Innovare Academic Sciences

International Journal of Pharmacy and Pharmaceutical Sciences
ISSN- 0975-1491 Vol 6, Issue 9, 2014

Original Article

EPITOPE PREDICTION AND STRUCTURAL ANALYSIS OF PRA ANTIGEN OF *COCCIDIOIDES*

PRASHANT SAXENA, JEMMY CHRISTY. H*

Department of Bioinformatics, Sathyabama University,Chennai 119, India.
Email: jemmyjacob@gmail.com

Received: 12 Jul 2014 Revised and Accepted: 18 Aug 2014

ABSTRACT

Objective: Coccidioidomycosis is a life threatening human respiratory disease caused by *Coccidioides immitis* and *C. posadasii*. As the incidence of symptomatic Coccidioidomycosis increases worldwide, vaccine development strategy against *Coccidioides* is needed for both immunocompetent and immunocompromized individuals. An insight about PRA antigens in the pathogenesis of Coccidioidomycosis has the potential to develop a therapeutic intervention.

Methods: In this study of ANN, SMM and QM based servers were used to predict promiscuous epitopes. Epitope structures were modeled using De novo based PEPFOLD server and minimized using Swiss model server. Further the modeled epitopes were tested for their binding affinity towards HLA alleles by means of peptide protein docking studies involving autodock.

Results: A total of six antigenic epitopes ARISVSNIV, GRPTASTPA, ALNLFVEAL, LVAAGLASA, FSHALJALV, AEPTHEPTE of PRA showed hydrogen bonding with HLA alleles HLA-A*02:03, HLA B*27:05, HLA-A*02:01 with considerable bond distance. Thus, this systematic study on epitopes of *Coccidioides* study would be helpful in designing and developing novel PRA antigen based vaccine and inhibitor.

Conclusion: These predicted sequences are potential vaccine candidates but functional/biological assays should be performed to verify whether they are indeed appropriate to be included in a vaccine formulation.

Keywords: HLA-A*02:03, HLA-B*27:05, HLA-A*02:01, PRA antigen, Epitope, Coccidioidomycosis,Vaccine.

INTRODUCTION

Coccidioides immitis and *C. posadasii* are the two nearly identical species of pathogenic fungal species are the prominent cause for Coccidioidomycosis (also known as Valley Fever), they grow in the top 2-8 inches of soil in semi-arid regions and produces arthroconidia, highly infectious propagules that can be disrupted from the ground and inhaled, where they can produce acute pulmonary infection in humans and many mammals [1-3]. Coccidioidomycosis infection will remain asymptomatic or present with mild respiratory symptoms among sixty percentages of infected people[4]. Approximately 40% of infections result in symptomatic disease, typically arising one to four weeks after exposure, which can resemble ordinary influenza, with fever, cough, fatigue, dyspnea, headache, myalgia and arthralgia.

The disease is considered non-transmissible, although rare instances of transmission have occurred, including via organ transplantation [5] and inhalation of spores growing in visecro-cutaneous fistulae. A low percentage of symptomatic infections progress, resulting in extrathoracic dissemination,with hematogenous spread to the central nervous system, skin, joints, major organs or bones and result in chronic (months to years) and even fatal disease [6,7]. Persons at increased risk of disseminated coccidioidomycosis include immune compromised persons, e. g., HIV/AIDS, diabetics, pregnant women [8,9] and persons of certain race/ethnicities, particularly Blacks Filipinos and Asians [10].

The Ag2/PRA- proline-rich coccidioidal antigen is a protein composed of 194-amino-acid residues which is a component of a glycopeptide. Several studies reveal, either protein vaccines using recombinant Ag2/PRA (rAg2/PRA) or DNA vaccines based on the sequence encoding Ag2/PRA would aid protection from lethal coccidioidal infection in mice [11,12, 13,14].

