Bioresearch Communications
Volume 01, Issue 02. July 2015.
Journal Homepage: www.bioresearchcommunications.com

Original Article

Immunoinformatics Aided Prediction of Cytotoxic T Cell Epitope of Respiratory Syncytial Virus

Md Nur Ahad Shah, Payal Barua and Md Kawsar Khan*

Department of Biochemistry and Molecular Biology, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh.

ABSTRACT: The Respiratory Syncytial Virus (RSV) poses the threat of lower respiratory tract infection to the infants as well as elderlies. As there is no licensed vaccine, alternative vaccine candidates like Epitope-based vaccines can be considered as a potential candidate. Being conserved among strains and reported to elicit cytotoxic T cell (CTL) response, the fusion glycoprotein of RSV (RSV-FP) is a first-rate target for epitope based vaccine designing. As RSV specific CD8+ CTLs are the central cell of viral clearance, the epitopes capable to generate CTL response are desirable. In this study, available immunoinformatics tools are utilized with a target to predict epitopes on the RSV-FP that elicit strong (CTL) responses. We report seven nine-mer peptides that bind strongly with 17 different HLA, have 100% sequence conservancy and is projected to provide 76.03% population coverage worldwide.

KEYWORDS: epitope based vaccine, RSV-F protein, CTL epitope, human respiratory syncytial virus, immunoinformatics.

CITATION: Shah, M. N. A., Barua, P. and Khan, M. K. 2015. Immunoinformatics Aided Prediction of Cytotoxic T Cell Epitope of Respiratory Syncytial Virus. *Biores Comm.* **1**(2), 99-104.

CORRESPONDENCE: kawsarkhan-bmb@sust.edu

INTRODUCTION

The Respiratory Syncytial Virus (RSV) is the single most important cause of lower respiratory tract infection among infants and young children with a yearly estimated case of 64 million illnesses and 160,000 deaths worldwide.[1] The Virus is classified as a member of the Paramyxoviridae family containing a negative-sense, single-stranded RNA genome having a unique immunopathological feature.[2] RSV infection causes acute bronchiolitis and pneumonia in children under age two years while there is also a significant rate of morbidity and mortality among the elderly people.[3-5] However, currently there is no licensed vaccine for RSV and effective vaccine is the demand of time.[6]

Epitope-based vaccine designing is apparently a new concept that offers a promising route for vaccine development for pathogen like RSV that has resisted traditional vaccine development.[7] Epitope vaccines permit the selection of a specific subset of epitopes capable to evoke the necessary immune response. A peptide must fulfill three criteria to be an ideal CTL epitope. First, the peptide must possess appropriate proteasomal cleavage sites that facilitate processing and generation of peptides of suitable lengths.[8] Secondly, the processed peptide must bind to transporters associated with antigen presentation (TAP) and transported to Golgi apparatus.[8] Finally, the peptide should bind with human leukocyte antigen (HLA) molecules and presented to the cell surface for the recognition of CD8+ T cells.[8] The HLA molecules are highly polymorphic and expressed at variable frequencies in different ethnic groups.[9] As a result of that, the efficacy of vaccine may vary in different geographical locations depending on its HLA coverage. Hence, ensuring broad range HLA coverage is one of the major challenges of the epitope-based vaccines. With the advent of immuno-informatics, we can now determine the CTL epitope with high precision. In this regard, a few web-based tools are freely available in Immune epitope database (IEDB) which has high accuracy in the proteasomal cleavage site, TAP binding and HLA binding prediction.[8] As only a few numbers of peptides meet all this criteria hence the use of immunoinformatics tools can assist to identify the CTL epitopes in a time and cost-effective way.

In human as well as in murine model, RSV-specific CD8+ T cell plays the principal role in the clearance of the virus and recovery from infection.[10-13] Among the RSV proteins, the fusion protein and nucleoprotein proteins are found to be the primary targets for MHC class I-restricted CTL responses.[14-17] RSV-FP is highly conserved among strains and is indispensable for the fusion of the viral envelope with host cell membranes during viral entry.[18] Previous studies in the murine model have reported an immunodominant CTL epitope (KYKNAVTEL) spanning from 83-95 amino acids of RSV-FP.[19] However, despite

the advancement of immunoinformatics, no studies have been performed considering the human HLA, TAP binding affinity and by using the sophisticated *in silico* tools. In this study, with the aid of a number of *in silico* algorithms we aim to predict the CTL epitopes on RSV-FP with a view to construct multi-epitope vaccine candidate for RSV.

