Momtaz & Foysal *Afr. J. Clin. Exper. Microbiol.* 2019; 20 (2): 104 – 116 https://www.afrjcem.org

African Journal of Clinical and Experimental Microbiology ISSN 1595-689X April 2019 Vol.20 No.2
AJCEM/1914: https://www.ajol.info/index.php/ajcem

Original Article **Open Access**

An immunoinformatic approach to design a novel vaccine against the human respiratory syncytial virus (hRSV) by targeting M2-1 protein

[1]Momtaz, F., and [2,3*]Foysal, M. J

[1]Department of Microbiology, University of Chittagong, Chittagong 4331, Bangladesh
[2]Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh
[3]School of Molecular and Life Sciences, Curtin University, WA 6845, Australia
Correspondence to: mjfoysal-geb@sust.edu and mdjaved.foysal@postgrad.curtin.edu.au

Abstract:

Background: Human respiratory syncytial virus (hRSV) is the leading cause of upper and lower respiratory infection in infants, adults and immunocompromised persons. The matrix protein, M2-1 of hRSV is a cofactor of viral RNA polymerase that plays a crucial role during replication. This programmed study was designed to scrutinize potential immunogens from the M2-1 protein characterized from four different continents.
Methods: Sequence data obtained from NCBI databases were analysed by using a series of web and software based bioinformatics tools to find out the best epitope against hRSV.
Results: The phylogenetic data revealed a homogenized clustering of M2-1 protein for the African, European, and Asian clades while proteins from North American collections found to have a significant evolutionary detachment compared to three other clusters. Using various web-based bioinformatics tools, the study identified four common B-cell epitopes present in all the M2-1 proteins from four different clusters with higher antigenicity and conservancy. Among the 17 M2-1 protein investigated for T-cell epitopes, "VLQNLDVGL" peptide from A2 super-type, and "QSACVAMSK" and "CLNGRRCHY" from A3 super-type showed the highest antigenicity at >0.80 conservancy cut-off value. After evaluation of all antigenic properties, only "CLNGRRCHY" peptide qualified as a potential vaccine candidate against hRSV. Molecular docking revealed strong and stable binding of the epitope to major histocompatibility complexes (MHC) molecules in terms of hydrogen bonding.
Conclusion: The designed epitope could be used as a possible vaccine candidate against hRSV.

Keywords: hRSV; M2-1 protein; phylogenetic cluster; BCL and CTL epitopes; molecular docking

Received January 29, 2019; Revised February 17, 2019; Accepted February 20, 2019

Une approche immuno-informatique pour concevoir un nouveau vaccin contre le virus respiratoire syncytial (VRSH) humain en ciblant la protéine M2-1

[1]Momtaz, F., and [2,3*]Foysal, M. J

[1]Département de microbiologie, Université de Chittagong, Chittagong 4331, Bangladesh
[2]Département de génie génétique et de biotechnologie, Université des sciences et technologies de Shahjalal, Sylhet 3114, Bangladesh
[3]Ecole des sciences moléculaires et de la vie, Université Curtin, WA 6845, Australie
Correspondance à: mjfoysal-geb@sust.edu and mdjaved.foysal@postgrad.curtin.edu.au

Abstrait

Contexte: Le virus respiratoire syncytial (VRSH) humain est la principale cause d'infection des voies respiratoires supérieures et inférieures chez les nourrissons, les adultes et les personnes immunodéprimées. La protéine matricielle M2-1 du hRSV est un cofacteur de l'ARN polymérase virale qui joue un rôle crucial lors de la réplication. Cette étude programmée a été conçue pour examiner les immunogènes potentiels de la
Méthodes: Les données de séquence obtenues des bases de données NCBI ont été analysées à l'aide d'une série d'outils bioinformatiques basés sur le Web et sur les logiciels, afin de déterminer le meilleur épitope contre le hRSV.
Résultats: Les données phylogénétiques ont révélé un regroupement homogénéisé de la protéine M2-1 pour les clades africain, européen et asiatique, tandis que les protéines des collections nord-américaines se sont révélées avoir un important détachement évolutif par rapport à trois autres groupes. À l'aide de divers outils bioinformatiques basés sur le Web, l'étude a identifié quatre épitopes de cellules B communs présents dans toutes les protéines M2-1 de quatre groupes différents présentant une antigénicité et une conservation plus élevées. Parmi les 17 protéines M2-1 étudiées pour les épitopes de lymphocytes T, le peptide «VLQNLDVGL» de super type A2, et «QSACVAMSK» et «CLNGRRCHY» de super type A3 ont montré la plus grande antigénicité à une valeur de coupure> 0,80. Après évaluation de toutes les propriétés antigéniques, seul le peptide «CLNGRRCHY» a été qualifié de candidat vaccin potentiel contre le hRSV. L'amarrage moléculaire a révélé une liaison forte et stable de l'épitope aux molécules du complexe majeur d'histocompatibilité (MHC) en termes de liaison hydrogène.
Conclusion: l'épitope conçu pourrait être utilisé comme vaccin potentiel contre le hRSV.

Mots-clés: hRSV; Protéine M2-1; groupe phylogénétique; Épitopes BCL et CTL; amarrage molèculaire

Introduction

Bronchiolitis is a severe respiratory tract infection caused by *Paramyxoviridae* family virus, human respiratory syncytial virus (hRSV). It is the leading cause of mortality in children under two years of age (1). It has also been reported to infect elderly and immunocompromised people, especially pregnant women and transplant patients (2, 3). Currently, there is no candidate vaccine or drug available to prevent hRSV infection in the young and adults. Although, some of the drugs showed promising preclinical trial results, their long-term application remained challenging due to problem of post-clinical trial development, adverse effects including allergenicity, and short-term efficacy in humans (4, 5). For instance, treatment with ribavirin is not only expensive but the drug also cause aerosol related side-effects (6).

M2-1 protein is an essential co-factor of hRSV viral RNA polymerase complex that plays a crucial function in transcription processivity, acting as an anti-termination factor. The M2-1 protein binds to RNA dependent RNA polymerase (RdRp) and also interacts with other components of viral RNA polymerase by its core domain at the N-terminus (7, 8). Inhibition of M2-1 protein activity caused a significant reduction of viral RNA transcription and viral assembly (7).