Despite numerous investigations, there is no real drug available for the treatment of the Coccidioidomycosis. In the absence of effective antiviral treatment, prevention through vaccination would greatly reduce morbidity and mortality associated with coccidiomycosis infections. Epitope mapping is the process of identifying the binding sites, or epitopes, of antibodies on their target antigens. Epitope mapping has become one of the key elements of both vaccine and drug development. Predicting immune epitope for T cell epitopes using computational methods assisting in assessing the infectious organism for imuunogenic antigens. An immune response elucidated by peptide based vaccines which in turn present immunogenic peptides to major histo compatability complex [15]. It is well established that T cells play a critical role in inducing immune response against foreign antigen. But for their activation, the antigenic fragment must bind to MHC molecules. There by predicting which antigenic fragment can bind to the MHC molecule is the first bottlenecks in vaccine design. As there is no experimentally resolved structure of PRA protein, it is challenging to select proper target that would lead to predict epitope and ligand binding sites in protein. Hence, this study aims to investigate the PRA gene with special focus on the structural and functional aspects through bioinformatics approach. Studies on HLA allelic diversity of among Asians population revealed some common alleles such as HLA-A*02:01. HLA-A*02:03[16],HLA-B*27:05[17]. We aimed to find out the epitope interaction with HLA-A*02:01, HLA-A*02:03, HLA-B*27:05 using the insilico tools and the protein-protein docking studies, a critical step in the development of vaccines candidates.

MATERIALS AND METHODS

Protein Sequence retrieval

The sequence of the complete CDS protein of *Coccidioides immitis* retrieved from Enterz protein database from NCBI (http://www. ncbi. nlm. nih. gov) has 194 amino acid residues. The sequence was saved in FASTA format and used for further analysis.

Potential epitope prediction of antigenic protein

To predict the T cell epitopes in PRA antigenic region of *Coccidioides* a systematic approach was adapted based on different algorithms. PRA antigen for the alleles HLA-A*02:03, HLA B*27:05, HLA-A*02:01 were predicted by using different Insilco tools like HLAPred(http://www. imtech. res. in/raghava/hlapred/), MHCPred[18], MAPPP[19], IEDB(www. iedb. org) [20] which are based on ANN (Artificial Neural Network), SMM (Stabilized Matrix Alignment Method) algorithm and “Quantitative Matrices” and “Additive Method” method.

De novo modeling of the epitopes

All the predicted epitopes were obtained in the form of small peptide sequences (nanomers). The three dimensional structures of small peptides (epitopes) were required to find out their prominent interaction with HLA alleles, so we constructed the three dimensional model for all predicted promiscuous epitopes using de novo based prediction server PEP-FOLD[21] that modeled linear peptides structure up to 36 amino acids with an averaged RMSD of 2.1 A from NMR structure, with user specified constraints such as disulfide bonds and inter-residue proximities. Further energy minimization was done for these nanomeric peptides using Swiss PDB viewer[22] (http://swissmodel. expasy. org).

HLA allele sequence retrieval and Homology modeling and structure refinement of HLA allele

The HLA-A*02:03, HLA-B*27:05. HLA-A*02:01 alleles sequences were retrieved from IMGT/HLA (http://www. ebi. ac. uk) database, and three dimensional structure of 1HHG (HLA-B*27:05) was retrieved from PDB database (http://www. rcsb. org),and there is no experimentally resolved three dimensional structure for HLA-A*02:03and HLA-A*02:01 in the protein data bank (PDB),structure was modeled by homology based modeling software MODELLER 9.13 [23],and we used 3OX8 which is an X-ray diffraction structure of the HLA-A2 allele with resolution of 2.16 and 1HHG an X-ray diffraction structure of HLA-A2 allele with resolution of 2.60 as templates for HLA-A*02:03,HLA-A*02:01 alleles respectively.

Energy minimization test was used to find out the energy parameters in comparison with the potential of mean force derived from a large set of known protein structures using Swiss PDB viewer. Then to evaluate backbone conformation Psi/Phi angles of modeled HLA allele structures Ramachandran plot obtained from PROCHECK [24] and checked for the presence of non-GLY residues in the disallowed regions.

The modeled quality was further determined based on Z-score,an model quality score which used to check whether the input structure is within the score range that is found in native protein of similar size using web based tools PROVE and ERRAT which are embedded in the SAVES server(http://nihserver. mbi. ucla. edu/SAVES/).

