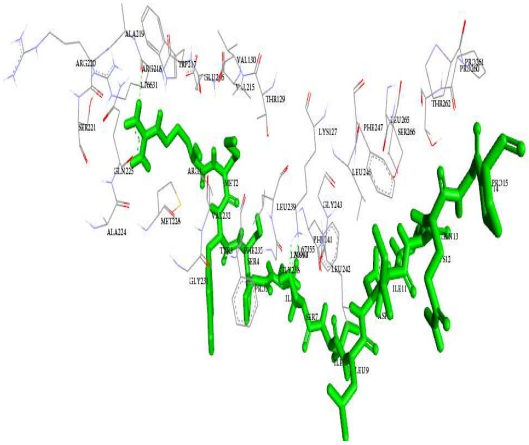
Binding affinity analysis of HLA-DRB1*10 allele and Gag Epitope:

Peptide and DRB1*10 allele complexes were generated using highly effective multistage docking approach CLUSPRO implemented with a simplified energy function and limited flexibility to discover regions of interest. The sizes and lowest energy values for largest clusters of balanced, electrostatic-favored, hydrophobic-favored and VdW+Elec were listed in the **Table 3.** The interacting residues of the representative models out of the largest clusters were analyzed and visualized by Discovery studio 4.0. Strong binding affinity between Gag epitopes and HLA alleles could implement CD_4 + cell specific immune induction. Based on binding groove occupancy of Gag epitope and then the molecular interaction type and bond distance and finally interacting residues position like anchor position especially in the c terminal region .Finally well-oriented peptide–MHC complexes from the top listed models, their binding free energy scores were compared and the peptide–MHC complex with the lowest score was selected for further molecular interaction analysis. Using Discovery studio visualizer 4.0 hydrogen bonding between the Gag epitopes and DRB1*10 allele were analyzed and bond distance upto 3.5 A^0 were considered for promiscuous epitope selection. Current study combines of epitope prediction processes as well a knowledge-based molecular docking technique thus provides structural insights of MHC class I CTL epitopes compared with the basic automated epitope-predicted servers.

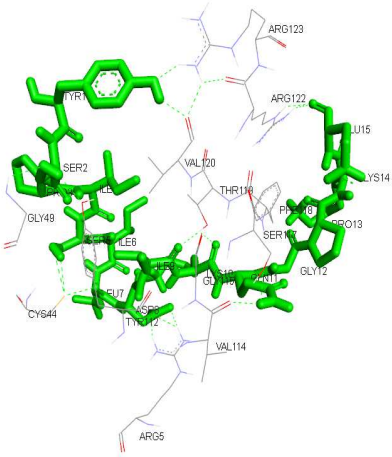
Figure 3:CD_4+ Gag epitopes binding affinity for DRB1*10 allele