MATERIALS AND METHODS

Retrieving the protein sequences

The protein sequences of the Respiratory Syncytial Virus Fusion Glycoprotein (RSV-FP) were retrieved from National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/). All of the available sequences were collected from different geographical locations.

Multiple sequence alignment

Multiple sequence alignments were generated using the retrieved sequences in order to get the conserved regions of the protein sequences. BLOSUM matrix algorithm from MEGA6.06[20] was utilized along with ClustalW tools. A weblogo was generated from the retrieved sequences using the 'Weblogo tool' (Version 2.8.2) (http://weblogo.berkeley.edu/logo.cgi) developed by Computational Genomics Research Group, Department of Plant and Microbial Biology, University of California, Berkeley.[21]

Construction of phylogenetic tree

To elucidate the evolutionary distances among the fusion glycoproteins, a phylogenetic tree were constructed from the generated multiple sequences alignments. MEGA 6.06 software package was employed to construct and visualize the un-rooted phylogenetic tree. The phylogenetic tree was constructed based on bootstrap method with the number of bootstrap replication at 1000.[22] Evolutionary distances that were used to infer the phylogenetic tree were computed using the Poisson correction method.[23]

Calculation of non-synonymous – synonymous substitution ratio

The gene sequences of the fusion glycoproteins from different variants (TaxID: 12814) were retrieved from the NCBI database (www.ncbi.nlm.nih.gov). Applying the Nei-Gojobori (Jukes-Cantor) method from MEGA6.06, an average value of non-synonymous and synonymous substitution was calculated.[24] Furthermore, the values were used to calculate the non-synonymous – synonymous substitution ratio (Ka/Ks).

Identification of cytotoxic T-lymphocyte (CTL) epitope

Available tools from the Immune Epitope Database (IEDB) (http://www.iedb.org) were employed to evaluate the immunogenicity of the RSV-F protein. The IEDB platform provides a range of different algorithms, including Artificial Neural Network (ANN).[25] Average Relative Binding (ARB)[26] and Stabilized Matrix Method (SMM)[27] for peptide-HLA Class I binding prediction.

The 'Proteasomal cleavage/TAP transport/MHC 1 Combined predictor' (http://tools.immuneepitope.org/processing/) tool was applied for the prediction of CTL epitope, which is reported to have the capability of detecting epitopes with 81–97% accuracy.[28,29] Among the available algorithms, the Artificial Neural Network (ANN) prediction method was used, which can predict successful CTL epitopes from a given protein sequence based on their ability to be processed by both the proteasome as well as transporters associated with antigen presentation (TAP). Furthermore, the epitopes binding affinities with available HLA Class I molecules served as an additional tool. The default cut-off value for MHC score -2.7, equivalent to the IC50 value of 500 nM, was used.[30-32] A threshold value 1.00 was selected for proteasomal cleavage and 1.14 for TAP binding, corresponding to expected specificity rate 76%.[33]

Epitope conservation analysis

To analyze the conservation status of the selected epitopes, 'Conservation Across Antigens' tool was applied from IEDB. The tool (http://tools.immuneepitope.org/tools/conservancy/iedb_input) detects the conservancy of an epitope from the provided protein sequences.[34]

Determination of cross-reacting human self-peptides in the predicted epitope

In order to determine the presence of human peptide analogs identical to the predicted epitopes, the NCBI BlastP program (http://www.ncbi.nlm.nih.gov/BLAST/) was used. The predicted epitopes were compared to the non-redundant protein sequence database restricted to human (TaxID: 9606). The parameters were adjusted as used in previous studies[31] briefly as follows: word size of two, expected threshold value of 30,000, 'no adjustment' setting for composition-based statistics and PAM30 as the matrix of choice.[35] Hits involving matches of fewer than six amino acids were ignored on account of the likely probability of such matches being random.