PRA antigenic peptide and HLA allele's interaction analysis using docking

The PRA antigen's promiscuous epitopes were further tested by *in silico* docking simulation to find out whether or not this peptide will bind to the HLA molecules when it will be applied *in vivo.* To accomplish peptide HLA allele interaction simulation, experimentally resolved and modeled HLA three dimensional structures of the HLA alleles HLA-A*02:03, HLA-B*27:05, HLA-A*02:01 were docked using Autodock 4.2, a suite of automated docking tools designed to predict how small molecules, such as nanomeric peptides bind to a receptor of known 3D structure [25] by employing an implementation of the Lamarckian genetic algorithm (GA),to model the peptide binding to HLA alleles, and in this The input ligands and the output data were mined by python scripts using the MGL Tools 1.5.4 package [26]. Here all retained poses considered in the study had an RMSD below 2.0Å . Further computational burden was minimized for calculating peptide–MHC interactions at positions not involved in the static docking. All coordinates were kept fixed apart from the peptide residues of interest. These were left flexible. Thus, GA settings were kept to their default values, apart from the number of energy evaluations and the number of generations which were set to 2 500 000 and 27 000, respectively. The docking grid was defined as a cuboid with respective dimensions of: $68\,Å \times 80\,Å \times 80\,Å$ for DP1, $72\,Å \times 80\,Å \times 82\,Å$ for DP41, $72\,Å \times 80\,Å \times 82\,Å$ for DP42 and $72\,Å \times 80\,Å \times 82\,Å$ for DP5 which encompassed the entire peptide binding site on DP. Output from ten independent GA runs for each ligand was processed and the pose (binding conformation) with the lowest Free Energy of Binding (FEB) was considered. FEB values represent the direct output from the AutoDock 4.2 scoring function which takes into consideration weighted terms for van der Waals dispersion/repulsion, hydrogen bonding, electrostatics, desolvation interactions, and the change in torsional free energy when the ligand goes from an unbound to a bound state.[25]

RESULTS AND DISCUSSION

The relative ability of a peptide epitope-MHC complex to elicit an immune response was predicted using "T cell -MHC immunogenicity predictors" that counts amino acid properties as well as their position within the peptide to predict the immunogenicity of a PRA antigenic epitope- HLA complex.

The predictions of viable epitopes were done using epitope prediction tools. The criterion of predicting epitopes was different for each tool. Prediction by IEDB tool is based on IC_{50} value. MHCPred is also based on IC_{50} value. If value is less than 50 (>50) consider as binder. MAPPP is based on score value, and minimum score is 0.5, and for HLAPred is based on threshold value which was 5 for the above prediction.

Epitopes model generations

Computationally predicted nanomeric epitopes from the pool of 182 epitopes, only six promiscuous epitopes were modeled by PEP-FOLD and their energy levels were minimized using swiss pdb viewer then labeled using PYMOL. Then the systematic analysis proceeded to assess their binding efficacy for HLA alleles which are common among ethnic group.

Table 1: Predicted promiscuous epitopes and their binding affinity calculations

HLA Allele	Epitope	IEDB IC_{50}	MAPPP	MHCPred	HLAPred
HLA-B*27:05	ARISVSNIV	27.9	1.7	37.84	10.460
HLA-B*27:05	GRPTASTPA	16.2	2.0	24.49	12.010
HLA-A*02:01	ALNLFVEAL	48.9	0.66	106.91	5.600
HLA-A*02:03	LVAAGLASA	31.3	1.49	45.29	9.090
HLA-A*02:03	FSHALIALV	37.1	1.33	85.31	7.770
HLA-A*02:03	AEPTHEPTE	44.7	1.0	92.04	6.460

Retrieval and Modelling of HLA alleles

Probably three-dimensional (3-D) structure of HLA alleles was generated by protein structure homology modeling server Modeller9.13. Further these modeled structures were minimized using Swiss PDB viewer. Then the quality of the predicted 3D structure of HLA-A*02:03 were confirmed by SAVE server. The perceived Ramchandran plot (Psi-Phi) pairs had 93.8% of residues in most favored regions,5.8 % core residues in additional allowed regions, 0.0 % residues in generously allowed regions and 0.4% residues in disallowed regions, finally overall quality factor of the model was 93.8%,where as another allele HLA-A*02:01 model quality was 86.54% which includes 96.2% of residues in the most favored region and rest of residues lies in additionally allowed and generously allowed region, none of the residues in the disallowed region. Considering these parameter the quality of the model was assessed as good.