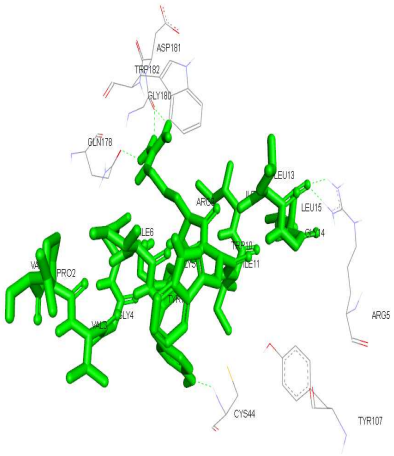
MYSPISILDIKQGPK-P1 & DRB1*10



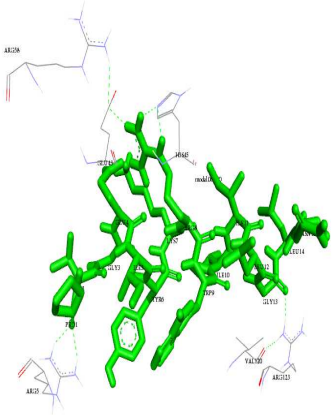
RMYSPISILDIKQGP-P2 & DRB1*10



YSPISILDIKQGPKE-P3 & DRB1*10



VPVGEIYKRWIILGL-P4 & DRB1*10



PVGEIYKRWIILGLN-P5 & DRB1*10

Table 3:Binding affinity of Gag epitopes for DRB1*10 allele

GAG Epitopes	Binding Energy (kcal/mol)	Interacting Atoms	Bond Distance ($A^{0)}$
RMYSPISILDIKQGP-P2	-818	DRB1*10:CYS11: H – Gag Epitope: SER3: O	1.90851
		DRB1*10:CYS11: SG – Gag Epitope: PRO4: O	2.86589
		DRB1*10:LEU28: H - Gag Epitope: PRO14: O	1.89994
		DRB1*10:ARG59:HH12 - Gag Epitope: ILE7: O	2.46928
		Gag Epitope:ILE7:H - DRB1*10:GLU38:OE2	2.36372
		Gag Epitope: LYS11:HZ3 - DRB1*10:THR18: O	2.07765
		Gag Epitope: LYS15:HZ3 - DRB1*10:PRO25: O	1.72076
MYSPISILDIKQGPK-P3	-910.2	DRB1*10:CYS11: H - Gag Epitope: SER3: O	1.90851
		DRB1*10:CYS11: SG - Gag Epitope: PRO4: O	2.86589
		DRB1*10:LEU28: H - Gag Epitope: PRO14: O	1.89994
		DRB1*10:ARG59:HH12 - Gag Epitope: ILE7: O	2.46928
		Gag Epitope: LYS11:HZ3 - DRB1*10:THR18: O	2.07765
		Gag Epitope: LYS15:HZ3 - DRB1*10:PRO25: O	1.72076
PVGEIYKRWIILGLN-P6	-966.2	DRB1*10:ARG5:HH12 - Gag Epitope: PRO1: O	2.04454
		DRB1*10:ARG5:HH22 - Gag Epitope: PRO1: O	1.89448
		DRB1*10:ARG123:HH12-Gag Epitope: GLY13: O	1.76769
		Gag Epitope:ARG8:HE - DRB1*10:HIS45:NE2	2.08473
		Gag Epitope:ARG8:HH11 - DRB1*10:HIS45:NE2	2.06537
		Gag Epitope:ARG8:HH12 - DRB1*10:GLU43:OE2	1.95325
VPVGEIYKRWIILGL-P5	-839.5	DRB1*10:CYS44:H - Gag Epitope:TYR7:OH	2.00485
		Gag Epitope:ARG9:HH11 - DRB1*10:GLN178:OE1	1.86235
		Gag Epitope: ARG9:HH12 - DRB1*10:GLY180: O	1.8262
		Gag Epitope: ARG9:HH22 - DRB1*10:GLY180: O	2.34405
YSPISILDIKQGPKE-P4	-757.4	DRB1*10:ARG5:HH12 - Gag Epitope:ASP8:OD1	2.29533
		DRB1*10:ARG5:HH12 - Gag Epitope:ASP8:OD	1.90733
		DRB1*10:ARG5:HH22 - Gag Epitope:ASP8:OD1	1.78382
		DRB1*10:CYS44:SG - Gag Epitope:SER5:OG	3.15502
		DRB1*10:CYS44: SG - Gag Epitope: LEU7: O	3.01267
		DRB1*10:GLY49:H - Gag Epitope:SER2:OG	2.48931
		DRB1*10:THR119:HG1 - Gag Epitope: ILE9: O	1.97642
		DRB1*10:THR119:HG1 - Gag Epitope: LYS10: O	2.48506
		DRB1*10:ARG122: HE - A: GLU15: O	2.06194
		DRB1*10:ARG122:HH11 - Gag Epitope: GLU15: O	1.81744
		DRB1*10:ARG123:HH12 - Gag Epitope:TYR1:OH	1.69644
		Gag Epitope: TYR1: HH - DRB1*10:VAL120: O	1.93038
		Gag Epitope:SER2:HG - DRB1*10:TYR112:OH	1.97987
		Gag Epitope:SER5:HG - DRB1*10:CYS44:SG	2.47207
		Gag Epitope: GLN11:HE21 - DRB1*10:VAL114: O	2.02438
		Gag Epitope:GLN11:HE21 - DRB1*10:SER117:OG	2.45162

CONCLUSION

Computational study on Insights of HLA-DRB1*10 Allele-Gag peptides would be helpful indentifying the Gag based HIV vaccine construct .Epitope based vaccine construction using conventional methods are time consuming and expensive, thus computation method for epitope based vaccine construct screening based on IEDB method identified the promiscuous epitopes namely P1-P5 and were specific for DRB1*10 allele since the stable hydrogen bonding and epitope's occupancy on binding groove of the DRB1*10 alleles and lower binding energy score epitopes were ranked .Binding of antigenic peptides on HLA allele signifies their role in cellular immune response .Complete systematic study suggests the promiscuous epitope candidates need to be further tested experimentally for conventional vaccine design.

