ISSN: 2475-6296

Research Article

Journal of Clinical & Experimental Immunology

In Silico Prediction of Multi-Epitopes Vaccine from the Fusion F Protein Against Respiratory Syncytial Virus

Marwa Karrar Hussein[1], Yassir A Almofti[2]*, Khoubieb Ali Abd-elrahman[3], Mashair AA Nouri[2] and Elsideeq EM Eltilib[2]

[1]*Alneelain Medical Research institute, Faculty of Medicine, Alneelain University, Khartoum- Sudan*

[2]*Department of Molecular Biology and Bioinformatics, College of Veterinary Medicine, University of Bahri, Khartoum- Sudan*

[3]*Department of pharmaceutical technology, College of Pharmacy, University of Medical Science and Technology (MUST) Khartoum-Sudan*

***Corresponding author:**
Yassir A Almofti, University of Bahri, College of Veterinary Medicine, Department of Molecular Biology and Bioinformatics- Khartoum, Sudan

Submitted: 25 Feb 2020; **Accepted**: 03 Mar 2020; **Published**: 17 Mar 2020

Abstract

Respiratory Syncytial Virus (RSV) is the major cause of the lower respiratory tract illness (RTI) in the elderly and in immunocompromised patients and children under 5 years of age. The disease causes deaths of approximately 500 infants each year. Conventional vaccine against the disease demonstrated immunological pitfalls to enhance T-helper responses and developed non-neutralising antibodies. This study aimed to predict epitopes from the fusion F protein of SRV that elicit the immune system and acted as safer efficacious vaccine. A total of 199 strains of RSV were retrieved from the NCBI database. The immune epitope database analysis resources (IEDB) were used for epitopes prediction against B and T cells. The population coverage was also calculated for the proposed epitopes against the whole world. Only two epitopes (441-YVSNK-445 and 440-DYVS-443) successfully passed all B cell prediction tools and demonstrated higher score in Emini and Kolaskar and tongaonker software. Thus were proposed as B cells epitopes. For T cells, a total of 177 epitopes were found to interact with MHC-I alleles. Among them four epitopes (53-YTSVITIEL-61; 250-YMLTNSELL-258, 198-YIDKQLLPI-206, and 450-VSVGNTLYY-458) were proposed since they interacted with the highest number of alleles and the best binding affinity to MHC-1 alleles. Moreover, a total of 397 core epitopes were found to interact with MHC-II alleles. However, the best four core proposed epitopes that interacted with higher number of MHC-II alleles were 217-IETVIEFQQ-226; 250-YMLTNSELL-258; 477-FYDPLVFPS-485 and 505-FIRKSDELL-513. Strikingly the epitope 250-YMLTNSELL-258 successfully interacted with both MHC-1and MHC-II alleles. The population coverage was 48.61% and 99.64% for MHC-I and MHC-II epitopes, respectively, and 100% for all T cells epitopes. Taken together ten epitopes successfully proposed as vaccine candidate against RSV. In vivo and in vitro clinical trials studies are required to elucidate the effectiveness of these epitopes as vaccine.

Keywords: Respiratory Syncytial Virus; NCBI; IEDB; Insilico Vaccine; B cells; T cells

Introduction

Respiratory Syncytial Virus (RSV) is the major cause of the lower respiratory tract illness (RTI) in the elderly and in immunocompromised patients [1, 2]. The disease was first recognized in 1955 in Chimpanzees and it was registered effectively in humans in 1956 [3]. It was named respiratory syncytial virus due to its symptomatic form and ability to generate and stimulate syncytia multinucleate mass of protoplasm produced by merging of cells in tissue culture [4, 5]. The RSV belongs to Paramyoxoviridae family, which is non-segmented, negative sense, single-stranded RNA genome viruses [2]. RSV is a very common etiologic agent of childhood respiratory infections throughout the world [6]. Beside that, the RSV is a seasonal virus causing annual outbreaks in winter, raining season in a temperate climate and tropical climate sequentially [7-9]. Re-infection with RSV happens constantly throughout life although infants are unlikely to get frequent bronchiolitis [10].

The disease causes about 150,000 hospital admissions per year in children under 5 years of age and about18% of all emergency departments visits in children under 5 years of age [11, 12]. The disease causes deaths of approximately 500 infants each year [13]. The incubation time of the disease is 3-4 days and most of the children infected with RSV that characterized by upper respiratory tract symptoms such as rhinitis, coryza, cough, nasal congestion and may also develop otitis media often with bacterial super infection [14].

The RSV genome composed of 10 open reading frames (ORFs) that coded for eleven structural and nonstructural proteins [15]. As in all paramyxoviruses, the first seven genes were translated into seven structural proteins. The first three proteins were phosphoprotein (P),

RNA-dependent RNA polymerase (L) and nucleoprotein (N) [15]. These proteins were responsible for encapsulating the viral RNA to form a helical construction that described the ribonucleoprotein complex (RNP) [16]. The aim of this structure is to protect the RNA and formed the lowest replication machinery [17, 18]. The three integral membrane proteins of RSV were the receptor attachment glycoprotein (G) which is responsible for viral attachment to cells and enhanced cell-to-cell fusion [19, 20]. The fusion protein (F) is responsible of viral fusion [21]. The third protein is a short hydrophobic (SH) protein that formed a pentameric ion channel [21].

The RSV F protein in its inactivated form (F0) comprises 574 amino acids and has a trimeric coiled-coil structure, similar to other viral fusion proteins [19]. F1 and F2 were the active form of F0 and the activation take place by cleavage of F0 into two disulphide linked subunits [19]. The main function of F protein is encouraging both the fusion of the viral to cell membranes and facilitated the transfer of viral genetic material [22]. Moreover, the F fusion protein of the infected and adjacent cell membranes caused the formation of syncytia [22]. The syncytia are the characteristics of the RSV cytopathic effect, besides, syncytia are necessary for cell-to-cell viral transmission. The interaction between the F protein and a small GTPase, RhoA, facilitates RSV-induced syncytium formation [23].

The importance of vaccination to combat the RSV infection could extend beyond the prevention of hospital admission of infants to long-term respiratory health. In the 1960s, the first vaccine was used. It was a formalin-inactivated whole virus but resulted in RSV enhanced disease with 80% hospitalization and two deaths [24]. Live-attenuated vaccine was shown to induce maximum immunogenicity. However, immunological pitfalls to enhance T-helper responses and the development of non-neutralising antibodies were reported, as induced by formalin-inactivated RSV and mimic exposure to wild-type virus [25]. As an alternative to live and attenuated vaccine to prevent the RSV infection, subunit vaccines were used because they were safer with no chance of reversion of the virus to wild type. However, they offer little immunogenicity in young children [26]. Accordingly, safety of the vaccine is a serious concern. Also, incomplete inactivation of the virus may cause outbreaks among the vaccinated populations. Furthermore, some viral proteins may induce harmful immune or inflammatory responses, even causing SRV like diseases [27, 28]. Thus, the need for a safer and efficacious vaccine without future complications is highly recommended. Therefore, this study aimed to use the immunoinformatics approaches found in the Immune Epitope Database (IEDB) to predict epitopes from the fusion F protein of SRV that elicit the immune system and acted as safer efficacious vaccine.

Material and Methods

Protein sequences retrieval and alignment tool

The protein sequences of Fusion protein (F) from 199 strains were retrieved from the NCBI database (https://www.ncbi.nlm.nih.gov/protein/NP_044596.1). The protein strains were retrieved according to their accession numbers, country and date of collection. All the retrieved strains were shown in table [1].

Determination of conserved regions

Multiple sequence alignment was used to obtain conserved regions in the retrieved sequences using Clustal-W as applied in the Bio-Edit program (version 7.2.5.0) [29-33].

Epitopes prediction

The conserved regions of the candidate epitopes were analyzed by different prediction software tools obtained by Immune Epitope Database (IEDB) analysis (https://www.iedb.org/). The reference sequence was used as an input in the IEDB software analysis.

B-lymphocytes epitopes prediction

Tools from IEDB at (http://tools.iedb.org/bcell/) were used to identify the B cell epitopes, including Bepipred for linear epitope analysis, Emini for surface accessibility and Kolaskar and Tongaonkar for antigenicity scale.

Table 1: The 199 retrieved strains with their accession numbers, countries and year of collection

Accession	Country	Year	Accession	Country	Year
AHW81420	USA: Tennessee	2013	AHE57683	USA: Tennessee	2002
AHW81410	USA: Tennessee	2013	AIA63634	USA: Tennessee	2013
AHW81400	USA: Tennessee	2013	AIA63624	USA: Tennessee	2013
AHW81390	USA: Tennessee	2003	AIA63614	USA: Tennessee	2013
AHW81370	USA: Tennessee	2003	AIA63604	USA: Tennessee	2013
AHW81360	USA: Tennessee	2003	AIA63594	USA: Tennessee	2013
AHW81350	USA: Tennessee	2003	AIA63584	USA: Tennessee	2013
AHW81330	USA: Tennessee	2003	AIA63564	USA: Tennessee	2013
AHW81320	USA: Tennessee	2003	AIA63554	USA: Tennessee	2013
AHW81310	USA: Tennessee	2003	AIA63544	USA: Tennessee	2013
AHW81290	USA: Tennessee	2004	AIA63534	USA: Tennessee	2013
AHW81280	USA: Tennessee	2004	AIA63514	USA: Tennessee	2013
AHW81270	USA: Tennessee	2004	AIA63504	USA: Tennessee	2013
AHW81260	USA: Tennessee	2004	AIA63494	USA: Tennessee	2013
AHW81250	USA: Tennessee	2004	AIA63484	USA: Tennessee	2013
AHW81240	USA: Tennessee	2004	AIA63474	USA: Tennessee	2013