HLA and Epitopes interaction analysis using Docking studies

It's worth noting all the parameters after docking, because it computes 11 types of energy values like Estimated free energy of

binding ($E_{FreeBind}$), Final Intermolecular Energy ($E_{InterMol}$), which is the sum of 4 energies such as vdW + Hbond + desolv Energy (E_{VHD}), Electrostatic Energy (E_{Elec}), Moving Ligand-Fixed Receptor (E_{MLFR}), and Moving Ligand-Moving Receptor (E_{MLMR}) Final Total Internal Energy (E_{FTot}), again the sum of 2 energy values such as Internal Energy Ligand (E_{IntL}), and Internal Energy Receptor (E_{IntR}) Torsional Free Energy (E_{Tors}) and unbound System's Energy (E_{Unb}). Minimum energy has greater stability and vice versa. Thus Docking studies would reveal the insight about the favorable interactions between epitope and HLA alleles and based on binding energy score and hydrogen bonds of all docked complexes epitopes efficacy as a vaccine candidate was assessed. Six antigenic epitopes and their binding energies,total internal energy as well torsional free energy were listed in the Table.1.

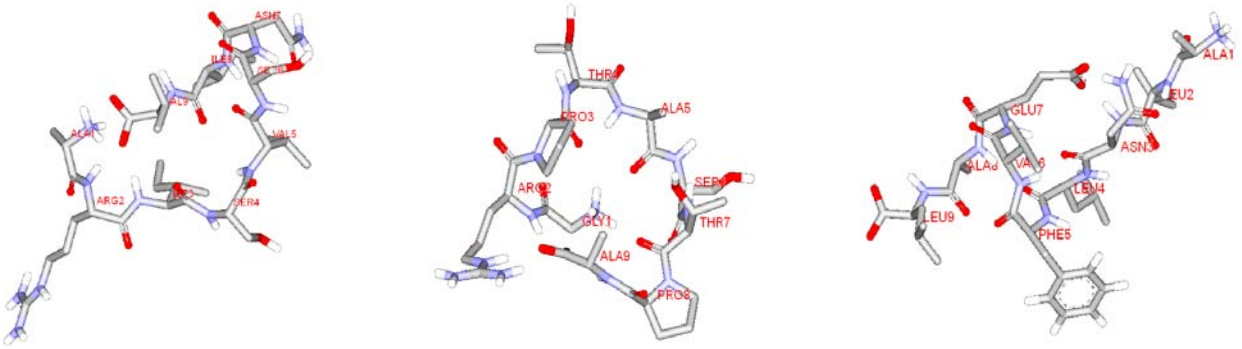

Fig. 1: Modeled nanomeric epitope-ARISVSNIV, GRPTASTPA, ALNLFVEA.