Three dimensional structures of the predicted epitopes

The three dimensional structures of the epitopes were designed using the PEP-FOLD Peptide structure prediction server (http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD/).[36] To view the epitope structures, PyMOL (Version 1.7.4) were used (http://www.pymol.org).

Population coverage prediction

To predict the population coverage, a web-based Population Coverage Analysis tool was utilized from IEDB (http://tools.immuneepitope.org/tools/population/iedb_input).[37] Using HLA allele genotypic frequencies obtained from The Allele Frequency Net Database,[38] this algorithm calculates the fraction of individuals predicted to respond to a given set of epitopes with known HLA restrictions. At present, The Allele Frequency Net Database contains frequencies of 3,245 alleles (including class I and class II) for the world, 16 geographical areas, 21 ethnicities, 115 countries and ethnicities by country.

RESULTS AND DISCUSSION

Evolutionary divergence in RSV-FP

A total of 94 sequences (supplementary file 1) were retrieved from different geographical locations. The conservancy of the retrieved sequences was observed from the multiple sequence alignments and from the

Figure 1. Web logo of the sequence variants of the RSV F protein. One of the conserved positions is shown with arrow and an unconserved position is shown with circle.

Table 1. CTL epitopes predicted by 'Proteasomal cleavage/ TAP transport/ MHC 1 Combined predictor'.

Allele	Peptide	Proteasome Score	TAP Score	MHC Score	Processing Score	Total Score	MHC IC50[nM]
HLA-A*01:01	VSVGNTLYY	1.23	1.35	-1.79	2.58	0.79	61
HLA-A*01:01	KTFSNGCDY	1.1	1.36	-2.67	2.46	-0.21	467
HLA-A*01:01	NTDIFNSKY	1.11	1.2	-0.85	2.3	1.46	7
HLA-A*03:01	KTFSNGCDY	1.1	1.36	-2.65	2.46	-0.19	450
HLA-A*11:01	VSVGNTLYY	1.23	1.35	-2.2	2.58	0.38	157
HLA-A*11:01	TVSVGNTLY	1.32	1.26	-2.65	2.58	-0.07	449
HLA-A*29:02	VSVGNTLYY	1.23	1.35	-0.9	2.58	1.68	8
HLA-A*29:02	TVSVGNTLY	1.32	1.26	-1.04	2.58	1.54	11
HLA-A*29:02	KTFSNGCDY	1.1	1.36	-1.83	2.46	0.63	68
HLA-A*29:02	SLGAIVSCY	1.19	1.23	-1.88	2.42	0.55	75
HLA-A*30:01	KTFSNGCDY	1.1	1.36	-2.29	2.46	0.17	196
HLA-A*30:02	VSVGNTLYY	1.23	1.35	-1.79	2.58	0.79	61
HLA-A*30:02	TVSVGNTLY	1.32	1.26	-2.18	2.58	0.4	151
HLA-A*30:02	KTFSNGCDY	1.1	1.36	-1.49	2.46	0.97	31
HLA-A*30:02	SLGAIVSCY	1.19	1.23	-2.21	2.42	0.22	162
HLA-A*30:02	NTDIFNSKY	1.11	1.2	-2.7	2.3	-0.39	498
HLA-A*30:02	LTRTDRGWY	1.42	1.28	-1.6	2.7	1.1	40
HLA-A*30:02	AYVVQLPIY	1.4	1.5	-1.75	2.9	1.16	56
HLA-A*68:01	TVSVGNTLY	1.32	1.26	-2.32	2.58	0.26	210
HLA-B*08:01	KRKRRFLGF	1.27	1.28	-1.81	2.55	0.74	65
HLA-B*15:01	VSVGNTLYY	1.23	1.35	-2.11	2.58	0.47	128
HLA-B*15:01	TVSVGNTLY	1.32	1.26	-2.69	2.58	-0.12	495
HLA-B*15:01	KTFSNGCDY	1.1	1.36	-2.09	2.46	0.37	124
HLA-B*15:01	SLGAIVSCY	1.19	1.23	-2	2.42	0.43	99
HLA-B*15:02	TVSVGNTLY	1.32	1.26	-2.64	2.58	-0.06	432
HLA-B*15:02	SLGAIVSCY	1.19	1.23	-2.55	2.42	-0.13	358
HLA-B*27:05	KRKRRFLGF	1.27	1.28	-1.72	2.55	0.83	52
HLA-B*35:01	VSVGNTLYY	1.23	1.35	-2.23	2.58	0.35	169
HLA-B*35:01	TVSVGNTLY	1.32	1.26	-2.01	2.58	0.56	103
HLA-B*57:01	KTFSNGCDY	1.1	1.36	-2.07	2.46	0.39	118
HLA-B*58:01	VSVGNTLYY	1.23	1.35	-1.36	2.58	1.22	23
HLA-B*58:01	KTFSNGCDY	1.1	1.36	-2.13	2.46	0.33	134
HLA-C*05:01	NTDIFNSKY	1.11	1.2	-2.06	2.3	0.24	115
HLA-C*14:02	AYVVQLPIY	1.4	1.5	-2.65	2.9	0.26	446