AHW81230	USA: Tennessee	2004	AIA63464	USA: Tennessee	2013
AHW81220	USA: Tennessee	2003	AIA63454	USA: Tennessee	2013
AHW81210	USA: Tennessee	2001	AIA63444	USA: Tennessee	2013
AHW81200	USA: Tennessee	2003	AIA63434	USA: Tennessee	2013
AHW81190	USA: Tennessee	2003	AIA63424	USA: Tennessee	2013
AHW81180	USA: Tennessee	2003	AIA63414	USA: Tennessee	2013
AHW81160	USA: Tennessee	2001	AIA63404	USA: Tennessee	2013
AHW81150	USA: Tennessee	2001	AIA63394	USA: Tennessee	2013
AHW81140	USA: Tennessee	2001	AIA63384	USA: Tennessee	2013
AHW81130	USA: Tennessee	2001	AIA63374	USA: Tennessee	2013
AHW81120	USA: Tennessee	2001	AIA63364	USA: Tennessee	2013
AHW81110	USA: Tennessee	2001	AIA63354	USA: Tennessee	2013
AHW81100	USA: Tennessee	2001	AIA63344	USA: Tennessee	2013
AHW81090	USA: Tennessee	2001	AIA63334	USA: Tennessee	2013
AHW81080	USA: Tennessee	2001	AIA63324	USA: Tennessee	2013
AHW81070	USA: Tennessee	2001	AIA63314	USA: Tennessee	2013
AHW81060	USA: Tennessee	2001	AIA63304	USA: Tennessee	2013
AHW81050	USA: Tennessee	2001	AIA63294	USA: Tennessee	2013
AHW81030	USA: Tennessee	2001	AIA63284	USA: Tennessee	2013
AHW81020	USA: Tennessee	2001	AIA63274	USA: Tennessee	2013
AHW81000	USA: Tennessee	2001	AIA63264	USA: Tennessee	2013
AHW80990	USA: Tennessee	2001	AIA63254	USA: Tennessee	2013
AHW80970	USA: Tennessee	2001	AIA63244	USA: Tennessee	2013
AHW80940	USA: Tennessee	2001	AIA63234	USA: Tennessee	2013
AHW80930	USA: Tennessee	2001	AIA63223	USA: Tennessee	2013
AHW80920	USA: Tennessee	2001	AIA63214	USA: Tennessee	2013
AHW80910	USA: Tennessee	2001	AIA63204	USA: Tennessee	2013
AHW80900	USA: Tennessee	2001	AIA63194	USA: Tennessee	2013
AHW80890	USA: Tennessee	2001	AIA63184	USA: Tennessee	2013
AHW80880	USA: Tennessee	2001	AIA63174	USA: Tennessee	2013
AHW80870	USA: Tennessee	2001	AIA63164	USA: Tennessee	2013
AHW80860	USA: Tennessee	2001	AIA63154	USA: Tennessee	2013
AHW80850	USA: Tennessee	2001	AIA63134	USA: Tennessee	2013
AHW80840	USA: Tennessee	2001	AIA63124	USA: Tennessee	2013
AHW80830	USA: Tennessee	2001	AIA63114	USA: Tennessee	2013
AHW80820	USA: Tennessee	2001	AIA63104	USA: Tennessee	2013
AHW80800	USA: Tennessee	2001	AIA63094	USA: Tennessee	2013
AHW80790	USA: Tennessee	2001	AIA63084	USA: Tennessee	2013
AHW80780	USA: Tennessee	2001	AIA63074	USA: Tennessee	2013
AHW80770	USA: Tennessee	2001	AIA63064	USA: Tennessee	2013
AHW80760	USA: Tennessee	2001	AIA63054	USA: Tennessee	2013
AHW80750	USA: Tennessee	2003	AIA63044	USA: Tennessee	2013
AHW80740	USA: Tennessee	2003	AIA63034	USA: Tennessee	2013
AHW80730	USA: Tennessee	2003	AIA63024	USA: Tennessee	2013
AHW80720	USA: Tennessee	2003	AIA63014	USA: Tennessee	2013
AHW80710	USA: Tennessee	2003	AIA63004	USA: Tennessee	2013

AHW80700	USA: Tennessee	2003	AIA62994	USA: Tennessee	2013
AHW80680	USA: Tennessee	2003	AIA62984	USA: Tennessee	2013
AHW80670	USA: Tennessee	2003	AIA62974	USA: Tennessee	2013
AHW80660	USA: Tennessee	2003	AIA62955	USA: Tennessee	2013
AHW80630	USA: Tennessee	2003	AIA62945	USA: Tennessee	2013
AHW80620	USA: Tennessee	2003	AIA62935	USA: Tennessee	2013
AHW80610	USA: Tennessee	2003	AIA62925	USA: Tennessee	2013
AHW80600	USA: Tennessee	2003	AIA62905	USA: Tennessee	2013
AHW80580	USA: Tennessee	2002	AIA62895	USA: Tennessee	2013
AHW80560	USA: Tennessee	2002	AIA62885	USA: Tennessee	2013
AHW80531	USA: Tennessee	2001	AIA62875	USA: Tennessee	2013
AHW80521	USA: Tennessee	2001	AIA62865	USA: Tennessee	2013
AHW80511	USA: Tennessee	2001	AIA62855	USA: Tennessee	2013
AHE57894	USA: Tennessee	2002	AIA62835	USA: Tennessee	2013
AHE57884	USA: Tennessee	2002	AIA62825	USA: Tennessee	2013
AHE57874	USA: Tennessee	2002	AIA62815	USA: Tennessee	2013
AHE57854	USA: Tennessee	2002	AIA62805	USA: Tennessee	2013
AHE57844	USA: Tennessee	2002	AIA62795	USA: Tennessee	2013
AHE57834	USA: Tennessee	2002	AIA62785	USA: Tennessee	2013
AHE57824	USA: Tennessee	2002	AIA62775	USA: Tennessee	2013
AHE57814	USA: Tennessee	2002	AIA62765	USA: Tennessee	2013
AHE57804	USA: Tennessee	2002	AIA62755	USA: Tennessee	2013
AHE57794	USA: Tennessee	2002	AIA62745	USA: Tennessee	2013
AHE57784	USA: Tennessee	2002	AIA62735	USA: Tennessee	2013
AHE57774	USA: Tennessee	2002	AIA62725	USA: Tennessee	2013
AHE57764	USA: Tennessee	2002	AIA62715	USA: Tennessee	2013
AHE57754	USA: Tennessee	2002	AIA62705	USA: Tennessee	2013
AHE57744	USA: Tennessee	2002	AIA62695	USA: Tennessee	2013
AHE57734	USA: Tennessee	2002	AIA62685	USA: Tennessee	2013
AHE57724	USA: Tennessee	2002	AIA62675	USA: Tennessee	2013
AHE57714	USA: Tennessee	2002	AIA62665	USA: Tennessee	2013
AHE57704	USA: Tennessee	2002	AIA62655	USA: Tennessee	2013
AHE57694	USA: Tennessee	2002	AIA62645	USA: Tennessee	2013

T-lymphocytes epitopes prediction

The IEDB tools were used for the identification of the T cell epitopes prediction. The prediction method includes the major histocompatibility complex class I and Π (MHC-I, MHC-Π).

MHC class I binding predictions

Analysis of epitopes binding to MHC-I molecules was assessed by the software of IEDB MHC-I prediction tools (http://tools.iedb.org/mhci/). The prediction method was obtained by Artificial Neural Network (ANN), Stabilized Matrix Method (SMM) or Scoring Matrices derived from combinatorial peptide libraries. Before the prediction step, epitopes length was set as 9mers. The conserved epitopes that bind to alleles at score equal to or less than 300 was considered as half-maximal inhibitory concentrations. Epitopes equal to or less than the IC50 were selected for further analysis

MHC-Π binding predictions

Analysis of epitopes binding to MHC-II molecules was achieved by the IEDB MHC-II prediction tools (http://tools.iedb.org/mhcii/). The neural networks align (NN-align) that allow for simultaneous identification of the MHC-II binding core epitopes and binding affinity was used. All the predicted conserved epitopes that bind to many alleles at score equal to or less than 3000 half-maximal inhibitory concentration (IC50 $\leq$ 3000) was considered. Epitopes that equal to or less than the IC50 were selected for further analysis.

Population coverage

For the calculation of the population coverage for all potential MHC-I and II epitopes bindings, the IEDB tools were used (http://tools.iedb.org/population/). The reference sequence of the virus was assessed for population coverage against the whole world with selected MHC-I and MHC-II interacted alleles.

Homology modeling

Raptor X protein structure prediction server was used for creation the 3D structure of the reference protein of the virus (http://raptorx.uchicago.edu/StructurePrediction/predict/). The reference sequence was used as an input and Chimera 1.8 was used as a tool (https://www.cgl.ucsf.edu/chimera/) to visualize the selected epitopes belonging to B cell and T cell (MHC-I and MHC-II). Homology modeling was used for visualization of the surface accessibility of the B-lymphocytes predicted candidate epitopes as well as for visualization of all predicted T cell epitopes in the structural level.

Results

Sequences alignment

Sequence alignment of all retrieved strains was performed using Clustal W that presented by Bioedit software. The retrieved strains, their accession numbers, country and date of collection were shown in (table 1). The retrieved sequences demonstrated conservancy when sequences were aligned. As shown in (figure 1) the conserved regions were recognized by identity of amino acid sequences among the retrieved sequences.