In-silico epitope-based drug design can provide rapid, reliable, inexpensive, and safe vaccine against targeted antigen using immunoinformatics platforms. Epitopes are usually 9–mer peptides from viral proteins that can generate potent antigenicity when directed against specific antigen(s) (9). In the post-genomic era, epitope-based computer-aided vaccine has proven to be successful against most of the pathogenic viruses including human viral pathogens such as Influenza, Chikungunya, Rota, Zika, Ebola, MARS-CoV, etc., and fish pathogens such as Edwardsiella, Flavobacterium, and shrimp

white spot virus (WSSV) (9–17). In humans particularly, proper definition and differentiation of major histocompatibility complex molecules (MHC class I and II), human leucocyte antigen (HLA), and readily available enhanced databases, make *in-silico* approach widely acceptable for vaccine design against viral and bacterial pathogens (17). The aim of the present study therefore was to discover potent B-cell and T-cell epitope(s) from the M2-1 matrix protein that can bind strongly to human MHC molecules and subsequently inhibit RSV replication in the host cell.

Materials and methods:

Retrieval of M2-1 protein sequences from the database

M2-1 protein sequences were downloaded in FASTA format from the National Center for Biotechnology Information (NCBI) protein data bases (https://www.ncbi.nlm.nih.gov/protein). M2-1 protein was selected for antigenicity screening and downstream bioinformatics analysis because of its role and association during respiratory infection (5,18,19). In the present study, we collected information on 17 M2-1 protein sequences from four different continents namely Asia, Africa, Europe, and North America for further immunoinformatics.

Although information about the epidemiology of hRSV infections is sparse in the literature, several outbreaks have been reported from these four continents (20, 21). Sequences having same isolation history and date were excluded in the present study.

Phylogenetic clustering

Extracted protein sequences were aligned using muscle alignment program in MEGA 7.0. Muscle alignment has been reported to be a highly efficient tool for protein sequence alignment and phylogenetic analysis (22). The phylogenetic tree was constructed by the neighbour-joining method in MEGA 7.0 using Kimura-J model where branches in the tree separated by 1000 bootstrap replicate (22). The evolutionary divergence in tree was calculated as p-distance for both transition and trans-version in default parameters.

Sorting antigenic protein

To determine the antigenic protein of M2-1 that can elicit necessary immune response, we used VaxiJen (v2.0) online based server (http://ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) for the classification of 17 sequences based on immunity score. This server can predict antigenicity of any protein sequence with 90% accuracy, and therefore, widely used for *in-silico* based drug design (23). The highest antigenic M2-1 sequences were selected for further studies.

Prediction of B-cell linear (BCL) epitopes

Immune Epitope Database (IDEB) (http://tools.iedb.org/main/) has different antigenicity analysis tools based on the protein sequences. Among the approaches, Kolaskar and Tongaonkar 1990 method has been widely used for B-cell linear epitope prediction due to its accuracy (>75%) (9). The tool can also efficiently stratify epitopes from the large group of protein sequences based on antigenic scores (17).

Helper T-lymphocytes (HTL) epitopes prediction

HTLs epitopes from the M2-1 protein sequences from four different clusters were screened by utilizing IEDB database for class II epitope prediction tool (http://tools.iedb.org/mhcii/) (24). In the webpage of IEDB, various prediction methods are available, but we chose the recommended option for the finding of best epitopes, and followed the process described earlier (25). The human was selected as the targeted species and recommended 27 MHC alleles were screened with >99% population coverage (26).

The selected MHC II alleles were; (i) HLA DRB1/01:01; (ii) HLA DRB1/03:01; (iii) HLA DRB1/04:01; (iv) HLA DRB1/04:05; (v) HLA DRB1/07:01; (vi) HLA DRB1/08:02; (vii) HLA DRB1/09:01;

(viii) HLA DRB1/11:01; (ix) HLA DRB1/12:01; (x) HLA DRB1/13:02; (xi) HLA DRB1/15:01; (xii) HLA DRB3/01:01; (xiii) HLA DRB3/02:02; (xiv) HLA DRB4/01:01; (xv) HLA DRB5/01:01; (xvi) HLA DQA1/05:01/DQB1/02:01; (xvii) HLA DQA1/05:01/DQB1/03:01; (xviii) HLA DQA1/03:01/DQB1/03:02; (xix) HLA DQA1/04:01/DQB1/04:02; (xx) HLA DQA1/01:01/DQB1/05:01; (xxi) HLA DQA1/01:02/DQB1/06:02; (xxii) HLA DPA1/02:01/DPB1/01:01; (xxiii) HLA DPA1/01:03/DPB1/02:01; (xxiv) HLA DPA1/01/DPB1/04:01; (xxv) HLA DPA1/03:01/DPB1/04:02; (xxvi) HLA DPA1/02:01/DPB1/05:01 and (xxvii) HLA DPA1/02:01/DPB1/14:01. During HTL epitope screening, allelic population coverage for four different clusters were taken into consideration so as to find more epitopes for wider coverage (>1 clusters).

Cytotoxic T lymphocyte epitopes prediction and conservancy analysis

Cytotoxic T-cells play a crucial role in generating MHC class I cellular response. They usually performed various functions including destruction of damaged, unresponsive, infected and cancerous cells by recognizing presented epitopes by MHC molecules on the cell surface. The CTL epitopes that can bind MHC class I molecule were predicted from CTLPred (http://crdd.osdd.net/raghava/ctlpred/) (27). The MHC binding ability of the predicted epitopes was measured by implementation of the artificial neural network (ANN) and stabilized matrix method implementation (SMM) at sensitivity cut-off value of 0.80. The 9-mer peptide (length) was selected because of its high binding ability (>75%) to MHC class I and class II molecules (28). The super-type A2, A3, and A7 were selected as a subtype to cover approximately 90% of the population around the globe (29). The C-terminal cleavage weight and tap transport efficiency cut-off value were set at 0.15 and 0.05, respectively.

The prediction threshold score was set at 0.80. The predicted binding scores of the CTL peptides against major human MHC molecules were calculated from the T-epitope designer portal http://www.bioinformation.net/script/hla_search.cgi. As conserved epitopes can provide border protection, therefore, we aimed to identify peptides with the higher conservancy. The conservancy values of the CTL epitopes were calculated from the IEDB conservancy analysis tool using NCBI protein reference data set NCBI (http://tools.iedb.org/ncbi_seq_browser/) (26).