generated weblogo (Figure 1). The generated phylogram shows the evolutionary divergence among RSV-FP (supplementary file 2).

Identification of evolutionarily conserved region in fusion glycoprotein

From the output of multiple sequence alignments, some peptide sequences were found to be conserved in all variants while the rest of the regions were variable in different sequences. The most conserved sequence of 574 amino acids in length was used for further analyses (http://www.ncbi.nlm.nih.gov/protein/38230490/).

To identify the conservancy of fusion glycoprotein under selective pressure non-synonymous – synonymous ratio was calculated using the MEGA6.06 software. The ratio greater than one implies positive or Darwinian selection, less than one implies purifying (stabilizing) selection and one indicates neutral selection. The average values for non-synonymous and synonymous mutations were 0.005 and 0.144 respectively. The ratio was 0.0347, which indicates the protein follows the purifying selection and hardly shows any mutation over the evolutionary history.

CTL epitope prediction in the RSV-F protein

The Artificial Neural Network (ANN) prediction method of 'Proteasomal cleavage/ TAP transport/ MHC 1 combined predictor' tool predicted 220 nine-merepitopes whose IC50 value was lower than 500nM. Among them, the proteasomal cleavage score of 166 epitopes was higher or equal to one. Finally, 21 of those 166 epitopes were found to have the TAP score higher or equal to 1.14 (Table 1). These 21 epitopes were further tested for their conservancy.

After epitope conservancy analysis, eight epitopes were selected as they showed 100% protein sequence match. Among the eight epitopes, one epitope (LTRTDRGWY) was found to bind with only one HLA (HLA-A*30:02) allele. The HLA-A*30:02 allele was also predicted to binds with six of the rest seven epitopes, hence the epitope fails to bind with any unique HLA alleles and thus was deliberately ignored from further analysis. The KTFSNGCDY epitope was predicted to binds with maximum eight different HLAs (HLA-A*01:01, HLA-A*03:01, HLA-A*30:01, HLA-B*58:01, HLA-B*15:01, HLA-B*57:01, HLA-A*29:02, HLA-A*30:02) (Table 2). On the other hand, AYVVQLPIY and KRKRRFLGF, each epitope was found to binds with lowest (two) number of epitopes (Table 2).

Cumulatively all seven epitopes were predicted to bind with 17 different HLA-A, HLA-B and HLA-C molecules. Among the seven epitopes, KTFSNGCDY, VSVGNTLYY and NTDIFNSKY was found to bind with HLA-A*01:01 and KTFSNGCDY was predicted to bind with HLA-B*57:01 allele. Both of the HLA (HLA-A*01:01 and HLA-B*57:01) alleles were previously reported and characterized to elicit strong CTL response against an RSV-FP epitope both *in vivo* and *in vitro* conditions.[39,40]

Structural configuration of epitopes

The epitopes three-dimensional structures were observed via PyMOL (Figure 2). All of the epitopes' secondary structures are in alpha-helical configuration except for the epitope NTDIFNSKY. For this epitope, the residues 4-IFNS-7 are in beta-sheet configuration and the remaining residues are in the alpha-helical configuration.