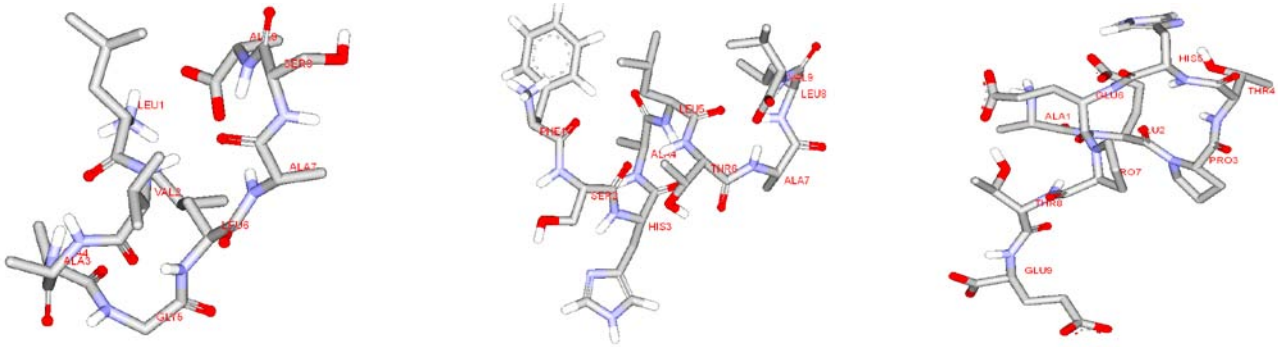

Fig. 2: Modeled nanomeric epitope- LVAAGLASA, FSHALIALV, AEPTHEPTE

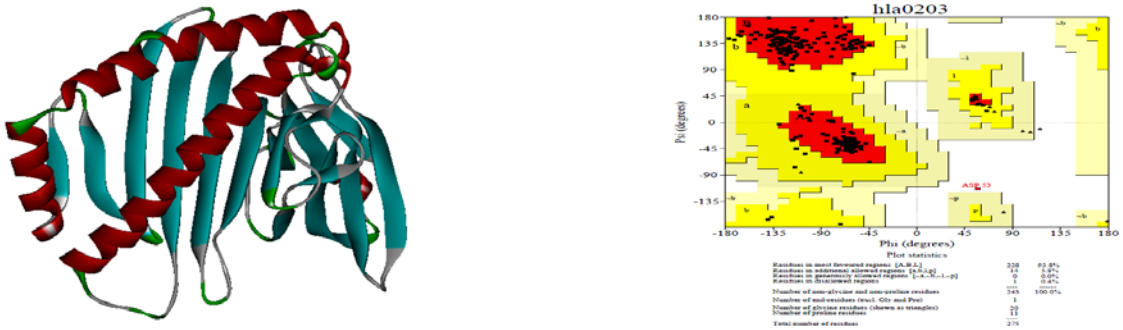

Fig. 3: Modeled HLA-A*02:01 three dimensional structure and Ramachandran plot evaluation

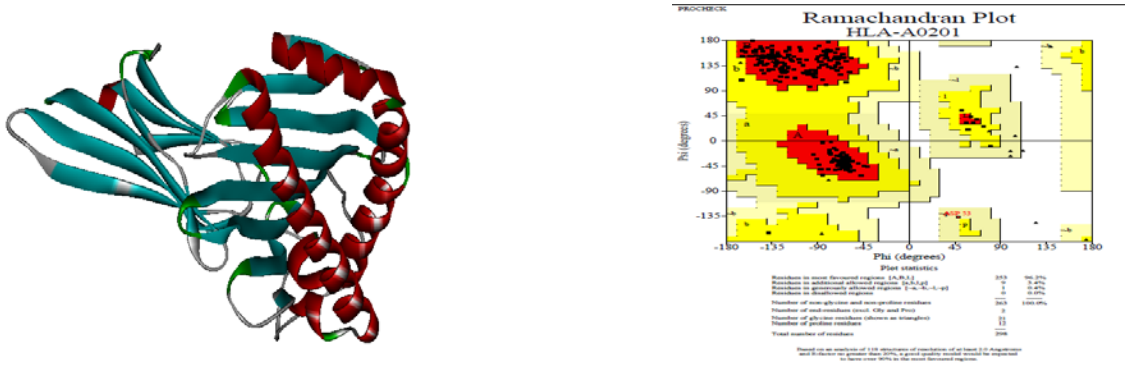

Fig. 4: Modeled HLA-A*02:01three dimensional structure and Ramachandran plot evaluation

Table 2: HLA alleles and nanomeric epitopes interaction based energy calculations

HLA Allele	Epitope	$E_{FreeBind}$	E_{FTot}	E_{Tors}	K^i
HLA-A*02:01	ALNLFVEAL	-2.44	-5.88	+8.78	16.17
HLA-A*02:01	AEPTHEPTE	-2.80	-4.98	+8.51	8.85
HLA-B*27:05	GRPTASTPA	-7.69	-2.47	+7.13	20.39
HLA-B*27:05	ARISVSNIV	-5.27	-1.18	+9.33	28.31
HLA-A*02:03	LVAAGLASA	-4.72	-3.89	+6.86	345.79
HLA-A*02:03	FSHALIALV	-3.22	-6.45	+8.51	4.40