Allergenicity of the predicted epitopes

The selection of non-allergen epitopes was one of the prime aims of this study. The allergenicity of the final epitopes those from four different clusters, having higher antigenicity scores and MHC binding affinity were evaluated using AllerTOP (v2.0) server at http://www.pharmfac.net/allertop/. This server transforms auto-cross covariance (ACC) into equal length vectors. Then the server classifies the protein to either known allergen or non-allergen based on the k-nearest neighbouring score after comparing 4420 allergenic and non-allergenic proteins from the same and different species (25).

Structure prediction and molecular docking of the peptide to MHC molecules

PEP-FOLD3 is a denovo peptide prediction server (http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/) from corresponding amino acid sequences. PEP-FOLD3 can generate 3D models of peptide-based on greedy strategy by using information from the input epitope sequences (30). Molecular docking is the widely used tool to measure the binding affinity of any ligand to receptor molecules (31). Molecular docking of predicted of the predicted peptide to human leucocyte antigens, HLA-A0201 (PDB code 1HLA) and HLA-B*3508 (PDB code 1ZHL) were performed in CABS-DOCK server (http://biocomp.chem.uw.edu.pl/CABSdock) with default parameters. CABS-dock provides efficient and flexible docking with high accuracies (>80%) without predefined localization of the binding sites

(32, 33). Visualization tool PyMol (v2.0.5) was used to extract the 3D structure of docked protein-peptide complexes. The PyMol graphics system is capable of providing excellent visualization platform for analysing the efficacy of computer-aided drugs (34).

Population coverage

Population coverage is usually used to check whether the predicted final epitope and its HLA alleles can cover a significant percent of the world population or not. The IEDB population coverage tool (http://tools.immuneepitope.org/tools/population/iedb) was used to calculate the cumulative percent of world wide population coverage for the final predicted epitope for both MHC classes, as described previously (35).

Results:

Phylogeny and antigenicity of M2-1 protein sequences from four different clusters

Analysis of 17 M2-1 hRSV proteins from four different geographic locations revealed uniform clustering of sequences from Asian, African, and European continents. However, North American (USA) cluster had significant evolutionary divergence and distributed erratically in the phylogenetic tree (Figure 1).

Multiple alignments using muscle found a highly conserved region in M2-1 protein for all four clusters in position 7-26 and 82-101. The protein sequence Q5MKM1.1 from North American (USA) cluster had the highest evolutionary detachment value of 0.579 with other M2-1 sequences, followed by Q84132.1 (detachment value 0.164). African M2-1 protein cluster was found to be phylogenetically more detached from the other three groups. Preliminary screening of antigenicity for 17 M2-1 protein revealed Q84132.1 as the highest antigenic protein with the value of 0.5063, followed by Q5MKM1.1 (0.4787), APW78912.1 (0.4647, European cluster), APW78659.1 (0.4647, European cluster) respectively (Table 1).

Table 1: VaxiJen score of RSV M2-1 protein from four different clusters

Accession number	VaxiJen Score	Cluster
NP_056864.1	0.4	North American Cluster (NAC)
Q84132.1	0.5063	NAC
Q5MKM1.1	0.4787	NAC
AIY60641.1	0.4134	NAC
ASV49500.1	0.4087	NAC
APW78912.1	0.4647	European Cluster (EC)
APW78692.1	0.4615	EC
APW78681.1	0.4627	EC
APW78659.1	0.4647	EC
APW78868.1	0.4615	EC
AGN92849.1	0.4189	Asian Cluster (AC)
AGN92838.1	0.4189	AC
AOD40569.1	0.4164	African Cluster (AFC)
AOD41194.1	0.4172	AFC
AOD41183.1	0.4172	AFC
AOD41018.1	0.4172	AFC
AOD40803.1	0.4172	AFC

BCL, HTL, and CTL epitopes prediction

The present study found 23 B-cell linear epitopes from various sequence positions of 17 M2-1 proteins from four different clusters (Table 2). Among them,

Table 2: Antigenic properties of 23 linear B-cell epitopes of RSV M2-1 protein from four different clusters

Rank	Sequence	Start position	Score	Cluster
1	YFEWPPHALLVRQNFMLNKI	27	0.88	NAC, EC
2	VYNTVISYIESNRKNNKQTI	127	0.86	NAC, EC, AC
3	KNNKQTIHLLKRLPADVLKK	140	0.85	NAC, EC
4	SYIGSINNITKQSACVAMSK	82	0.84	NAC, EC, AC, AFC
5	CKFEIRGHCLNGRRCHYSHN	7	0.84	NAC, EC, AC, AFC
6	SACVAMSKLLIEINSDDIKK	94	0.83	NAC, AC
7	LIEINSDDIKKLRDNEEPNS	103	0.83	NAC, EC, AC, AFC
8	KTIKNTLDIHKSIIISNPKE	159	0.82	NAC, EC, AC, AFC
9	NGRRCHYSHNYFEWPPHALL	17	0.81	NAC, EC
10	PHALLVRQNFMLNKILKSMD	32	0.88	NAC, AC, AFC
11	RNPCKYEIRGHCLNGKKCHF	4	0.84	NAC
12	VISYIDSNKRNPKQTIHLLK	131	0.82	NAC
13	LGSVNNITKQSACVAMSKLL	84	0.8	NAC
14	KYSHKYWEWPLKTLMLRQNY	21	0.86	NAC
15	IACGSLITVLQNLDVGLVIQ	93	0.82	NAC
16	FDAPQRTAEYALGTIGVLKS	62	0.82	NAC
17	NTDAMSDVSGFDAPQRTAEY	52	0.82	NAC
18	GVLCNLIQSVISIEEKINSS	153	0.82	NAC
19	ESNRKNNKQTIHLLKRLPAD	130	0.84	AFC
20	PNSPKVRVYNTVISYIESNR	114	0.84	AFC
21	LLKRLPADVLKKTIKNTLDI	142	0.81	AFC
22	VRVYNTVISYIESNRKNNKQ	125	0.84	AFC
23	NRKNNKQTIHLLKRLPADVL	138	0.83	AC