Human cross-reactivity analysis

Sometimes whole cell and subunit vaccines in their natural forms may contain protein sequences mimicking native human peptides. We tested the occurrence of

Table 2. Selected CTL epitopes with their corresponding HLAs.

Number	Epitope	HLA to interact with
1	VSVGNTLYY	HLA-B*35:01, HLA-A*11:01, HLA-B*15:01, HLA-A*01:01, HLA-A*30:02, HLA-B*58:01, HLA-A*29:02
2	TVSVGNTLY	HLA-B*15:01, HLA-A*11:01, HLA-B*15:02, HLA-A*68:01, HLA-A*30:02, HLA-B*35:01, HLA-A*29:02
3	KTFSNGCDY	HLA-A*01:01, HLA-A*03:01, HLA-A*30:01, HLA-B*58:01, HLA-B*15:01, HLA-B*57:01, HLA-A*29:02, HLA-A*30:02
4	SLGAIVSCY	HLA-B*15:02, HLA-A*30:02, HLA-B*15:01, HLA-A*29:02
5	NTDIFNSKY	HLA-A*30:02, HLA-C*05:01, HLA-A*01:01
6	AYVVQLPIY	HLA-C*14:02, HLA-A*30:02
7	KRKRRFLGF	HLA-B*08:01, HLA-B*27:05

Table 3. Determination of human peptides analogous to the predicted epitopes. Amino acids in bold correspond to matched human peptide sequences. Positions of matched sequences within the human protein are denoted by numbers on either side of the peptide.

Epitope Sequence	Human Peptide	Human Protein Name	GenBank ID
VSVGNTLYY	35 VSVGNT 40	T cell receptor alpha	AIE10864.1
V**SVGNTL**YY	238 SVGNTL 243	tetratricopeptide repeat domain 30A	AAH42848.1
T**VSVGNT**LY	35 VSVGNT 40	T cell receptor alpha, partial	AIE10864.1
TV**SVGNTL**Y	238 SVGNTL 243	Tetratricopeptide repeat domain 30A	AAH42848.1
TVSVGNTLY	715 TVSVGN 720	vault protein	AAD47250.1
TVSVGNTLY	716 TVSVGN 721	KIAA0177 protein	BAA11494.2
A**YVVQLP**IY	252 YVVQLP 257	L1 cell adhesion molecule, partial	AAO17583.1
KRKRRFLGF	127 KRKRRF 132	histone 1, H2bn, isoform CRA_b	EAX03115.1

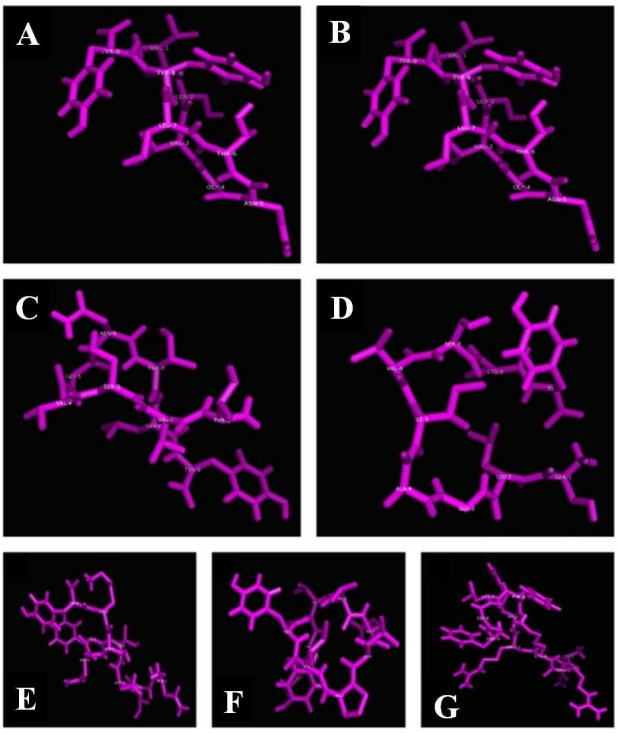

Figure 2. 3D structures of the predicted epitopes. **A:** VSVGNTLYY, **B:** TVSVGNTLY, **C**: KTFSNGCDY, **D**: SLGAIVSCY, **E**: NTDIFNSKY, **F**: AYVVQLPIY and **G**: KRKRRFLGF.