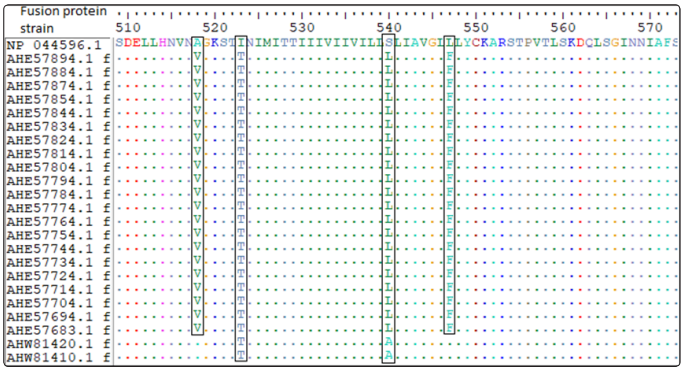

Figure 1: Multiple sequence alignment showed the conservancy between sequences of the retrieved strains of F fusion protein. The alignment was performed using BioEdit software tool. Dots showed the conserved regions while rectangular within the sequences showed the mutated region between strains

B cell Epitopes Prediction

The Fusion (F) protein was subjected to Bepipred linear epitope prediction, Emini surface accessibility and Kolaskar and Tongaonkar antigenicity prediction methods from IEDB. The thresholds of Bepipred linear epitope, Emini surface accessibility and Kolaskar and Tongaonkar antigenicity were shown in Figure 2. In Bepipred linear epitope prediction method; the average score of Fusion protein to B-lymphocytes was -0.212 (minimum: -0.001 and maximum: 1.580). Values equal to or greater than the default threshold 0.350 were predicted as conserved linear epitope. Table 2 showed that 12 epitopes were predicted by Bepipred method as a linear epitopes. In Emini surface accessibility prediction, the average score of Fusion F protein was 1.000 (minimum: 0.030 and maximum: 14.390). Values equal to or greater than the default threshold 1.000 were regarded potentially on the surface. Emini surface accessibility method predicted 7 epitopes on the surface that have potential binding to B-lymphocytes cells. In Kolaskar and Tongaonkar antigenicity prediction method; the average score of Fusion F protein was 1.041 (minimum: -0.836 and maximum: 1.246). Values equal to or greater than the default threshold 1.041 were considered as antigenic epitopes. This method predicted 5 antigenic epitopes with potential binding to B-lymphocytes cells. Accordingly, two conserved epitopes were successfully predicted to elicit the B cell lymphocytes since they were conserved among all retrieved strains, got higher score values in Emini surface accessibility and Kolaskar and Tongaonkar antigenicity prediction methods. These two 2 epitopes were 441-YVSNK-445 and 440-DYVS-443. The three dimension structural (3D) level of these epitopes was shown in Figure 3.

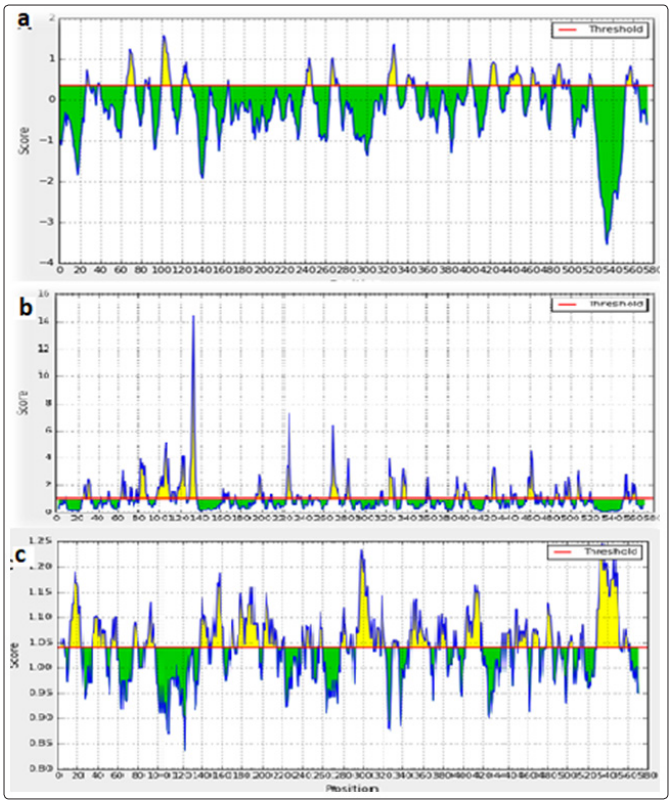

Figure 2: Prediction of B-cell epitopes by different IEDB scales (A- Bepipred linear epitope prediction (0.350), B- Emini surface accessibility (1.000), C- Kolaskar and Tongaonkar antigenicity prediction(1.041). Regions above threshold (red line) are proposed to be a part of B cell epitope while regions below the threshold (red line) are not

Table 2: B-cell epitopes prediction, the position of peptides is according to the position of amino acids in the fusioF protein of RSV virus

Peptide	Start	End	Length	Emini a	Kolaskar b
DYVSNKGVDTVS	440	451	12	0.91	1.046
#YVSNK	441	445	5	1.522	1.052
#DYVS	440	443	4	1.004	1.106
ENKCNGTDA	66	74	9	1.434	0.94
GVTTPV	242	247	6	0.42	1.087
KCTASNKN	421	428	8	1.685	0.976
LCTTNTKEG	321	329	9	0.908	0.98
RGWYCD	339	344	6	0.685	1.013
SDEFDA	485	490	6	1.356	0.958
TDVS	400	403	4	0.924	1.042
KQEGKS	461	466	6	3.809	0.935
RARR	106	109	4	2.927	0.921

#the proposed epitopes that pass all the criteria for B cells

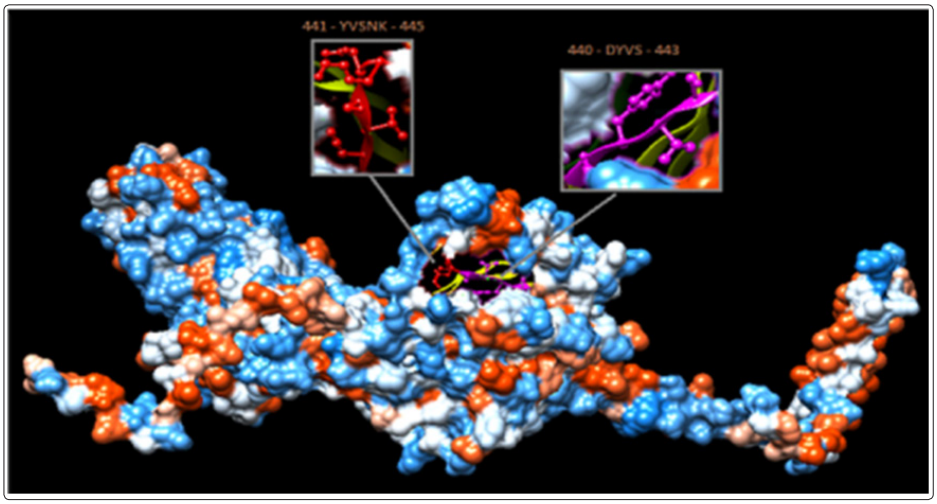

Figure 3: Position of proposed conserved B cell epitopes in structural level of fusion F protein of RSV. Two epitopes were predicted to interact with B cell. The epitopes showed conservancy, surface accessibility and antigenicity using IEDB software

T Lymphocytes Epitopes Binding Prediction

MHC-I Binding Predictions

RSV Fusion F protein was analyzed using IEDB MHC-1 binding prediction tool to predict T lymphocytes epitopes that have binding affinity with MHC-I alleles based on Artificial Neural Network (ANN) with half-maximal inhibitory concentration (IC50) ≤ 300. As shown in table 3, a total of 177 epitopes were found to interact with MHC-I alleles. The epitopes 53-YTSVITIEL-61, 198-YIDKQLLPI-206, 250-YMLTNSELL-258 and 450-VSVGNTLYY-458 interacted with the highest number of alleles and the best binding affinity to MHC-1 alleles. The three dimensional structural level (3D) of these epitopes within Fusion F protein of RSV was shown in figure 4.

Table 3: List of epitopes that had binding affinity with MHC-I alleles. The position of peptides is according to position of amino acids in fusion F protein of RSV virus

Peptide	Start	End	Allele	ic50	Percentile
AFIRKSDEL	504	512	HLA-C*03:03	268.24	0.54
ALRTGWYTS	47	55	HLA-A*30:01	242.72	0.58
ASISQVNEK	490	498	HLA-A*11:01	25.31	0.15
			HLA-A*30:01	154.51	0.46
AVVSLSNGV	177	185	HLA-A*02:06	14.44	0.19
			HLA-A*68:02	28.9	0.26
AYVVQLPLY	298	306	HLA-A*29:02	44.17	0.29
			HLA-A*30:02	110.12	0.4
CNIDIFNPK	382	390	HLA-A*68:01	55.61	0.49
CSISNIETV	212	220	HLA-A*02:06	275.77	1.9
			HLA-A*68:02	56.51	0.43
			HLA-C*12:03	180.39	0.27
CWKLHTSPL	313	321	HLA-B*08:01	138.32	0.32
			HLA-C*14:02	111.52	0.2
DEFDASISQ	486	494	HLA-B*18:01	12.67	0.04
DTVSVGNTL	448	456	HLA-A*68:02	72.09	0.51
EGSNICLTR	328	336	HLA-A*68:01	90.71	0.7
ETCKVQSNR	356	364	HLA-A*68:01	8.38	0.05
ETVIEFQQK	218	226	HLA-A*68:01	50.15	0.44
EVLAYVVQL	295	303	HLA-A*68:02	103.07	0.63
EVNLCNIDI	378	386	HLA-A*68:02	158.97	0.8
FCDTMNSLT	366	374	HLA-C*05:01	13.38	0.02
FPSDEFDAS	483	491	HLA-B*35:01	17.75	0.08
FQQKNNRLL	223	231	HLA-B*39:01	217.86	0.19
			HLA-C*06:02	104.82	0.04
			HLA-C*07:01	194.22	0.06
FSVNAGVTT	237	245	HLA-C*12:03	242.75	0.33
GAIVSCYGK	411	419	HLA-A*11:01	136.15	0.86
			HLA-A*68:01	115.73	0.81
GVTTPVSTY	242	250	HLA-A*30:02	244.45	0.9
IASGIAVSK	148	156	HLA-A*11:01	65.09	0.45
			HLA-A*68:01	25.6	0.21
IIKEEVLAY	291	299	HLA-B*15:01	51.73	0.3
IINFYDPLV	474	482	HLA-A*02:01	44.48	0.48
			HLA-A*02:06	264.63	1.8
IKEEVLAYV	292	300	HLA-A*02:06	298.4	1.9
IMITTIIIV	525	533	HLA-A*02:01	13.6	0.13
			HLA-A*02:06	48.47	0.53
IMSIIKEEV	288	296	HLA-A*02:01	130.38	1.3
ITREFSVNA	233	241	HLA-A*30:01	15.65	0.08
IVNKQSCSI	206	214	HLA-A*32:01	153.74	0.2
IVRQQSYSI	280	288	HLA-A*30:01	59.8	0.25
			HLA-A*32:01	124.58	0.17
			HLA-B*07:02	531.11	1.4