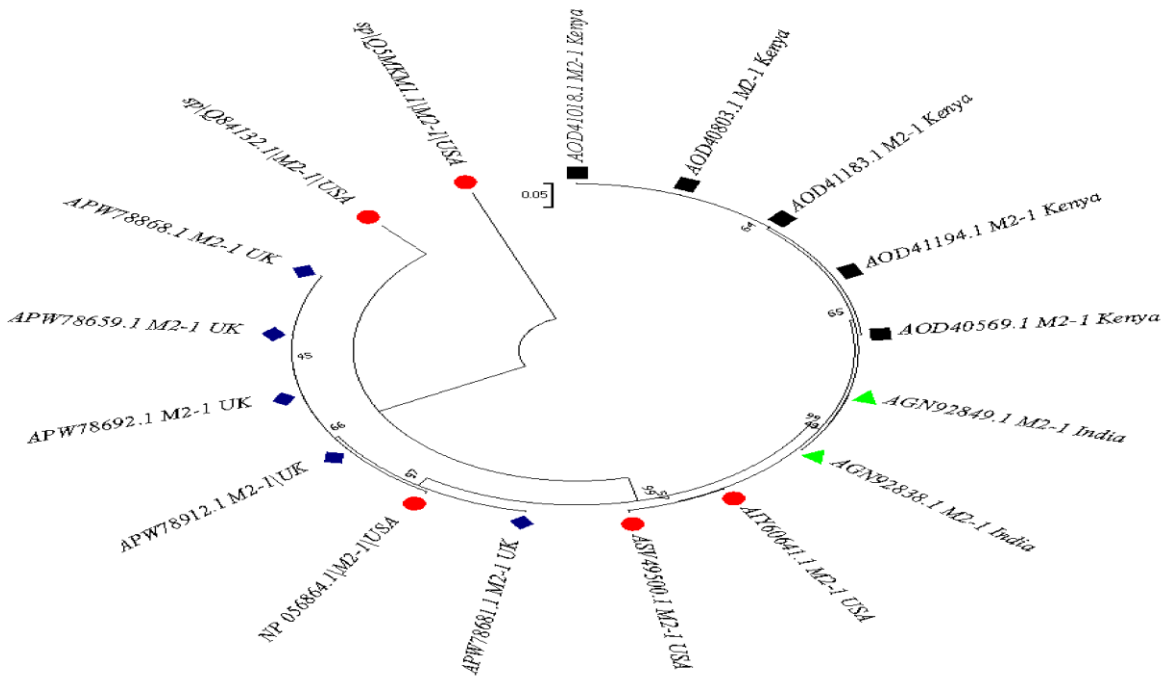

Figure 1: Phylogenetic clustering of 17 HRSV M2-1 Proteins from four different continents. African, Asian, North America, and European clusters are symbolized in black, green, red, and blue colours. The percentages of the replicate in trees in which the associated taxa clustered together by the bootstrap test (1000 replicates) are shown next to the branches. The tree was constructed using neighbour-joining method in mega7.

four were common in all four clusters (AC, AFC, EC, and NAC). Out of the four, epitope "SYIGSINNITKQSACVAMSK" and "CKFEIRGHCLNGRRCHYSHN" had higher antigenicity value of 0.84 compared to other two in the clusters. Among these two, the threshold of antigenicity for "CKFEIRGHCLNGRRCHYSHN" (1.039, only one residue had a score of below 1) was noticeably higher compared to "SYIGSINNITKQSACVAMSK" (1.028, seven residues had a score of below 1).

We found ten potential HTL epitopes candidate after the screening of a diverse set of alleles (Table 3). However, the majority of the epitopes were from HLA-DRB1*01:01 allele and some them had sequence similarity with BCL linear epitopes. CTL peptides are the potential vaccine candidates for the control of diseases. After screening of three main super type class, we found one common epitope "LLVRQNFML" in A2 with higher antigenicity (0.7122) and conservancy (1.03).

Table 3: HTL epitopes among matrix protein of hRSV

ID	Epitope	Allele	Method
74635	YLEKESIYY	HLA-DRB1*01:01	Consensus
28122	IPYSGLLLV	HLA-DRB3*01:01	Consensus
95757	RFAIKPME	HLA-DRB1*11:01	Consensus
121876	SRSALLAQM	HLA-DRB4*01:01	Consensus
122004	VRNKCLNGRR	HLA-DRB1*01:01	Consensus
144866	AITNAKII	HLA-DPA1*02:01/ DPB1*01:01	Consensus
542935	ILVKQISTPKGPS	HLA-DPA1*02:01	Consensus
546171	VNILVKQISTPKGPS	HLA-DRB4*01:01	Consensus
546172	VNILVKQISTPKGPSL	HLA-DRB5*01:01	Consensus
546173	VNILVKQISTPKGPSLR	HLA-DRB1*01:01	Consensus

Two epitopes "CLNGRRCHY" and "QSACVAMSK" from A3 super type were common in 15 out of 17 M2-1 proteins from four different clusters with antigenicity and conservancy score of 1.2599, 1.4418 and 0.8527, 1.3972, respectively (Table 4)

Table 4: Antigenic properties of major T-cell epitopes of RSV M2-1 protein from four different clusters at 0.80 conservancy

Accession number	T-cell antigen	Super type	VaxiJen Score	Conservancy (<-E)
NP056864.1	LLVRQNFML	A2 super type	0.7122	1.0329
	CLNGRRCHY	A3 super type	1.2599	0.8527
	QSACVAMSK		1.4418	1.3972
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701
	KSIIISNPK		0.5318	1.1139
Q84132.1	LLVRQNFML	A2 super type	0.7122	1.0329
	SMDRSNDTL		0.428	0.8539
	QSACVAMSK	A3 super type	1.4418	1.4542
	KQTIHLLKR		0.4361	0.8558
Q5MKM1.1	LMLRQNYML	A2 super type	0.6751	1.1108
	VLQNLDVGL		2.1681	0.9847
	KTLMLRQNY	A3 super type	1.025	0.9069
	ALGTIGVLK		0.8243	1.3255
	RQIIHILKR		0.5421	0.8231
	LPVGVLCNL	B7 super type	0.4009	0.8792
AIY60641.1	LLVRQNFML	A2 super type	0.7122	1.03366
	QSACVAMSK	A3 super type	1.4418	1.4532
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701
ASV49500.1	LLVRQNFML	A2 super type	0.7122	1.0366
	QSACVAMSK	A3 super type	1.4418	1.4532
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701
APW78912.1	LLVRQNFML	A2 super type	0.7122	1.0329
	CLNGRRCHY	A3 super type	1.2599	0.8527
	QSACVAMSK		1.4418	1.3972
	ISYIESNRK		0.6024	1.282
	KQTIHLLKR		0.4361	0.8657
	KSITISNPK		0.8562	1.2084
APW78692.1	LLVRQNFML	A2 super type	0.7122	1.0329
	CLNGRRCHY	A3 super type	1.2599	0.8527
	QSACVAMSK		1.4418	1.3972
	ISYIESNRK		0.6024	1.282
	KQTIHLLKR		0.4361	0.8657
	KSITISNPK		0.8562	1.2091
APW78681.1	LLVRQNFML	A2 super type	0.7122	1.0329