Table 4. Percentage of population covered by the selected epitopes in different geographical areas.

Geographical area	Population coverage[a]	Average hit[b]	PC90[c]
East Asia	57.64%	1.71	0.24
Northeast Asia	73.15%	2.26	0.37
South Asia	74.44%	2.07	0.39
Southeast Asia	60.59%	1.58	0.25
Southwest Asia	61.64%	1.65	0.26
Europe	83.64%	2.64	0.61
East Africa	64.88%	2.33	0.28
West Africa	68.60%	2.34	0.32
Central Africa	58.46%	1.96	0.24
North Africa	69.55%	2.15	0.33
South Africa	78.04%	2.95	0.46
West Indies	71.78%	2.52	0.35
North America	74.26%	2.33	0.39
Central America	4.14%	0.13	0.10
South America	48.83%	1.18	0.20
Oceania	61.33%	1.56	0.26

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.

contiguous sequences identical to the epitopes in the human non-redundant protein database. The maximum identical protein sequences with six amino acids were found for four epitopes (Table 3). In case of the rest three epitopes (KTFSNGCDY, SLGAIVSCY, NTDIFNSKY) the maximum consecutive identical amino acids were less than six amino acids. The absence of the identical sequences of more than six amino acids greatly reduces the likelihood of triggering an autoimmune response by the epitopes.

Population coverage of the epitopes

Due to the vast polymorphism of HLA molecules in different ethnic groups, the immune responses elicited by T-cell epitopes may vary from one population to another. Also, the smaller size of the vaccine often results in a limited number of epitopes with low HLA allele coverage. These factors often restrict the broad use of epitopes as potential vaccines. We have calculated the population coverage of the seven epitopes in 16 different geographical regions. The epitopes are predicted to provide 76.03% population coverage worldwide which signifies the goodness of the proposed epitopes. Moreover, highest 83.64% coverage was projected in Europe followed by 78.04% in South Africa, 74.44% in South Asia and 74.26% in North America. On the other hand lowest 4.14% population coverage was predicted in Central America (Table 4).

CONCLUSION

With the advancement of immunoinformatics, a plethora of new vaccine candidates is being identified, characterized and deposited in the public databases now. In current study, we have predicted a cohort of seven epitopes in the conserve region of RSV-FP which is predicted to provide CTL responses in human. The epitope cohort is predicted to bind with 17 different HLA alleles and projected to provide 76.03% population coverage worldwide. We propose a multi-epitope vaccine of the seven epitopes of RSV-FP which might induce protective immune response against RSV infection. However, *in vivo* and *in vitro* studies are required to determine the true effectiveness of the proposed multi-epitope vaccine.

ABBREVIATIONS

Respiratory Syncytial Virus (RSV), Cytotoxic T Lymphocytes (CTL), Respiratory Syncytial Virus Fusion Protein (RSV-FP), Major Histocompatibility Complex (MHC), Human Leukocyte Antigen (HLA), Transporters Associated with antigen Presentation (TAP), National Center for Biotechnology Information (NCBI), Immune Epitope Database (IEDB), Artificial Neural Network (ANN), Average Relative Binding (ARB), Stabilized Matrix Method (SMM), Point Accepted Mutation (PAM), BLOck SUbstitution Matrix (BLOSUM).