IVSCYGKTK	413	421	HLA-A*11:01	175.57	1.2
KEEVLAYVV	293	301	HLA-B*40:01	73.25	0.19
			HLA-B*40:02	20.59	0.03
KGEPIINFY	470	478	HLA-A*30:02	72.3	0.25
KINQSLAFI	498	506	HLA-A*02:01	219.85	1.8
			HLA-A*02:06	83.07	0.81
			HLA-A*30:01	63.82	0.26
KLIKQELDK	77	85	HLA-A*03:01	161.2	0.68
KQLLPIVNK	201	209	HLA-A*03:01	94.25	0.38
			HLA-A*11:01	41.73	0.27
			HLA-A*30:01	262.22	0.59
			HLA-A*31:01	209.37	1.9
KQSCSISNI	209	217	HLA-A*02:06	39.09	0.45
KSDELLHNV	508	518	HLA-A*02:01	214.59	1.8
			HLA-A*02:06	23.94	0.29
			HLA-C*15:02	123.26	0.07
KTKCTASNK	419	427	HLA-A*03:01	69.7	0.3
			HLA-A*11:01	111.77	0.75
			HLA-A*30:01	2.66	0.02
			HLA-A*31:01	35.5	0.45
KVKLIKQEL	75	83	HLA-A*30:01	55.05	0.24
KVLDLKNYI	191	199	HLA-A*02:01	211.06	1.8
			HLA-A*02:06	22.15	0.26
			HLA-A*30:01	296.44	0.65
LAYVVQLPL	297	305	HLA-B*35:01	166.22	0.41
			HLA-B*58:01	28.18	0.14
			HLA-C*03:03	5.77	0.04
			HLA-C*14:02	260.89	0.36
LEITREFSV	231	239	HLA-B*40:02	91.55	0.19
LPLYGVIDT	303	311	HLA-B*35:01	270.45	0.56
LPSEVNLCN	375	383	HLA-B*35:01	190.67	0.45
			HLA-B*53:01	257.6	0.24
LRTGWYTSV	48	56	HLA-B*27:05	159.38	0.58
LSALRTGWY	43	53	HLA-A*01:01	40.52	0.12
			HLA-A*30:02	85.85	0.31
LSKKRKRRF	129	137	HLA-B*08:01	126.51	0.3
LTLPSEVNL	373	381	HLA-A*02:06	262.19	1.8
LTNSELLSL	252	260	HLA-A*02:06	232.8	1.7
			HLA-C*15:02	186.88	0.1
LTSKVLDLK	188	196	HLA-A*11:01	22.63	0.13
			HLA-A*68:01	19.85	0.15
MITTIIIVI	526	534	HLA-A*02:01	237.08	1.9
			HLA-A*68:02	39.5	0.34
MSIIKEEVL	289	297	HLA-B*58:01	299.05	0.76
NIDIFNPKY	383	391	HLA-A*01:01	81.41	0.2
NQSLAFIRK	500	508	HLA-A*11:01	161.14	1

NSLTLPSEV	371	379	HLA-C*15:02	271.21	0.16
QLLMQSTPA	94	102	HLA-A*02:06	118.75	1.2
QQSYSIMSI	283	291	HLA-A*02:06	47.28	0.52
			HLA-B*15:01	220.31	0.91
			HLA-B*39:01	277.72	0.22
QSYSIMSII	284	292	HLA-A*68:02	117.41	0.67
			HLA-B*58:01	176.48	0.57
			HLA-C*12:03	149.32	0.23
			HLA-C*15:02	176.53	0.1
REFSVNAGV	235	243	HLA-B*40:01	23.04	0.08
			HLA-B*40:02	53.53	0.1
			HLA-B*44:03	212.12	0.38
RLLEITREF	229	237	HLA-A*32:01	6.49	0.02
			HLA-B*15:01	247.12	0.99
RTGWYTSVI	49	57	HLA-A*32:01	26.88	0.05
			HLA-B*58:01	194.36	0.6
			HLA-C*15:02	132.94	0.07
SAIASGIAV	146	154	HLA-A*02:06	56.64	0.6
			HLA-A*68:02	57.1	0.43
			HLA-C*03:03	32.74	0.15
			HLA-C*12:03	254.94	0.33
SLGAIVSCY	409	417	HLA-A*29:02	155.82	0.7
			HLA-A*30:02	164.35	0.61
			HLA-B*15:01	118.94	0.6
SLSNGVSVL	180	188	HLA-A*02:01	113.4	1.2
			HLA-B*15:01	231.11	0.94
SNKNRGIIK	425	433	HLA-A*30:01	95	0.32
SQNITEEFY	25	33	HLA-A*30:02	139.93	0.53
			HLA-B*15:01	102.77	0.54
SSQNITEEF	24	32	HLA-A*32:01	203.23	0.24
STNKAVVSL	173	181	HLA-A*32:01	75.89	0.11
			HLA-A*68:02	97.03	0.61
SVGNTLYYV	451	459	HLA-A*02:01	29.86	0.33
			HLA-A*02:06	10.51	0.13
			HLA-A*68:02	27.15	0.25
SVITSLGAI	405	413	HLA-A*02:06	37.83	0.43
			HLA-A*68:02	76.53	0.53
SYSIMSIIK	285	293	HLA-A*11:01	246.55	1.5
			HLA-A*30:01	237.55	0.58
TELQLLMQS	91	99	HLA-B*18:01	155.34	0.35
TLSKKRKRR	128	136	HLA-A*31:01	278.9	2.2
TRTDRGWYC	335	343	HLA-C*07:01	222.69	0.07
TTIIIVIIV	528	536	HLA-A*68:02	13.42	0.14
TTPVSTYML	244	252	HLA-A*68:02	21.39	0.21
			HLA-C*14:02	233.1	0.34
TYMLTNSEL	249	257	HLA-A*23:01	87.38	0.28

			HLA-A*24:02	177.57	0.3
			HLA-C*14:02	4.61	0.02
VFCDTMNSL	365	373	HLA-C*14:02	202.55	0.31
VNAGVTTPV	239	247	HLA-A*68:02	40.2	0.34
VQIVRQQSY	278	286	HLA-A*30:02	132.93	0.49
			HLA-B*15:01	34.13	0.21
VQLPLYGVI	301	309	HLA-A*02:06	66.48	0.67
VSLSNGVSV	179	187	HLA-C*12:03	285.38	0.35
VSSSVITSL	402	410	HLA-B*58:01	50.88	0.24
VSVGNTLYY	450	458	HLA-A*01:01	63.5	0.17
			HLA-A*11:01	88.82	0.62
			HLA-A*29:02	10.26	0.08
			HLA-A*30:02	54.19	0.15
			HLA-B*35:01	190.21	0.45
			HLA-B*58:01	19.43	0.11
			HLA-C*12:03	164.92	0.25
VTLSKKRKR	127	135	HLA-A*31:01	38.83	0.48
VTTPVSTYM	243	251	HLA-A*68:02	133	0.74
			HLA-B*58:01	19.13	0.11
			HLA-C*03:03	46.96	0.2
			HLA-C*15:02	37.21	0.02
VVQLPLYGV	300	308	HLA-A*02:06	103.62	0.96
			HLA-A*68:02	204.32	0.94
WYCDNAGSV	341	349	HLA-C*12:03	155.48	0.24
			HLA-C*14:02	51.22	0.09
YCDNAGSVS	342	350	HLA-C*05:01	54.19	0.07
YIDKQLLPI	198	206	HLA-A*02:01	20.94	0.23
			HLA-A*02:06	10.09	0.12
			HLA-C*05:01	136.31	0.12
			HLA-C*08:02	240.89	0.02
YMLTNSELL	250	258	HLA-A*02:01	9.14	0.08
			HLA-A*02:06	22.36	0.27
			HLA-B*39:01	95.21	0.12
			HLA-C*03:03	65.56	0.24
			HLA-C*14:02	81.09	0.15
YTSVITIEL	53	61	HLA-A*02:01	134.33	1.3
			HLA-A*02:06	146.93	1.3
			HLA-A*68:02	6.06	0.05
			HLA-C*03:03	22.41	0.13
			HLA-C*14:02	122.81	0.21
			HLA-C*15:02	292.2	0.17