	QSACVAMSK	A3 super type	1.4418	1.3972
	ISYIESNRK		0.6024	1.282
	KQTIHLLKR		0.4361	0.8652
	KSITISNPK		0.8562	1.2091
APW78659.1	LLVRQNFML	A2 super type	0.7122	1.0329
	CLNGRRCHY	A3 super type	1.2599	0.8527
	QSACVAMSK		1.4418	1.3972
	ISYIESNRK		0.6024	1.282
	KQTIHLLKR		0.4361	0.8652
	KSITISNPK		0.8562	1.2084
APW78868.1	LLVRQNFML	A2 super type	0.7122	1.0329
	CLNGRRCHY	A3 super type	1.2599	0.8527
	QSACVAMSK		1.4418	1.3972
	ISYIESNRK		0.6024	1.282
	KQTIHLLKR		0.4361	0.8657
	KSITISNPK		0.8562	1.2091
AGN92849.1	LLVRQNFML	A2 super type	0.7122	1.0366
	QSACVAMSK	A3 super type	1.4418	1.4532
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701
AGN92838.1	LLVRQNFML	A2 super type	0.7122	1.0366
	QSACVAMSK	A3 super type	1.4418	1.4532
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701
AOD40569.1	LLVRQNFML	A2 super type	0.7122	1.0366
	QSACVAMSK	A3 super type	1.4418	1.4532
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701
AOD41194.1	LLVRQNFML	A2 super type	0.7122	1.0366
	QSACVAMSK	A3 super type	1.4418	1.4532
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701
AOD41183.1	LLVRQNFML	A2 super type	0.7122	1.0366
	QSACVAMSK	A3 super type	1.4418	1.4532
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701
AOD41018.1	LLVRQNFML	A2 super type	0.7122	1.0366
	QSACVAMSK	A3 super type	1.4418	1.4532
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701
AOD40803.1	LLVRQNFML	A2 super type	0.7122	1.0366
	QSACVAMSK	A3 super type	1.4418	1.4532
	ISYIESNRK		0.6024	1.278
	KQTIHLLKR		0.4361	0.8701

Epitope "LLVRQNFML" failed to produce a satisfactory binding score with MHC molecule during scoring from the T-epitope designer website while positive scores recorded for "CLNGRRCHY" and "QSACVAMSK" epitopes against major HLA types.

Allergenicity of the predicted epitopes

The predicted peptide should be non-allergen in order to be considered as the safe vaccine. Among the final two selected epitopes, "QSACVAMSK" classified as "probable allergen" by the AllerTop database while "CLNGRRCHY" categorized as "probable non-allergen". Therefore, epitope "QSACVAMSK" was excluded for further studies.

Molecular docking of the peptide to MHC molecules

The structure of the predicted CTL epitope, "CLNGRRCHY" in PEP-FOLD3 with residues is shown in Figure 2. Molecular docking image of "CLNGRRCHY" to human MHC class I antigen, HLA-A2 (PDB code 1HLA) and MHC class II antigen, HLA-DRB1*04:01 (PDB code 5NIG) are presented in Figure 3. The designed peptide produced strong and stable binding with both MHC class I and II regarding hydrogen bonds. In the first case (peptide to 1HLA), the designed epitope formed seven hydrogen bonds, six (CYS, ARG, TYR, HIS) of which had a bond resolution of more than 3 °A. While in second interaction (peptide to 1ZHL), peptide-protein interaction aided by six hydrogen bonds, four (CYS, GLY, ARG) of which were more robust and stable (>3 °A).