REFERENCES

[1] M. D Moore; W. S Hu. *AIDS Reviews*. **2009,** 11(2), 91–102.
[2] Factsheet HIV and AIDS in Asia and the Pacific http://www.aids2014.org/ webcontent/ file/ AIDS2014_Fact_sheet_ Asia_Pacific.pdf
[3] United Nations Programme on HIV/AIDS. "Global Report fact sheet: the global AIDS epidemic, **2013**". http://www.unaids.org/en/resources/campaigns/globalreport2013/factsheet/

[4] NACO annual report 2013-14.http://naco.gov.in/NACO/Quick_ Links/Publication/ Annual_ Report/NACO_Annual_Report/Annual_Report_2013-14/,2014.
[5] Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents http://aidsinfo.nih.gov/contentfiles/lvguidelines/adultandadolescentgl.pdf,**2014**.
[6] Paolo Palma, Lindvi Gudmundsdotter , Andrea Finocchi et al.,*Vaccines* **2014**, 2, 563-580.
[7]HIV sequence database (http://www.hiv.lanl.gov)
[8] DJ Shedlock, Shen *H. Science*. **2003**; 300:337–339.
[9] LH Carvalho, G Sano, Hafalla JC, A Morrot, MA Curotto de Lafaille, *F* Zavala. *Nat Med.* **2002**; 8:166–170.
[10]CC Kemball, CD Pack, HM Guay, ZN Li, DA Steinhauer, E Szomolanyi-Tsuda et al. *J Immunol*. **2007**; 179:1113–1121.
[11] AL Marzo, V Vezys, KD Klonowski, SJ Lee, G Muralimohan, M Moore et al. *J Immunol.* **2004**; 173:969–975.
[12] Christy and Anand.*BMC Infectious Diseases* **2014**, 14(Suppl3):O4.
[13] H Jemmy Christy, D Alex Anand. *Int J Pharm Bio Sci* **2014** Oct; 5(4): (B) 1212 – 1222.
[14]H Jemmy Christy, D Alex Anand . *J. Chem. Pharm. Res.*, **2014**, 6(10):762-770.
[15] K Balakrishnan; C Rathika; R Kamaraj. *Int J Hum Genet*. **2012**, 12(1): 45-55.
[16] P Wang, J Sidney, Y Kim, A Sette, O Lund O, M Nielsen ,B Peters **2**. *BMC Bioinformatics.* 2**010, 11:568**.
[17] M Nielsen ,O Lund .*BMC Bioinformatics.* **2009 ,**10:296.
[18] M Nielsen, C Lundegaard ,O Lund . *BMC Bioinformatics.* **2007,8**:238.
[19] T Sturniolo , E Bono , J Ding , L Raddrizzani ,O Tuereci , U Sahin ,M Braxenthaler , F Gallazzi , MP Protti , F Sinigaglia , J Hammer . *Nat Biotechnol*. **1999,**17(6):555-561.
[20] E1 Karosiene,M Rasmussen,T Blicher,O Lund,S Buus,M Nielsen,*Immunogenetics* , **2013,** 65(10)711.
[21] J Robinson , JA Halliwell , H McWilliam , R Lopez , P Parham , SG Marsh , *Nucleic Acids Research* ., **2013**, 41,D1222-7.
[22] J Robinson , JA Halliwell , SG Marsh , *Mol Biol*., **2014**, 1184,109-21.
[23] Yang Zhang. *BMC Bioinformatics* **2008**, 9,40.
[24] A Roy, A Kucukural, Y Zhang. *Nature Protocols*, 5: 725-738 (**2010**)
[25] Chen et al. (**2010**) , *Acta Crystallographica* D66:12-21.
[26] Wiederstein & Sippl, *Nucleic Acids Research*, **2007**, 35, W407-W410
[27]Accelrys Software Inc., Discovery Studio Modeling Environment, Release 4.0, San Diego: Accelrys Software Inc., **2007**.
[28] BR Brooks , RE Bruccoleri , BD Olafson ,DJ States ,S Swaminathan ,M Karplus , J. Comp.Chem. **1983**, 4,187–217
[29] Kozakov D, Beglov D, Bohnuud T, Mottarella S, Xia B, Hall DR, Vajda, S. *Proteins: Structure, Function, and Bioinformatics*, **2013** Aug
[30] Kozakov D, Brenke R, Comeau SR, Vajda S. *Proteins*. **2006** Aug 24
[31] Accelrys Software Inc., Discovery Studio Modeling Environment, Release 4.0, San Diego: Accelrys Software Inc., **2013**