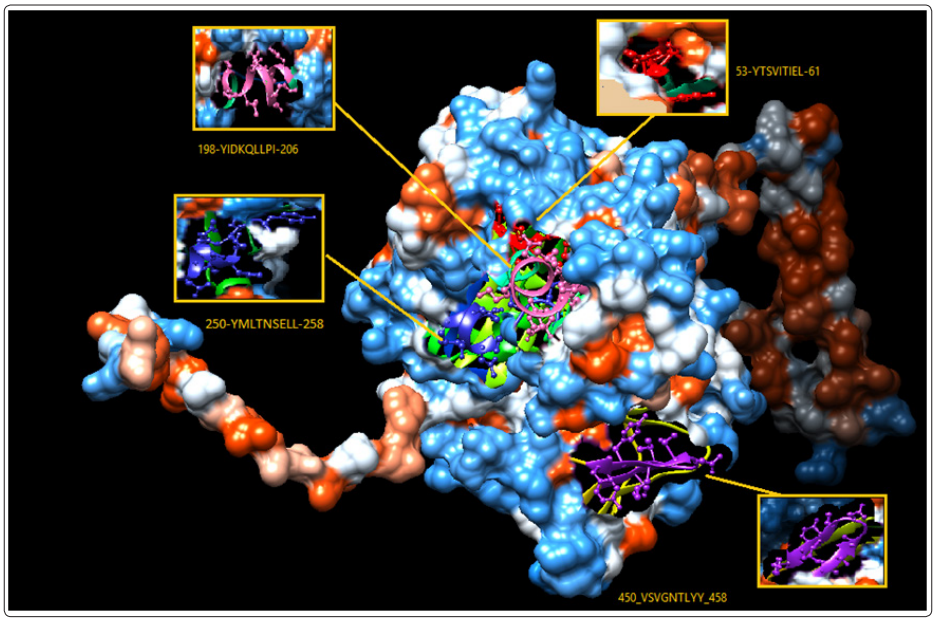

Figure 4: T cell proposed epitopes that interact with MHC-I The epitopes 53-YTSVITIEL-61,198-YIDKQLLPI-206, 250-YMLTNSELL-258 and 450-VSVGNTLYY-458 were found to interact with MHC-1 and only one epitope (250-YMLTNSELL-258) was found interacting with both MHC-1 and MHC-II alleles. The positions of proposed epitopes were shown in the 3D structural level of fusion F protein of RSV

MHC-II binding predictions

RSV Fusion F protein was analyzed using IEDB MHC- II binding prediction tool to predict T lymphocytes epitopes that have binding affinity with MHC-II alleles based on NN-align with half-maximal inhibitory concentration (IC50) ≤ 1000. Total of 397 core epitopes were found to interact with MHC-II alleles. Table 4 demonstrated the best four core epitopes that interacted with MHC-II alleles. The core epitopes 217-IETVIEFQQ-226, 250-YMLTNSELL-258, 477-FYDPLVFPS-485 and 505-FIRKSDELL-513 interacted with higher number of MHC-II alleles. The three dimensional structural level (3D) of these epitopes within Fusion F protein of the RSV was shown in figure 5.

Table 4: List of top four epitopes that had binding affinity with MHC-II alleles. The position of peptides is according to position of amino acids in fusion F protein of RSV virus

Core Sequence	Peptide Sequence	Start	End	Allele	IC50	Rank
IETVIEFQQ	SCSISNIETVIEFQQ	211	225	HLA-DPA1*02:01/DPB1*01:01	151.6	14.51
				HLA-DQA1*01:01/DQB1*05:01	2338.8	25.38
				HLA-DQA1*01:02/DQB1*06:02	292.2	18.94
				HLA-DQA1*05:01/DQB1*02:01	372.1	8.41
				HLA-DRB4*01:01	703.6	33.59
	CSISNIETVIEFQQK	212	226	HLA-DPA1*02:01/DPB1*01:01	100.2	10.42
				HLA-DPA1*02:01/DPB1*05:01	747.9	14.34
				HLA-DPA1*03:01/DPB1*04:02	867.2	30.05
				HLA-DQA1*01:01/DQB1*05:01	2133.7	24.01
				HLA-DQA1*01:02/DQB1*06:02	270.8	17.88
				HLA-DQA1*03:01/DQB1*03:02	99	0.86
				HLA-DQA1*05:01/DQB1*02:01	498.9	11.15
				HLA-DRB1*08:02	179.7	3.82
				HLA-DPA1*02:01/DPB1*01:01	80.6	8.59
				HLA-DPA1*02:01/DPB1*05:01	650.1	12.86
				HLA-DPA1*03:01/DPB1*04:02	865.7	30.03
				HLA-DQA1*01:02/DQB1*06:02	282.3	18.46
				HLA-DQA1*03:01/DQB1*03:02	107.9	1.01

				HLA-DQA1*05:01/DQB1*02:01	661.5	14.46
				HLA-DRB1*08:02	190.1	4.1
				HLA-DRB1*11:01	2530.4	55.93
				HLA-DPA1*02:01/DPB1*01:01	75.6	8.11
				HLA-DPA1*02:01/DPB1*05:01	505.5	10.47
				HLA-DPA1*03:01/DPB1*04:02	714.8	27.69
				HLA-DQA1*01:01/DQB1*05:01	2687.7	27.59
				HLA-DQA1*01:02/DQB1*06:02	285.1	18.6
				HLA-DQA1*03:01/DQB1*03:02	123.9	1.29
				HLA-DQA1*05:01/DQB1*02:01	970.9	19.97
				HLA-DRB1*08:02	253.8	5.84
				HLA-DRB1*11:01	1962.2	51.56
				HLA-DPA1*02:01/DPB1*01:01	95	9.95
				HLA-DPA1*02:01/DPB1*05:01	597.3	12.02
				HLA-DPA1*03:01/DPB1*04:02	927.5	30.9
				HLA-DQA1*01:02/DQB1*06:02	474	26.42
				HLA-DQA1*03:01/DQB1*03:02	167.4	2.1
				HLA-DQA1*05:01/DQB1*02:01	1335	25.48
				HLA-DRB1*08:02	403.1	9.76
				HLA-DRB1*11:01	2333.2	54.54
	NIETVIEFQQKNNRL	216	230	HLA-DPA1*02:01/DPB1*01:01	101	10.49
				HLA-DPA1*02:01/DPB1*05:01	873.9	16.1
				HLA-DPA1*03:01/DPB1*04:02	1111.6	33.26
				HLA-DQA1*01:02/DQB1*06:02	698.4	33.28
				HLA-DQA1*03:01/DQB1*03:02	338.7	5.49
				HLA-DQA1*05:01/DQB1*02:01	1828.5	31.75
				HLA-DRB1*08:02	598.3	14.44
	IETVIEFQQKNNRLL	217	231	HLA-DPA1*01:03/DPB1*02:01	1000.7	30.48
				HLA-DPA1*02:01/DPB1*01:01	92.3	9.7
				HLA-DPA1*03:01/DPB1*04:02	1223.9	34.57
				HLA-DQA1*01:02/DQB1*06:02	1255.7	44.82
				HLA-DQA1*03:01/DQB1*03:02	798.3	14.25
				HLA-DQA1*05:01/DQB1*02:01	2552.3	39.17
				HLA-DRB1*08:02	1314.9	27.43
YMLTNSELL	TTPVSTYMLTNSELL	244	258	HLA-DPA1*01/DPB1*04:01	572.2	16.07
				HLA-DPA1*01:03/DPB1*02:01	458.7	20.53
				HLA-DPA1*02:01/DPB1*01:01	142.4	13.85
				HLA-DPA1*02:01/DPB1*05:01	1567.3	24.19
				HLA-DPA1*03:01/DPB1*04:02	152	12.59
				HLA-DQA1*05:01/DQB1*02:01	389.4	8.8
				HLA-DRB1*01:01	7.8	3.36
				HLA-DRB1*04:01	68.9	5.53
				HLA-DRB1*07:01	14.7	2.73
				HLA-DRB1*09:01	206.9	13.32
				HLA-DRB1*13:02	827.9	23.32
				HLA-DRB1*15:01	571.5	31.09

				HLA-DRB3*01:01	288.6	8.51
				HLA-DRB5*01:01	31.3	7.1
				HLA-DPA1*01/DPB1*04:01	284.2	10.63
				HLA-DPA1*01:03/DPB1*02:01	188.4	12.41
				HLA-DPA1*02:01/DPB1*05:01	1285.6	21.23
				HLA-DPA1*03:01/DPB1*04:02	86.6	8.78
				HLA-DQA1*05:01/DQB1*02:01	402.7	9.09
				HLA-DRB1*01:01	6.3	1.99
				HLA-DRB1*04:01	49	3.71
				HLA-DRB1*07:01	20.5	3.95
				HLA-DRB1*09:01	162	10.79
				HLA-DRB1*11:01	2259.2	53.99
				HLA-DRB1*13:02	564	18.69
				HLA-DRB1*15:01	515.3	29.51
				HLA-DRB3*01:01	269.2	8.18
				HLA-DRB5*01:01	19.9	4.86
	PVSTYMLTNSELLSL	246	260	HLA-DPA1*01/DPB1*04:01	82.4	4.59
				HLA-DPA1*02:01/DPB1*05:01	787.5	14.92
				HLA-DQA1*05:01/DQB1*02:01	406.1	9.16
				HLA-DRB1*01:01	5	0.79
				HLA-DRB1*04:01	40.7	2.92
				HLA-DRB1*07:01	22	4.25
				HLA-DRB1*09:01	131	8.93
				HLA-DRB1*11:01	1244.1	44.08
				HLA-DRB1*13:02	328	13.48
				HLA-DRB1*15:01	404	25.98
				HLA-DRB3*01:01	221.8	7.3
				HLA-DRB5*01:01	11.6	2.82
	VSTYMLTNSELLSLI	247	261	HLA-DPA1*01/DPB1*04:01	54	3.29
				HLA-DPA1*02:01/DPB1*01:01	20.7	1.67
				HLA-DPA1*02:01/DPB1*05:01	571	11.59
				HLA-DQA1*05:01/DQB1*02:01	477.2	10.7
				HLA-DRB1*01:01	4.6	0.47
				HLA-DRB1*04:01	38.6	2.72
				HLA-DRB1*07:01	29.2	5.56
				HLA-DRB1*09:01	148.9	10.04
				HLA-DRB1*11:01	771	36.97
				HLA-DRB1*13:02	217.6	10.37
				HLA-DRB1*15:01	385	25.31
				HLA-DRB3*01:01	193.4	6.72
				HLA-DRB5*01:01	8.7	1.98
	STYMLTNSELLSLIN	248	262	HLA-DPA1*01/DPB1*04:01	54.8	3.33
				HLA-DPA1*02:01/DPB1*05:01	547.8	11.19
				HLA-DQA1*05:01/DQB1*02:01	555	12.31
				HLA-DRB1*01:01	5.3	1.06
				HLA-DRB1*04:01	56.7	4.43