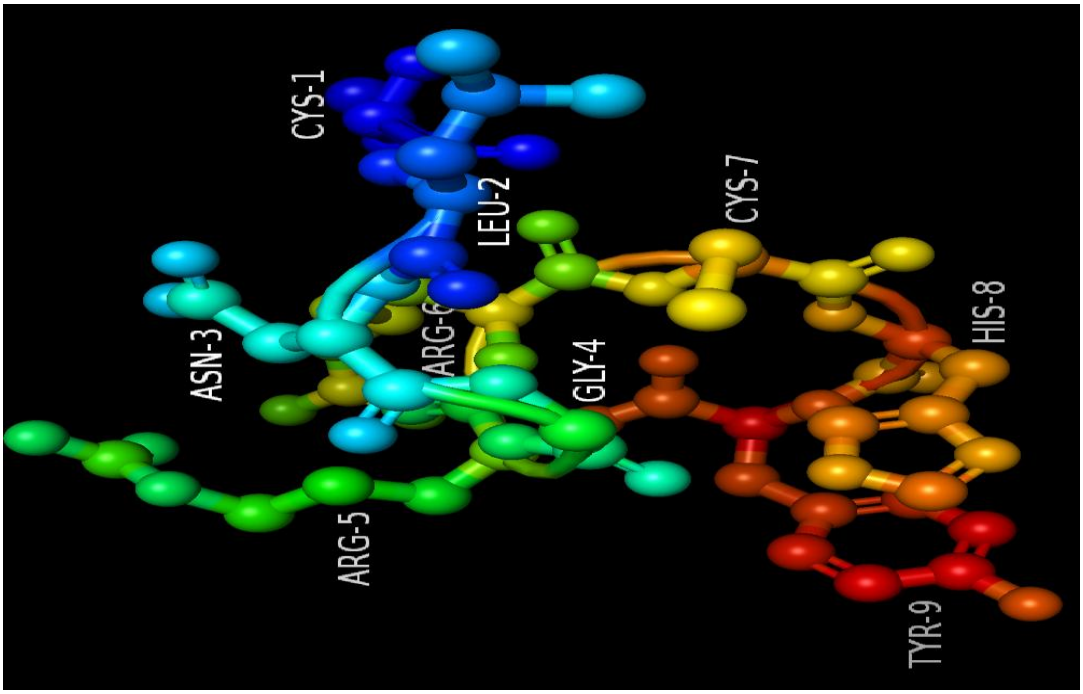

Figure 2: Modelled 3D structure of predicted "CLNGRRCHY" epitope using pep-fold3

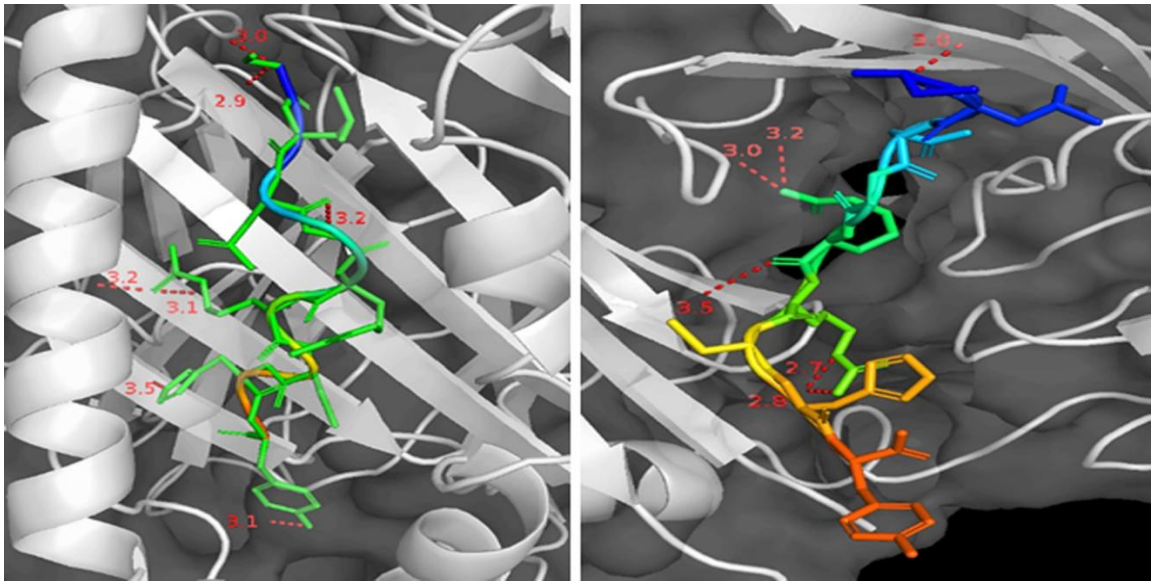

Figure 3: Molecular docking of the predicted epitope to HLA-A*0201, PDB-1HLA (left) and HLA-B*3508, PDB-1ZHL (right). The epitope "CLNGRRCHY" bound perfectly to the binding groove of MHC molecules in terms of strong hydrogen bonds (>3 °A).

Binding properties and population coverage of the predicted epitope

The binding properties of "CLNGRRCHY" revealed the percentile rank of epitope ranged from 0.4-0.75 in consensus method (ANN and SMM) and IC50 ranged from 135.87-553.88 for both ANN and SMM (Table 5). The deigned epitope covered 98.84% of world population with average hit of 2.57 and PC90 value of 1.04 by taking both class I and class II in counts where all major alleles (HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-B*07:02, HLA-B*08:01, HLA-C*01:02, HLA-C*02:02, HLA-C*05:01) were taken into consideration (Figure 4).

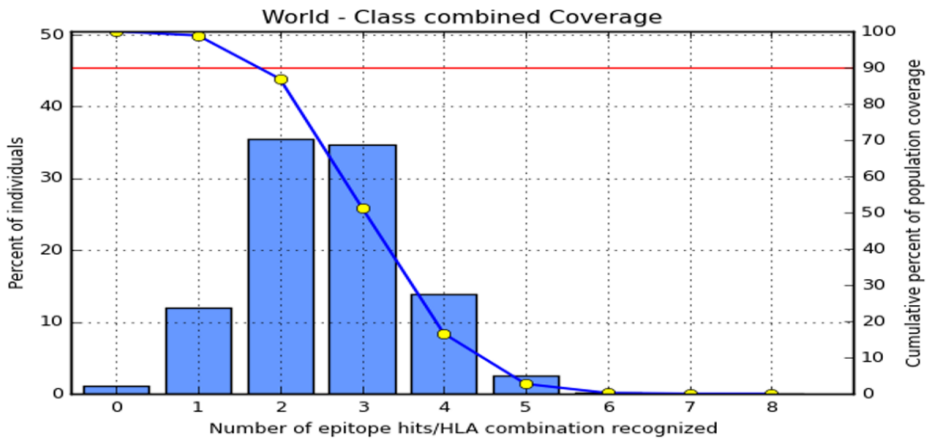

Figure 4: The world population coverage of predicted epitope based on MHC Restriction data. The cumulative coverage was 98.84% based on recommended ten HLA class. Individual bar represents number of hits for the corresponding HLA class generated by the predicted epitope

Discussion

In the present study, we aimed to identify the key antigenic epitope(s) with all necessary properties that can eventually bind strongly to MHC molecules to generate immunogenicity against hRSV. Several studies have been conducted to proffer solution to this, however no vaccine is yet to be licensed for hRSV (36). Transcription anti-termination factor M2-1 has been reported to be involved in viral replication and inactivation of M2-1 protein triggered inhibition of hRSV *in-vivo* (5). Hence, we selected this protein for *in-silico* drug design study.

The phylogenetic and evolutionary data suggested that the M2-1 protein from North America had more detachment and clustered everywhere. Therefore, we assumed that North America (NA) could be a potential source of hRSV transmission around the globe. However, these clusters had significantly higher bootstrap value (99%) and that means taxa's are well-supported by the reported data. We found more antigenicity in the M2-1 protein from the NA cluster, a well-studied sequence that can be a candidate vaccine. The four common BCL epitopes in 17 M2-1 protein sequences was one of the major findings in this study. A vaccine from one of these epitopes hereafter can inhibit hRSV from anywhere in the world. The sequence similarity of HTL and CTL epitopes with the BCL epitopes revealed more accuracies and wider coverage of the predicted peptides.