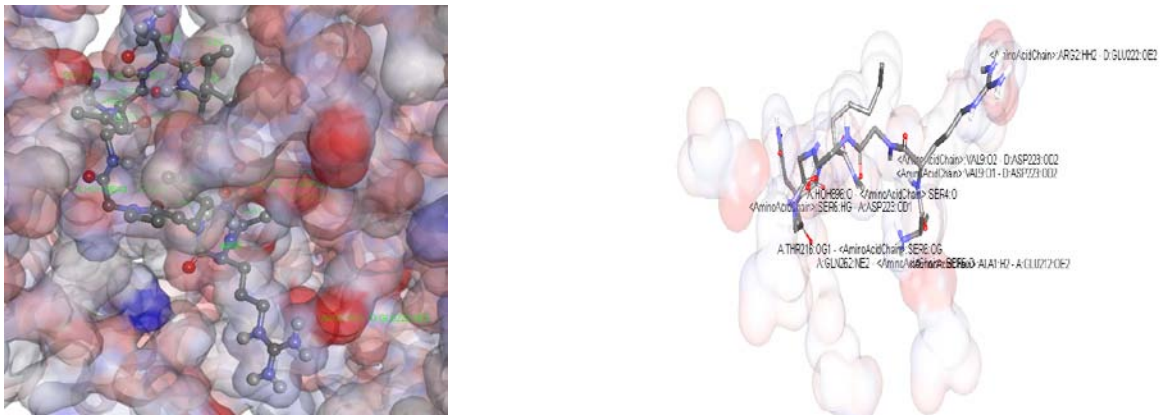

Fig. 5: Binding pose of ARISVSNIV nanomeric peptide on HLA-B*27:05

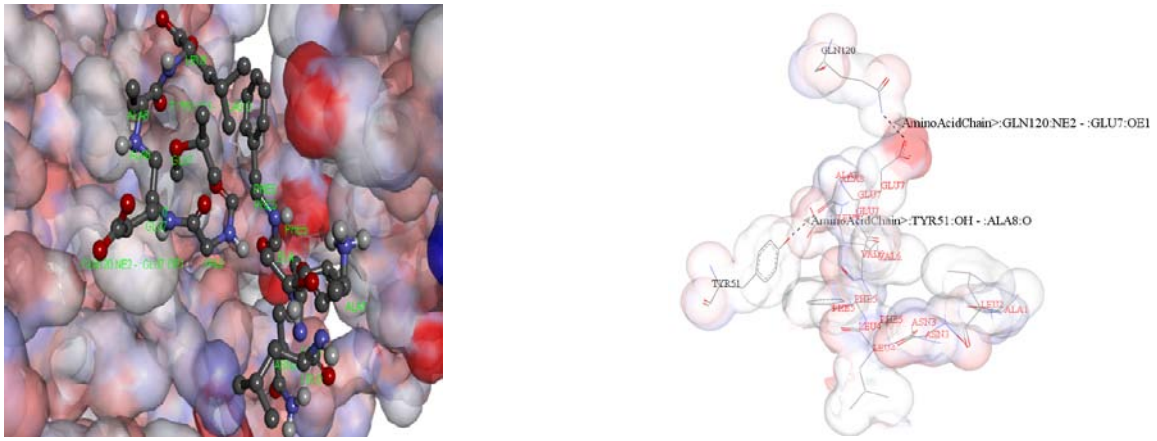

Fig. 6: Binding pose of ALNLFVEAL nanomeric peptide on HLA-A*02:01

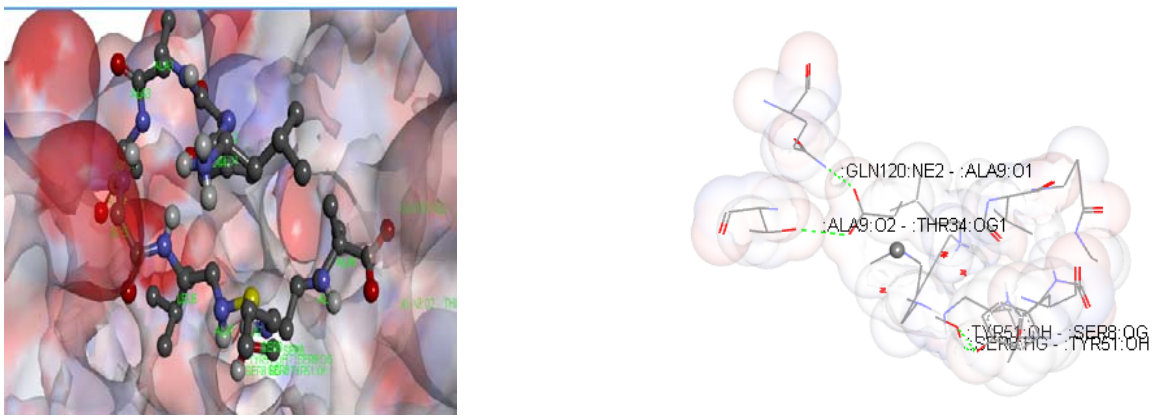

Fig. 7: Binding pose of LVAAGLASA nanomeric peptide on HLA-A*02:03

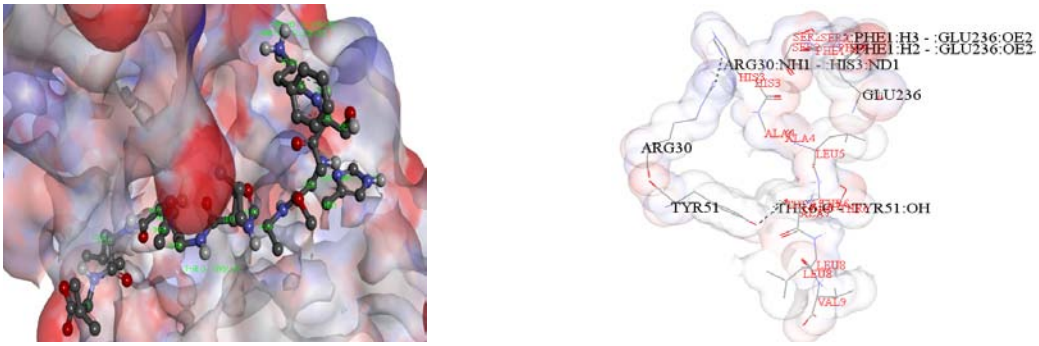

Fig. 8: Binding pose of FSHALIALV nanomeric peptide on HLA-A*02:03

Fig. 9: Binding pose of AEPTHEPTE nanomeric peptide on HLA-A*02:01

Fig. 10: Binding pose GRPTASTPA of nanomeric peptide on HLA-B*27:05

Table 3: Interacting residues of HLA alleles with nanomeric peptides

Allele	Epitope	H-Bonds
HLA-B*27:05	ARISVSNIV	ALA1- GLU212 ARG2- GLU222 GLN262- SER6 THR216-SER6 SE6-ASP223 VAL9-ASP223
HLA-B*27:05	GRPTASTPA	THR225- SER6 GLN262- ARG2 THR214- ARG2 GLY1- ASP223
HLA-A*02:01	ALNLFVEAL	ARG30- LEU4 ARG30- ALA8 THR257- ALA1
HLA-A*02:03	LVAAGLASA	THR51- SER8 GLN120- ALA9 ALA9-THR34 SER8-TYR51
HLA-A*02:03	FSHALIALV	PHE1- ASP54 PHE1- ASP53 TYR51- SER2
HLA-A*02:01	AEPTHEPTE	TYR51- GLU2 ILE237- GLU9

Here we attempt to explore the use of intermolecular bonding patterns of HLA alleles and it is clear that peptide/HLA structural interaction patterns vary among different alleles and may be based on family dependent manner. The results obtained here would provide an insight of peptide binding