				HLA-DRB1*07:01	44.2	7.84
				HLA-DRB1*09:01	219.1	13.93
				HLA-DRB1*13:02	276.4	12.09
				HLA-DRB1*15:01	470.9	28.19
				HLA-DRB3*01:01	481.8	11.5
				HLA-DRB5*01:01	12.6	3.08
	TYMLTNSELLSLIND	249	263	HLA-DPA1*01/DPB1*04:01	60.2	3.58
				HLA-DPA1*02:01/DPB1*05:01	517.8	10.68
				HLA-DQA1*05:01/DQB1*02:01	659	14.41
				HLA-DRB1*01:01	6.7	2.37
				HLA-DRB1*04:01	105.5	8.55
				HLA-DRB1*04:04	133.1	15.1
				HLA-DRB1*04:05	150.8	13.44
				HLA-DRB1*07:01	68.3	10.96
				HLA-DRB1*09:01	352.7	20.05
				HLA-DRB1*13:02	426.5	15.83
				HLA-DRB1*15:01	713.6	34.58
				HLA-DRB3*01:01	1198.5	19.92
				HLA-DRB5*01:01	20.8	5.06
	YMLTNSELLSLINDM	250	264	HLA-DPA1*01/DPB1*04:01	91.1	4.95
				HLA-DPA1*02:01/DPB1*01:01	35.4	3.53
				HLA-DPA1*02:01/DPB1*05:01	631.5	12.57
				HLA-DRB1*01:01	9.4	4.72
				HLA-DRB1*04:01	185.7	14.11
				HLA-DRB1*04:05	189.4	15.94
				HLA-DRB1*07:01	114.6	15.45
				HLA-DRB1*09:01	875.3	36.09
				HLA-DRB1*13:02	592.7	19.24
				HLA-DRB1*15:01	1411.2	46.11
				HLA-DRB3*01:01	2111.1	28.45
				HLA-DRB4*01:01	477.1	26.61
				HLA-DRB5*01:01	38.4	8.3
FYDPLVFPS	GEPIINFYDPLVFPS	471	485	HLA-DPA1*01:03/DPB1*02:01	482.1	21.08
				HLA-DPA1*02:01/DPB1*05:01	758.3	14.5
				HLA-DQA1*03:01/DQB1*03:02	1521.8	25.67
				HLA-DQA1*05:01/DQB1*03:01	1895.1	60.69
				HLA-DRB5*01:01	658.7	36.92
	EPIINFYDPLVFPSD	472	486	HLA-DPA1*01:03/DPB1*02:01	437.6	20.02
				HLA-DPA1*02:01/DPB1*05:01	653.8	12.92
				HLA-DQA1*03:01/DQB1*03:02	867.2	15.45
				HLA-DQA1*05:01/DQB1*03:01	1906.8	60.82
				HLA-DRB5*01:01	632.8	36.34
	PIINFYDPLVFPSDE	473	487	HLA-DPA1*01:03/DPB1*02:01	454.1	20.42
				HLA-DPA1*02:01/DPB1*05:01	550	11.23
				HLA-DQA1*05:01/DQB1*03:01	1876.9	60.49
				HLA-DRB1*01:01	424.9	54.09

				HLA-DRB1*04:01	387.9	24.42
				HLA-DRB1*11:01	1263.2	44.32
				HLA-DRB5*01:01	551.9	34.4
	IINFYDPLVFPSDEF	474	488	HLA-DPA1*01/DPB1*04:01	403.7	13.14
				HLA-DPA1*01:03/DPB1*02:01	430.2	19.85
				HLA-DPA1*02:01/DPB1*01:01	123.6	12.4
				HLA-DPA1*02:01/DPB1*05:01	485.1	10.11
				HLA-DPA1*03:01/DPB1*04:02	382	20.75
				HLA-DQA1*05:01/DQB1*03:01	1881.8	60.55
				HLA-DRB1*01:01	285.4	47.25
				HLA-DRB1*04:01	361.5	23.27
				HLA-DRB1*11:01	873.4	38.74
				HLA-DRB5*01:01	399.5	30.09
	INFYDPLVFPSDEFD	475	489	HLA-DPA1*01:03/DPB1*02:01	501.4	21.52
				HLA-DPA1*02:01/DPB1*01:01	290.4	22.6
				HLA-DPA1*02:01/DPB1*05:01	883.5	16.24
				HLA-DPA1*03:01/DPB1*04:02	720	27.78
				HLA-DQA1*05:01/DQB1*03:01	2385.7	65.52
				HLA-DRB1*01:01	357.1	51.16
				HLA-DRB1*04:01	625.9	33.05
				HLA-DRB1*11:01	1314.6	44.97
				HLA-DRB5*01:01	639.9	36.5
	NFYDPLVFPSDEFDA	476	490	HLA-DPA1*01:03/DPB1*02:01	320.9	16.91
				HLA-DPA1*02:01/DPB1*01:01	635.1	34.68
				HLA-DPA1*03:01/DPB1*04:02	1549.1	37.9
				HLA-DRB1*01:01	529.5	57.65
				HLA-DRB1*04:01	973.8	42.36
				HLA-DRB1*09:01	2627	62.7
				HLA-DRB1*11:01	2031.5	52.16
				HLA-DRB5*01:01	1156.4	45.54
	FYDPLVFPSDEFDAS	477	491	HLA-DPA1*02:01/DPB1*01:01	830.5	39.27
				HLA-DPA1*03:01/DPB1*04:02	2669.5	46.26
				HLA-DRB1*01:01	626.5	60.28
				HLA-DRB1*04:01	1107.2	45.33
				HLA-DRB5*01:01	2002.3	54.92
FIRKSDELL	INQSLAFIRKSDELL	499	513	HLA-DPA1*01/DPB1*04:01	2543.3	35.2
				HLA-DPA1*02:01/DPB1*01:01	163.5	15.35
				HLA-DPA1*03:01/DPB1*04:02	1195.1	34.24
				HLA-DRB1*01:01	102.5	30.53
				HLA-DRB1*03:01	455	13.3
				HLA-DRB1*13:02	1510.6	32.51
	NQSLAFIRKSDELLH	500	514	HLA-DPA1*01/DPB1*04:01	1913.6	30.54
				HLA-DPA1*02:01/DPB1*01:01	151.8	14.53
				HLA-DPA1*03:01/DPB1*04:02	589.8	25.44
				HLA-DQA1*05:01/DQB1*02:01	345.8	7.81
				HLA-DRB1*01:01	55.7	22.57

				HLA-DRB1*07:01	178.6	20.13
				HLA-DRB1*09:01	926.7	37.27
				HLA-DRB1*13:02	1048.3	26.6
				HLA-DRB1*15:01	1650.8	48.9
				HLA-DRB5*01:01	454.7	31.78
	QSLAFIRKSDELLHN	501	515	HLA-DPA1*01/DPB1*04:01	1654.8	28.33
				HLA-DPA1*02:01/DPB1*01:01	128.8	12.81
				HLA-DPA1*03:01/DPB1*04:02	308.1	18.63
				HLA-DQA1*05:01/DQB1*02:01	473.2	10.62
				HLA-DRB1*01:01	38.4	18.17
				HLA-DRB1*07:01	310.1	26.68
				HLA-DRB1*08:02	579.4	14.05
				HLA-DRB1*09:01	806.6	34.41
				HLA-DRB1*13:02	811.5	23.05
				HLA-DRB1*15:01	1645.3	48.84
				HLA-DRB3*01:01	2934.3	35.02
				HLA-DRB5*01:01	328.6	27.68
	SLAFIRKSDELLHNV	502	516	HLA-DPA1*01/DPB1*04:01	1157.8	23.57
				HLA-DPA1*01:03/DPB1*02:01	945	29.63
				HLA-DPA1*02:01/DPB1*01:01	84.9	9.01
				HLA-DPA1*03:01/DPB1*04:02	146.8	12.34
				HLA-DQA1*05:01/DQB1*02:01	592.1	13.08
				HLA-DRB1*01:01	23.7	13
				HLA-DRB1*04:01	329.1	21.82
				HLA-DRB1*07:01	364.2	28.81
				HLA-DRB1*08:02	761	17.91
				HLA-DRB1*09:01	660.7	30.49
				HLA-DRB1*13:02	570.3	18.82
				HLA-DRB1*15:01	1729.9	49.73
				HLA-DRB3*01:01	2750.5	33.61
				HLA-DRB5*01:01	206.2	22.31
	LAFIRKSDELLHNVN	503	517	HLA-DPA1*01/DPB1*04:01	1147.7	23.46
				HLA-DPA1*01:03/DPB1*02:01	1558.3	37.62
				HLA-DPA1*02:01/DPB1*01:01	89	9.4
				HLA-DQA1*05:01/DQB1*02:01	768.8	16.46
				HLA-DRB1*01:01	41.6	19.09
				HLA-DRB1*07:01	541.1	34.44
				HLA-DRB1*08:02	1022.1	22.77
				HLA-DRB1*13:02	656.5	20.42
				HLA-DRB1*15:01	2226.4	54.37
				HLA-DRB5*01:01	262.8	25.06
	AFIRKSDELLHNVNA	504	518	HLA-DPA1*01/DPB1*04:01	1321.3	25.24
				HLA-DPA1*01:03/DPB1*02:01	2608.2	47.25
				HLA-DPA1*02:01/DPB1*01:01	144.2	13.98
				HLA-DQA1*05:01/DQB1*02:01	1102.3	22.05
				HLA-DRB1*01:01	100.8	30.29