B-lymphocytes and hyper T-lymphocytes (HTLs) are the key players of adaptive immune response. They are the primary factors of the immune response mediated by activation of B-cells, cytotoxic T-cells and macrophages (9, 25). CTL peptides play a central role in governing adaptive immune response against broad range of infections (17). We comprehensively screened BCL, HTL, and CTL epitopes from M2-1 protein with consistency and flow strictly maintained. The two selected BCL epitopes were separately evaluated by a wide range of antigenic properties where "CKFEIRGHCLNGRRCHYSHN" showed its superiority over other.

Finally, among four CTL epitopes, only "CLNGRRCHY" qualified for all the necessary properties to be used as a candidate vaccine. Other three were excluded due to allergenicity, low affinity, and coverage. The conservancy values of selected BCL and CTL epitopes were more than 80%, a standard value to generate broad-spectrum antigenicity, and well-desired criteria for any predicted epitopes (17,37). A potential binder (epitope) should have three characteristics in order to generate strong immunogenicity, (i) an IC50 value of less than 500nM, (ii) percentile rank of below 1.5%, and (iii) population coverage of more than 80% (35,38). We found satisfactory IC50 value and percentile rank in our predicted epitope and population coverage of more than 98%, therefore, could be a potential immunogen against hRSV.

In molecular docking, epitope "CLNGRRCHY" formed strong and stable binding with MHC molecules. Usually, four hydrogen bonds are required to consider peptide as a strong binder of MHC molecules (16). Here we found seven (HLA-A2) and six hydrogen bonds (HLA-B*3508) in our protein-peptide dock for where nine formed a very strong and steady bond (> 3 °A) with both classes of HLA molecules. In both cases, arginine (R), cysteine (C), histidine (H), and tyrosine (Y) play a key role in protein binding. Arginine is strong protein binders which stabilize protein-protein interaction from aggregation, especially during the process of folding (39). These four amino acids have been used for the induction of protein binding ability of peptides in many vaccine preparation against broad range of diseases including cancer, pneumonia, and malaria (40–43). The introduction of cysteine zipper in the candidate vaccine leads to formation of inter-promoter disulfide rings that enabled stable coiled-coil trimers, generating satisfactory immunogenicity against hRSV (44).

Recent advances in bioinformatics have led to the rapid design of epitope-based vaccine against many human

pathogens. This method saves significant amount of time compared to previous lengthier vaccine design and long-lasting clinical trials. After analysing all the parameters, we hope that our designed epitope has all the necessary criteria as a future candidate vaccine that could produce protective neutralizing antibodies and cell-mediated immune responses to hRSV. Further synthesis and *in-vivo* laboratory trials are required to determine the exact potency of the designed epitope before commercial release.

Conflict of interest

The authors declare no conflict of interest

References

1. Falsey, A. R. Respiratory syncytial virus infection in adults. Crit Care. 2007; 1 (212): 171–181
2. Nair, H., Nokes, D. J., Gessner, B. D., et al. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. Lancet. 2010; 375 (9725): 1545–1555
3. Resch, B., Kurath, S., and Manzoni, P. Epidemiology of respiratory syncytial virus infection in preterm infants. Open Microbiol J. 2011; 5: 135–143
4. Empey, K. M., Peebles, R. S., and Kolls, J. K. Pharmacologic advances in the treatment and prevention of respiratory syncytial virus. Clin Infect Dis. 2010; 50(9): 1258–1267
5. Bailly, B., Richard, C. A., Sharma, G., et al. Targeting human respiratory syncytial virus transcription anti-termination factor M2-1 to inhibit in vivo viral replication. Sci Rep. 2016; 6: 1–11
6. Park, S., Baek, S., Lee, S., et al. Efficacy of oral ribavirin in hematologic disease patients with paramyxovirus infection: Analytic strategy using propensity scores. Antimicroial Agents Chemother. 2013; 57 (2):983–989
7. Tran, T-L., Castagne, N., Dubosclard, V., et al. The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner. J Virol. 2009; 83 (13): 6363–6374
8. Blondot, M. L., Dubosclard, V., Fix, J., et al. Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2-1 protein. PLoS Pathog. 2012; 8(5): e1002734
9. Shi, J., Zhang, J., Li, S., et al. Epitope-based vaccine target screening against highly pathogenic MERS-CoV: An in silico approach applied to emerging infectious diseases. PLoS One. 2015; 10(12): 1–16.
10. Muñoz-Medina, J. E., Sánchez-Vallejo, C. J., Méndez-Tenorio, A., et al. In silico identification of highly conserved epitopes of influenza A H1N1, H2N2, H3N2, and H5N1 with diagnostic and vaccination potential. Biomed Res Int. 2015;2015:ID813047.
11. Ali, M. T., Morshed, M. M., and Hassan, F. A computational approach for designing a universal epitope-based peptide vaccine against Nipah virus. Interdiscip Sci Comput Life Sci. 2015; 7(2): 177–185
12. Kori, P., Sajjan, S. S., and Madagi, S. B. In silico prediction of epitopes for Chikungunya viral strains. J Pharm Investig. 2015; 45(6): 579–591
13. Mirza, M. U., Rafique, S., Ali, A., et al. Towards peptide vaccines against Zika virus: Immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins. Sci Rep. 2017; 7: 44633.
14. Dash, R., Das, R., Junaid, M., Akash, M. F. C., Islam, A., and Hosen, S. M. Z. In silico-based vaccine design against Ebola virus glycoprotein. Adv Appl Bioinforma Chem. 2017; 10(1): 11–28
15. Ghosh, A., Chattopadhyay, S., Chawla-Sarkar, M., Nandy, P., and Nandy, A. In silico study of rotavirus VP7 surface accessible conserved regions for antiviral drug/vaccine design. PLoS One. 2012; 7 (7): e40749
16. Mahendran, R., Jeyabaskar, S., Sitharaman, G., Michael, R. D., and Paul, A. V. Computer-aided vaccine designing approach against fish pathogens *Edwardsiella tarda* and *Flavobacterium columnare* using bioinformatics softwares. Drug Des Devel Ther. 2016; 10: 1703–1714
17. Momtaz, F., Foysal, J., Rahman, M., and Fotedar, R. Design of epitope based vaccine against shrimp white spot syndrome virus (WSSV) by targeting the envelope proteins: An immunoinformatic approach. Turkish J Fish Aquat Sci. 2019; 19(2): 149–159
18. Cartee, T. L., and Wertz, G. W. Respiratory syncytial virus M2-1 protein requires phosphorylation for efficient function and binds viral RNA during infection. J Virol. 2001; 75 (24):12188–12197
19. Reimers, K., Buchholz, K., and Werchau, H. Respiratory syncytial virus M2-1 protein induces the activation of nuclear factor kappa B. Virology. 2005; 331 (2):260–268
20. Weber, M. W., Mulholland, E. K., and Greenwood, B. M. Respiratory syncytial virus infection in tropical and developing countries. Trop Med Int Heal. 1998; 3 (4): 268–280.
21. Bloom-feshbach, K., Alonso, W. J., Charu, V., Tamerius, J., Simonsen, L., and Miller, M. A. Latitudinal variations in seasonal