motifs or receptor information. From the molecular docking energy table, we observed that epitope GRPTASTPA have the minimum binding energy i.e. -7.69 Kcal/Mol among all other epitopes and AEPTHEPTE epitope has the maximum energy i. e. -2.80 Kcal/Mol. From the docking interaction table, it is clear that we have 6 epitopes, out of which five nanomeric epitopes have more hydrogen bonding with considerable bond distance (ARISVSNIV, ALNLFVEAL, LVAAGLASA, FSHALJALV, AEPTHEPTE,) with the HLA alleles.

CONCLUSION

Vaccine development for coccidiomycosis is based on screening of multiple antigenic epitopes that direct the immune system to protect the host from fungal infection. The purpose of the present study was screening the new immunogenic HLA class I restricted cytotoxic T cell (CTL) epitopes of PRA antigen and analyzing their interaction with HLA alleles. Promiscuous epitopes like ARISVSNIV,ALNLFVEAL, LVAAGLASA, FSHALJALV, AEPTHEPTE interacting with HLA class I alleles like HLA-B*27:05, HLA-A*02:03, HLA-A*02:01 and exhibiting minimal binding energy for MHC-I receptor. So these epitopes are the good binders for MHC-I which can be the key factor for developing the diagnostic tool as well as a vaccine candidate for coccidiomycosis. Epitopes binding sites analysis would be beneficial for protein function annotation and designing the structure based drugs.