REFERENCES

1. World Health Organization. Acute respiratory infection. World Health Organization; Geneva:[Accessed on 28 May 2015].
2. Bueno, S.M., González, P.A., Pacheco, R., et al. 2008. Host immunity during RSV pathogenesis. *Int. Immunopharmacol.* **8(10)**, 1320–9.
3. Stott, E.J. and Taylor, G. 1985. Respiratory syncytial virus. Brief review. *Arch. Virol.* **84(1–2)**, 1–52.
4. Falsey, A.R., Cunningham, C.K., Barker, W.H., et al. 1995. Respiratory syncytial virus and influenza A infections in the hospitalized elderly. *J. Infect. Dis.* **172(2)**, 389–394.
5. Thompson, W.W., Shay, D.K., Weintraub, E., et al. 2003. Mortality associated with influenza and respiratory syncytial virus in the United States. *JAMA.* **289**, 179–186.
6. Kim, H.W., Canchola, J.G., Brandt, C.D., et al. 1969. Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. *Am. J. Epidemiol.* **89**, 422–434.

7. Burton, D.R., Poignard, P., Stanfield, R.L., et al. 2012. Broadly neutralizing antibodies present new prospects to counter highly antigenically diverse viruses. *Science*. **337**, 183–186.
8. Larsen, M., Lundegaard, C., Lamberth, K., et al. 2005. An integrative approach to CTL epitope prediction: A combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions. Eur. *J. Immunol.* **35**: 2295–2303.
9. Ovsyannikova, I.G., Dhiman, N., Jacobson, R.M., et al. 2006. Human leukocyte antigen polymorphisms: variable humoral immune responses to viral vaccines. Expert Rev. *Vaccines*. **5**, 33-43.
10. Cannon, M. J., Openshaw, P. J., &Askonas, B. A. 1988. Cytotoxic T cells clear virus but augment lung pathology in mice infected with respiratory syncytial virus. *The Journal of experimental medicine*, ***168*(3)**, 1163-1168.
11. Graham, B. S., Bunton, L. A., Wright, P. F., &Karzon, D. T. 1991. Role of T lymphocyte subsets in the pathogenesis of primary infection and rechallenge with respiratory syncytial virus in mice. *Journal of Clinical Investigation*, **88(3)**, 1026.
12. Chiba, Y., Higashidate, Y., Suga, K., Honjo, K., Tsutsumi, H., &Ogra, P. L. 1989. Development of cell-mediated cytotoxic immunity to respiratory syncytial virus in human infants following naturally acquired infection. *Journal of medical virology*, **28(3)**, 133-139.
13. Bangham, C. R., & McMichael, A. J. 1986. Specific human cytotoxic T cells recognize B-cell lines persistently infected with respiratory syncytial virus. *Proceedings of the National Academy of Sciences*, **83(23)**, 9183-9187.
14. Bangham, C. R., Openshaw, P. J., Ball, L. A., King, A. M., Wertz, G. W., &Askonas, B. A. 1986. Human and murine cytotoxic T cells specific to respiratory syncytial virus recognize the viral nucleoprotein (N), but not the major glycoprotein (G), expressed by vaccinia virus recombinants. *The Journal of Immunology*, **137(12)**, 3973-3977.
15. Pemberton, R. M., Cannon, M. J., Openshaw, P. J. M., Ball, L. A., Wertz, G. W., &Askonas, B. A. 1987. Cytotoxic T cell specificity for respiratory syncytial virus proteins: fusion protein is an important target antigen. *J. Gen. Virol*, **68(8)**, 2177-2182.
16. Goulder, P. J., Lechner, F., Klenerman, P., McIntosh, K., & Walker, B. D. 2000. Characterization of a novel respiratory syncytial virus-specific human cytotoxic T-lymphocyte epitope. *Journal of virology*, **74(16)**, 7694-7697.
17. Brandenburg, A. H., de Waal, L., Timmerman, H. H., Hoogerhout, P., De Swart, R. L., & Osterhaus, A. D. M. E. 2000. HLA class I-restricted cytotoxic T-cell epitopes of the respiratory syncytial virus fusion protein. *Journal of virology*, **74(21)**, 10240-10244.
18. Kahn, J.S., Schnell, M.J., Buonocore, L., Rose, J.K. 1999. Recombinant vesicular stomatitis virus expressing respiratory syncytial virus (RSV) glycoproteins: RSV fusion protein can mediate infection and cell fusion. *Virology*. **254(1)**, 81–91