activity of influenza and respiratory syncytial virus (RSV): A global comparative review. PLoS One. 2013; 8 (2): 3–4
22. Kumar, S., Stecher, G., and Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016; 33 (7): 1870–1874
23. Doytchinova, I. A., and Flower, D. R. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007; 8(1): 4
24. Wang, P., Sidney, J., Dow, C., Mothé, B., Sette, A., and Peters, B. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput Biol. 2008; 4(4): e1000048
25. Pandey, R. K., Bhatt, T. K., and Prajapati, V. K. Novel Immunoinformatics Approaches to Design Multi-epitope Subunit Vaccine for Malaria by Investigating Anopheles Salivary Protein. Sci Rep. 2018; 8(1): 1–11.
26. Bui, H. H., Sidney, J., Dinh, K., Southwood, S., Newman, M. J., and Sette, A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinformatics. 2006; 7: 1–5
27. Bhasin, M., and Raghava, G. P. S. Prediction of CTL epitopes using QM, SVM and ANN techniques. Vaccine. 2004; 22 (23–24): 3195–3204
28. Haug, B. E., Camilio, K. A., Eliassen, L. T., et al. Discovery of a 9-mer Cationic Peptide (LTX-315) as a Potential First in Class Oncolytic Peptide. J Med Chem. 2016; 59 (7): 2918–2927
29. Lund, O., Nielsen, M., Kesmir, C., et al. Definition of supertypes for HLA molecules using clustering of specificity matrices. Immunogenetics. 2004; 55 (12):797–810.
30. Thévenet, P., Shen, Y., Maupetit, J., Guyon, F., Derreumaux, P., and Tufféry, P. PEP-FOLD: An updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res. 2012; 40 (W1): 288–293
31. Pandey, R. K., Kumbhar, B. V., Sundar, S., Kunwar, A., and Prajapati, V. K. Structure-based virtual screening, molecular docking, ADMET and molecular simulations to develop benzoxaborole analogs as potential inhibitor against Leishmania donovani trypanothione reductase. J Recept Signal Transduct. 2017; 37 (1): 60–70
32. Ciemny, M., Kurcinski, M., Kozak, K., Kolinski, A., and Kmiecik, S. Highly flexible protein-peptide docking using CABS-Dock. Methods Mol Biol. 2012; 857: 375–398
33. Kurcinski, M., Jamroz, M., Blaszczyk, M., Kolinski, A., and Kmiecik, S. CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. Nucleic Acids Res. 2015; 43 (W1): W419–424
34. Yuan, S., Chan, H. C. S., and Hu, Z. Using PyMOL as a platform for computational drug design. WIREs Comput Mol Sci. 2017; 7 (2): 1–10
35. Misra, N., Panda, P. K., Shah, K., Sukla, L. B., and Chaubey, P. Population coverage analysis of T-Cell epitopes of Neisseria meningitidis serogroup B from Iron acquisition proteins for vaccine design. Bioinformation. 2011; 6 (7): 255–261
36. Neuzil, K. M., Acosta, P. L., Caballero, M. T., Polack, F. P, and Immunol, C. V. Progress toward a respiratory syncytial virus vaccine. Clin Vaccine Immunol. 2016; 23 (3): 186–188
37. Sumo, U., Tambunan, F., Ruth, F., Sipahutar, P., Parikesit, A. A., and Kerami, D. Vaccine design for H5N1 based on B- and T-cell epitope predictions. Bioinform Biol Insights. 2016; 10: 27–35.
38. Fleri, W., Paul, S., Dhanda, S. K., et al. The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. Front Immunol. 2017; 8: 1–16.
39. Baynes, B. M., Wang, D. I. C., and Trout, B. L. Role of arginine in the stabilization of proteins against aggregation. Biochemistry. 2005; 44 (12): 4919–4925
40. Mattner, F., Fleitmann, J. K., Lingnau, K., et al. Vaccination with poly-L-arginine as immunostimulant for peptide vaccines: Induction of potent and long-lasting T-cell responses against cancer antigens. Cancer Res. 2002; 62(5): 1477–1480
41. Huebener, N., Fest, S., Strandsby, A., et al. A rationally designed tyrosine hydroxylase DNA vaccine induces specific antineuroblastoma immunity. Mol Cancer Ther. 2008; 7 (7): 2241–2251
42. Ochs, M. M., Williams, K., Sheung, A., et al. A bivalent pneumococcal histidine triad protein D-choline-binding protein A vaccine elicits functional antibodies that passively protect mice from *Streptococcus pneumoniae* challenge. Hum Vaccines Immunother. 2016; 12 (11): 2946–2952.
43. Jin, J., Hjerrild, K. A., Silk, S. E., et al. Accelerating the clinical development of protein-based vaccines for malaria by efficient purification using a four amino acid C-terminal "C-tag." Int J Parasitol. 2017; 47 (7): 435–446
44. Stewart-Jones, G. B. E., Thomas, P. V., Chen, M., et al. A cysteine zipper stabilizes a pre-fusion F glycoprotein vaccine for respiratory syncytial virus. PLoS One. 2015; 10 (6): 1–16.