				HLA-DRB1*04:04	907.1	45.45
				HLA-DRB1*04:05	947.5	40.36
				HLA-DRB1*07:01	906.1	42.65
				HLA-DRB1*08:02	1412.7	28.86
				HLA-DRB1*13:02	904.1	24.5
				HLA-DRB5*01:01	455.7	31.81
	FIRKSDELLHNVNAG	505	519	HLA-DPA1*01/DPB1*04:01	1664.8	28.42
				HLA-DPA1*02:01/DPB1*01:01	203.7	17.95
				HLA-DPA1*03:01/DPB1*04:02	177.6	13.83
				HLA-DQA1*05:01/DQB1*02:01	1482	27.47
				HLA-DRB1*01:01	199.1	40.81
				HLA-DRB1*04:04	758.8	42.14
				HLA-DRB1*04:05	1656.6	51.42
				HLA-DRB1*07:01	1386.3	49.88
				HLA-DRB1*08:02	1558.5	30.86
				HLA-DRB1*13:02	1119.5	27.59
				HLA-DRB5*01:01	832.5	40.39

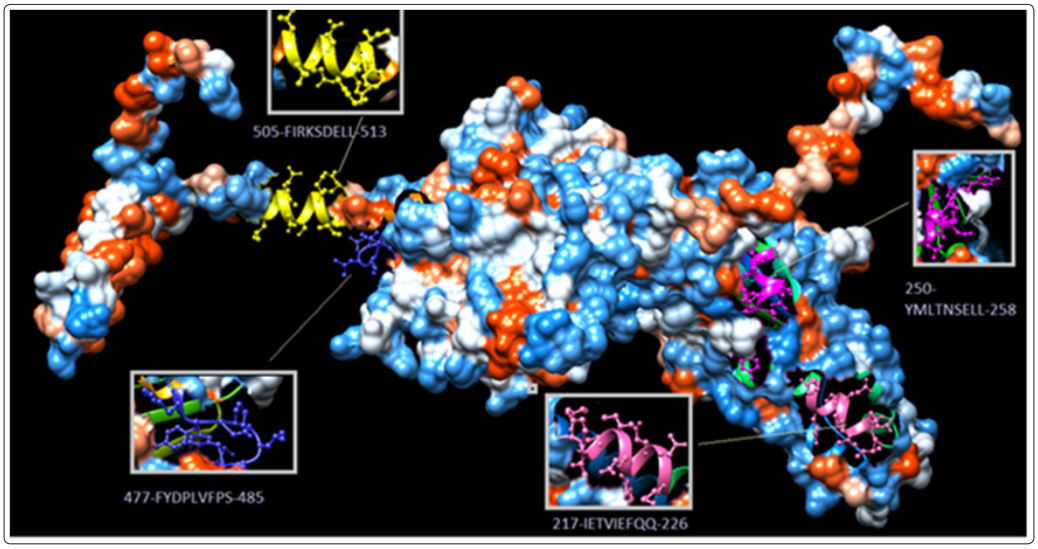

Figure 5: T cell proposed epitopes that interact with MHC-II. The epitope 250-YMLTNSELL-258 interacted with MHC-I and MHC-II alleles while the epitopes 217-IETVIEFQQ-226, 477-FYDPLVFPS-485 and 505-FIRKSDELL-513 interacted only with MHC-II alleles. The positions of proposed epitopes were shown in the 3D structural level of fusion F protein of RSV

Overlapping of T cell epitopes residues in MHC class I and II

In this study four epitopes (53-YTSVITIEL-61, 198-YIDKQLLPI-206, 250-YMLTNSELL-258 and 450-VSVGNTLYY-458) were proposed as T cytotoxic epitopes and four epitopes (The core epitopes 217-IETVIEFQQ-226, 250-YMLTNSELL-258, 477-FYDPLVFPS-485 and 505-FIRKSDELL-513) were proposed as T helper epitopes since they interacted with MHC-I and MHC-II alleles. Each epitope and the number of the interacted alleles were shown in table [5]. Only one epitope (250-YMLTNSELL-258) was shown to be interacting with both MHC-I and MHC-II alleles. This epitope showed the higher number of MHC-II interacted alleles.

Table 5: The predicted epitopes and the number of the MHC-1 and/or MHC-II interacted alleles

Epitope	MHC-1 interacted alleles	Epitope	MHC-11 interacted alleles
450-VSVGNTLYY-458	7	250-YMLTNSELL-258	90
53-YTSVITIEL-61	6	217-IETVIEFQQ-226	52
250-YMLTNSELL-258	5	477-FYDPLVFPS-485	49
198-YTSVITIEL-61	4	505-FIRKSDELL-513	75

Population Coverage

The suggested epitopes that demonstrated higher affinity to interact with MHC-I and MHC-II alleles and that bound to different sets of alleles were selected for population coverage analysis. Table 6 demonstrated the population coverage percentages for each epitope and their epitopes sets. The epitopes 450-VSVGNTLYY-458, 250-YMLTNSELL-258, 198-YIDKQLLPI-206 and 53-YTSVITIEL-61 interacted with most frequent MHC-I alleles and they demonstrated high percentage against the whole world population coverage with epitope set 82.85%. Strikingly the first four epitopes that interacted with MHC-1 in addition to the epitope 217-FYDPLVFPS-266 were interacted with most frequent MHC-II alleles and they demonstrated high percentage against the whole world population coverage with epitopes set of 99.98%. The overall MHC-1 and MHC-II whole world population coverage for all suggested epitopes was 100%.

Table 6: The population coverage against the whole world for the predicted epitopes. The first four epitopes were found to be interacted with both MHC-1 and MHC-11 with good population coverage, while the epitopes 53-YTSVITIEL-61 and 217-IETVIEFQQ-226 demonstrated higher population coverage only in MHC-1 and MHC-11 respectively. The overall population coverage epitope set for all predicted epitopes in MHC-1 and MHC-11 was 100%

MHC-1		MHC-11		MHC-1/ MHC-11	
Epitope	**Population coverage**	**Epitope**	**Population coverage**	**Epitope**	**Population coverage**
450-VSVGNTLYY-458	49.76%	250-YMLTNSELL-258	99.64%	450-VSVGNTLYY-458	49.76%
53-YTSVITIEL-61	51.26%	217-IETVIEFQQ-226	99.69%	53-YTSVITIEL-61	51.26%
250-YMLTNSELL-258	48.61%	505-FIRKSDELL-513	99.56%	250-YMLTNSELL-258	48.61%
198-YIDKQLLPI-206	47.68%	477-FYDPLVFPS-485	99.63%	198-YIDKQLLPI-206	47.68%
Epitope set	82.85%	Epitope set	99.98%	505-FIRKSDELL-513	99.56%
				217-IETVIEFQQ-226	99.69%
				477-FYDPLVFPS-485	99.63%
				250-YMLTNSELL-258	99.64%
				Epitope set	100.00%

The population coverage against the whole world for the predicted epitopes. Only one epitope (250-YMLTNSELL-258) interacted with both MHC-1 and MHC-11. The overall population coverage epitope set for all predicted epitopes in MHC-1 and MHC-11 was 100%

Discussion

Infectious diseases have largely been eliminated through vaccine by inducing protective neutralizing antibodies against vulnerable epitopes. Many considerable pathogens have to thwart traditional vaccine development, despite those vulnerable epitopes targeted by neutralizing antibodies have been identified for numerous such cases. For that reason, new vaccine design methods to promote epitope-specific neutralizing antibodies are needed. The Vaccine usually contains two classes of epitopes the combination of two classes of vaccine epitopes: B-cell epitopes and T-cell epitopes can make the vaccine either induce specific humoral or cellular immunity against specific pathogens [29]. Accordingly, in the present study the B and T cells were analyzed systemically against 191 retrieved strains of the Fusion F protein of RSV virus to obtain epitope vaccine candidates.

For better determination of best vaccine candidate epitopes, the whole Fusion F protein was analyzed including the receptor-binding domain (RBD). Noticeably the amino acids sequence from 1 to 574 of the fusion F glycoprotein represents a unique domain during RSV infection. Fusion requires a conformational change from the pre-fusion to the post-fusion form of F20. Targeting the transition from the pre-fusion F to the post-fusion F form can prevent the fusion process and block RSV infection in vitro and in vivo [30-32]. The F protein contains several antigenic sites (Ø, I-V) [33, 34]. Two antigenic sites, II (target of Palivizumab) and IV (target of 101F and MAb19), are present in both the pre- and post-fusion F proteins with the defined binding epitope (residues 262–276), (422_ 471) respectively thus antigenic sites were used in the major target of neutralizing antibodies, making it an attractive target for vaccine development in multiple studies [35-39].

In this study, nine epitopes were successfully predicted to interact with B and T cells. In our results two epitopes demonstrated high affinity to interact with B cell and were found within the region from 1 to 574 of fusion F glycoprotein. These epitopes were 441-YVSNK-445, and 440-DYVS-396. Most importantly the latter epitope (441-YVSNK-445) was positioned within the RBD region and antigenic site IV, Furthermore the two epitopes demonstrated conservancy in Bepipred as linear epitopes and got high scores (over the thresholds) in Emini surface accessibility and Kolaskar and Tongaonkar antigenicity method. Therefore, they were recommended as promising peptides vaccine candidates against B cells. The epitopes illustrated in Table 2, were the only conserved epitopes from all retrieved strains of fusion F glycoprotein that were available in NCBI database until 2013 and have high probability of activating humoral immune response.