19. Chang, J., Srikiatkhachorn, A., & Braciale, T. J. (2001). Visualization and characterization of respiratory syncytial virus F-specific CD8+ T cells during experimental virus infection. *The Journal of Immunology*, **167(8)**, 4254-4260.
20. Tamura, K., Stecher, G., Peterson, D., et al. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*.**30**, 2725-2729.
21. Crooks, G.E., Hon, G., Chandonia, J.M., et al. 2004. WebLogo: A sequence logo generator. *Genome Research.* **14(6)**, 1188-1190.
22. Nei, M. and Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York. p. 333.
23. Kumar, S., Nei, M., Dudley, J., et al. 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.* **9**, 299-306.
24. Nei, M. and Gojobori, T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution.* **3(5)**, 418-426.
25. Nielsen, M., Lundegaard, C., Worning, P., et al. 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci.* **12**, 1007–1017.
26. Bui, H.H., Sidney, J., Peters, B., et al. 2005. Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications. *Immunogenetics.* **57**, 304–314.
27. 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.
28. Tenzer, S., Peters, B., Bulik, S., et al. 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**, 1025-1037.
29. Schellens, I.M., Kesmir, C., Miedema, F., et al. 2008. An unanticipated lack of consensus cytotoxic T lymphocyte epitopes in HIV-1 databases: the contribution of prediction programs. *AIDS*. **22**, 33–37.
30. Assarsson, E., Sidney, J., Oseroff, C., et al. 2007. A quantitative analysis of the variables affecting the repertoire of T cell specificities recognized after vaccinia virus infection. *J. Immunol.* **178**, 7890–7901.
31. Peters, B., Bui, H.H., Frankild, S., et al. 2006. A community resource benchmarking predictions of peptide binding to MHC-I molecules. *PLoSComput. Biol.* **2**, e65.
32. Khan, M. K., Zaman, S., Chakraborty, S., et al. 2014. In silico predicted mycobacterial epitope elicits in vitro T-cell responses. *Molecular immunology*, **61(1)**, 16-22.
33. Burroughs, N.J., de Boer, R.J., Kesmir, C. 2004. Discriminating self from nonself with short peptides from large proteomes. *Immunogenetics.* **56(5)**, 311–320.
34. Bui, H.H., Sidney, J., Li, W., et al. 2007. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics.* **8(1)**, 361.
35. Tan, P.T., Heiny, A.T., Miotto, O., et al. 2010. Conservation and diversity of influenza A H1N1 HLA-restricted T cell epitope candidates for epitope-based vaccines. *PLoS ONE* .**5**, e8754.
36. Thévenet, P., Shen, Y., Maupetit, J., et al. 2012. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Res.* **40**, W288-293.
37. Bui, H.H., Sidney, J., Dinh, K., et al. 2006. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics.***7**, 153.
38. Gonzalez-Galarza, F.F., Takeshita, L.Y., Santos, E.J., et al. 2015. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acid Research.* **43**, D784-788.
39. Venter, M., Rock, M., Puren, A.J., et al. 2003. Respiratory syncytial virus nucleoprotein-specific cytotoxic T-cell epitopes in a South African population of diverse HLA types are conserved in circulating field strains. *J. Virol.* **77**, 7319–7329.
40. Jurcevic, S., Travers, P.J., Hills, A., et al. 1996. Distinct conformations of a peptide bound to HLA-DR1 or DRB5*0101 suggested by molecular modelling. *Int. Immunol* .**8**, 1807–1814.
41. Vordermeier, H.M., Harris, D.P., Moreno, C., et al. 1994. Promiscuous T cell recognition of an H-2 IA-presented mycobacterial epitope. *Eur. J. Immunol.* **24(9)**, 2061–2067.
42. Schipper, R.F., Van Els, C.A., D'Amaro, J., et al. 1996. Minimal phenotype panels. A method for achieving maximum population coverage with a minimum of HLA antigens. *Hum. Immunol.* **51(2)**, 95-98.