CONFLICT OF INTERESTS

Declared None

REFERENCES

1. Fisher F, Bultman MW, Pappagianis D, Washington DC. Operational Guidelines for Geological Fieldwork in Areas Endemic for Coccidioidomycosis (Valley Fever);US Geological Survey Open-File Report; US Department of the Interior: USA 2000;1-16.
2. Stevens DA. Coccidioidomycosis. The New England J of Medicine 1995;332(16):1077-82.
3. Chiller TM, Galgiani JN, Stevens DA. Coccidioidomycosis. J Infect Dis Clin North Am 2003;17(1):41-57.
4. Saubolle MA, McKellar PP, Sussland D. Epidemiologic, clinical, and diagnostic aspects of coccidioidomycosis. J Clin Microbiol 2007;45(1):26-30.
5. Blair JE. Coccidioidomycosis in patients who have undergone transplantation. J Ann NY Acad Sci 2007;1111:365-76.
6. Kirkland TN, Fierer J. Coccidioidomycosis: a reemerging infectious disease. J Emerg Infect Dis 1996;2(3):192-9.
7. Ampel NM. The complex immunology of human coccidioidomycosis. Ann N Y Acad Sci 2007;1111:245-58.
8. Masannat FY, Ampel NM. Coccidioidomycosis in patients with HIV-1 infection in the era of potent antiretroviral therapy. Clin Infect Dis 2010, *50*, 1-7.
9. Santelli AC, Blair JE, Roust LR. Coccidioidomycosis in patients with diabetes mellitus. The American J of Medicine 2006;119(11):964-9.
10. Pappagianis D, Lindsay S, Beall S, Williams P. Ethnic background and the clinical course of coccidioidomycosis. Am Rev Resp Dis1979;120:959-61.
11. Abuodeh RO, LF Shubitz, E Siegel, S Snyder, T Peng, KI Orsborn, E Brummer, DA Stevens, JN Galgiani. Resistance to *Coccidioides immitis* in mice after immunization with recombinant protein or DNA vaccine of a proline-rich antigen. Infect Immun 1999;67:2935-40.
12. Jiang C, Magee DM, Quitugua TN, Cox RA. Genetic vaccination against Coccidioides immitis: comparison of vaccine efficacy of recombinant antigen 2 and antigen 2 cDNA. Infect Immun 1999;67(2):630-5.
13. Kirkland TN, Finley F, Orsborn KI, Galgiani JN. Evaluation of the proline-rich antigen of Coccidioides immitis as a vaccine candidate in mice. Infect Immun 1998;66(8):3519-22.
14. Schneider E, Hajjeh RA, Spiegel RA, Jibson RW, Harp EL, Marshall GA, *et al*. A coccidioidomycosis outbreak following the Northridge, Calif, earthquake. J of the American Medical Association 1997;277(11):904-8.
15. GCBM Arun, Susumu Goto. Prediction of CTL epitope in silico modeling and functional analysis of cytolethal distending toxin CDT protein of Campylobacter jejuni Notes 92 2014;7.
16. Menchaca L, Rood H, Komerofsky R D. New allele frequency database: http: //www. allelefrequencies. net. Middleton Tissue Antigens 2003:403-7.
17. Taurog JD. The mystery of HLA-B if it isn't one thing, it's another. Arthritis Rheum 2007;27:2478-81.
18. Guan P, Hattotuwagama CK, Doytchinova IA, Flower DR. MHCPred 2.0:an updated quantitative T-cell epitope prediction server. Applied Bioinformatics 2006;5(1):55-61.
19. Alexander H, Hans-Georg S, Hermann-Georg K, Stefan HE, I Jörg Hansjörg Christina tter, Peter-M, *et al.* MAPPP-MHC-Peptide Processing Prediction. Applied Bioinformatics 2003;2(3):155-8.
20. Vita R, Zarebski L, Greenbaum JA, Emami H, Hoof I, Salimi N, *et al*. The immune epitope database 2. 0 Nucleic Acids Res 38Database issue; 2010:854-62.
21. Thévenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tufféry P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res 2012;40:288-93.
22. TJN. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res Oxford Univ Press 2003.
23. Webb B, Sali A. Protein structure modeling with MODELLER. Methods in molecular biology (Clifton, NJ) 2014;1137:1-15.
24. Laskowski RA, McArthur MW, Moss DS, Thornton JM J. PROCHECK: a program to check the stereochemical quality of protein structures. Cryst 1993;26:283-91.
25. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, *et al*. AutoDock4 and AutoDockTools4:Automated docking with selective receptor flexibility. J Comput Chem 2009;30(16):2785-91.
26. Sanner MF J. Python: a programming language for software integration and development. Graphics Mod 1999;17(2):57-61.