The immune response of T cell is considered as a long lasting response compared to B cell, where the antigen can easily escape the antibody memory response [40]. Thus, designing vaccine against T cell is much more important. In this study, 95 and 1576 conserved T cell epitopes were predicted to interact with both MHC-1 and MHC-II alleles respectively. Among them, seven epitopes 450-VSVGNTLYY-458, 198-YIDKQLLPI-206,

53-YTSVITIEL-61, 505-FIRKSDELL-513, 477-FYDPLVFPS-485, 217-IETVIEFQQ-226 and 250-YMLTNSELL-258 demonstrated good binding affinity against MHC-1 or MHC-II alleles.

The epitopes 198-YIDKQLLPI-206 and 450-VSVGNTLYY-458 were found positioned in the region from 1 to 574 that considered as a unique domain during RSV infection. This epitope successfully interacted with both MHC-1 and MHC-II alleles. Moreover, it is position found with antigenic site IV; it interacted with the high number of alleles (Table 5) and provided the high population coverage epitope set (Table 6). Also, it was found tandemly positioned with 441-YVSNK-445 (epitope that predicted for B-cell). This result may reflect the importance of this region in the fusion F protein during RSV infection since it contains these tandem epitopes. Also, the epitope 250-YMLTNSELL-258 demonstrated unique features. Firstly, it was found to be located within the unique domain (residues 1 to 510) of fusion F protein [41]. Secondly it provided high binding affinity with MHC-II and provided excellent results concerning binding affinity with MHC-I favorable epitope set (Table 6).with good population coverage against whole world.

Moreover, the epitope 53-YTSVITIEL-61 showed a high binding affinity for both MHC-1 and MHC-II alleles with high population coverage against the whole world. While the epitope 505-FIRKSDELL-513, 217-IETVIEFQQ-226, 477-FYDPLVFPS-485 provided excellent results concerning binding affinity to MHC-II alleles. It is noteworthy that this epitope interacted with 75, 52 and 49 alleles in MHC-II respectively (Table 6) with population coverage 99.56%, 99.69 %, 99.63%, respectively. However, they did not interact with MHC-1 alleles. Therefore, they were only chosen as MHC-II epitope. Accordingly, the six predicted epitopes demonstrated very promising features against T lymphocytes as a vaccine candidate against RSV. Moreover, the population coverage against the whole world for the epitopes that interacted with MHC-I and MHC-II alleles provided coverage of 79.54% and 99.98% respectively. The overall epitope set for the MHC-I and MHC-II predicted epitopes showed excellent population coverage against whole world population (100%). Accordingly, these epitopes were strongly recommended as promising epitopes vaccine candidate against T cell.

Conclusion

Respiratory Syncytial Virus (RSV) is the major cause of the lower respiratory tract illness. Developing an effective and safe vaccine is strongly needed to prevent the infection of RSV. In this study, only two epitopes (441-YVSNK-445 and 440-DYVS-443) successfully passed all B cell prediction tools and were proposed as B cells epitopes. For T cells, four epitopes (53-YTSVITIEL-61; 250-YMLTNSELL-258, 198-YIDKQLLPI-206, and 450-VSVGNTLYY-458) were proposed as a vaccine candidates since they interacted with the highest number of alleles and the best binding affinity to MHC-1 alleles. Moreover, four core proposed epitopes interacted with higher number of MHC-II alleles (217-IETVIEFQQ-226; 250-YMLTNSELL-258; 477-FYDPLVFPS-485 and 505-FIRKSDELL-513. The population coverage was 48.61% and 99.64% for MHC-I and MHC-II epitopes, respectively, and 100% for all T cells epitopes. In this study, ten epitopes successfully proposed as vaccine candidate against RSV. In vivo and in vitro clinical trials studies are required to elucidate the effectiveness of these epitopes as vaccine.

Acknowledgments

Authors would like to thank the staff members of Department of Molecular Biology and Bioinformatics, College of Veterinary Medicine, University of Bahri/ Sudan for their cooperation and support.

Competing Interest

The authors declare that they have no competing interests.

Funding

No funding was received.

References

1. 1. Lamb RA, Parks GD (2007) Paramyxoviridae: the viruses and their replication. In Knipe DMH, Peter M, editors. (ed), Fields virology, 5th ed, Lippincott Williams & Wilkins, Philadelphia, PA 2: 1449-1496.
2. 2. Collins PL, McIntosh K, Chanock RM (1996) Fields Virology, eds. Fields, B. N., Knipe, D. M. & Howley, P. M. (Lippincott, Philadelphia) 1996: 1313-1351.
3. 3. Morris JA, RE Blount Jr, RE Savage (1956) Recovery of cytopathogenic agent from chimpanzees with coryza. Proc Soc Exp Biol Med 92: 544-550.
4. 4. Chanock R, Roizman B, Myers R (1957) Recovery from infants with respiratory illness of virus related to chimpanzee coryza agent (CCA): I Isolation properties and characterization. Am J Hyg 66: 281-290.
5. 5. Chanock RM, Finberg L (1957) Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA): II. Epidemiologic aspects of infection in infants and young children. Am J Hyg 66: 291-300.
6. 6. Olli Ruuskanen MD (2014) Lower Respiratory Tract Virus Findings in Mechanically Ventilated Patients with Severe Community-Acquired Pneumonia. Clin Infect Dis 59: 62-70.
7. 7. Ducoire G, Cauchi P, Hendrickx E (2001) Respiratory syncytial virus epidemiology in Belgium in 1998, 1999 and 2000. Abstract presented at the 5th World Congress of Perinatal Medicine 2001: 23-27.
8. 8. Carbonell-Estrany X, Quero J, IRIS Study Group (2001) RSV hospitalization rates in premature infants born over two consecutive seasons. Pediatr Infect Dis J 20: 874-879.
9. 9. Cherian T, Simoes EA, Steinhoi MC, K Chitra, M John, et al. (1990) Bronchiolitis in tropical south India. Am J Dis Child 144: 1026-1030.
10. 10. Simoes EA (1999) Respiratory syncytial virus infection. Lancet 354: 847-852.
11. 11. Shay DK, Holman RC, Newman RD, Liu LL, Stout JW, et al. (1999) Bronchiolitis-associated hospitalizations among US children, 1980–1996. JAMA 282: 1440-1446.
12. 12. Hall CB, Weinberg GA, Iwane MK, Aaron K Blumkin, Kathryn M Edwards, et al. (2009) The burden of respiratory syncytial virus infection in young children. N Engl J Med 360: 588-598.
13. 13. Shay DK, Holman RC, Newman RD, Liu LL, Stout JW, et al. (1999) Bronchiolitis-associated hospitalizations among US children, 1980–1996. JAMA 282: 1440-1446
14. 14. Hall CB (1998) Respiratory syncytial virus. In: Feigin RD, Cherry JD. (1998) (eds) Textbook of Pediatric Infectious Diseases. Philadelphia, PA: WB Saunders 1998: 2084-2111.
15. 15. Huang YT, Collins PL, Wertz GW (1985) Characterization

of the 10 proteins of human respiratory syncytial virus: identification of a fourth envelope-associated protein. Virus Res 2: 157-173.
16. 16. Ruuskanen O, Ogra PL (1993) Respiratory syncytial virus. Curr Probl Pediatr 23: 50-79.
17. 17. Grosfeld H, Hill M, Collins PL (1995) RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA. J Virol 69: 5677-5686.
18. 18. Yu Q, Hardy RW, Wertz GW (1995) Functional cDNA clones of the human respiratory syncytial (RS) virus N, P, and L proteins support replication of RS virus genomic RNA analogs and define the minimal trans-acting requirements for RNA replication. J Virol 69: 2412-2419.
19. 19. Meanwell NA, Krystal M (2000) Respiratory syncytial virus: recent progress towards the discovery of effective prophylactic and therapeutic agents. Drug Discov Today 5: 241-252.
20. 20. Graham BS, Johnson TR, Peebles RS (2000) Immune-mediated disease pathogenesis in respiratory syncytial virus infection. Immunopharmacology 48: 237-247.
21. 21. Gan SW, Tan E, Lin X, Yu D, Wang J, et al. (2012) The small hydrophobic protein of the human respiratory syncytial virus forms pentameric ion channels. J Biol Chem 287: 24671-24689.
22. 22. Domachowske JB, Rosenberg HF (1999) Respiratory syncytial virus infection: immune response, immunopathogenesis, and treatment. Clin Microbiol Rev 12: 298-309.
23. 23. Pastey MK, Crowe Jr JE, Graham BS (1999) RhoA interacts with the fusion glycoprotein of respiratory syncytial virus and facilitates virus-induced syncytium formation. J Virol 73: 7262-7270.
24. 24. Kim HW, Canchola JG, Brandt CD, G Pyles, RM Chanock, et al. (1969) Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. Am J Epidemiol 89: 422-434.
25. 25. Empey KM, Peebles RS Jr, Kolls JK (2010) Pharmacologic advances in the treatment and prevention of respiratory syncytial virus. Clin Infect Dis 50: 1258-1267.
26. 26. Anderson LJ, Dormitzer PR, Nokes DJ, Rappuoli R, Roca A, et al. (2013) Strategic priorities for respiratory syncytial virus (RSV) vaccine development. Vaccine 31: B209-B215.
27. 27. Marshall E, M Enserink (2004) Caution urged on SARS vaccines. Science 303: 944.
28. 28. Black M, Trent A, Tirrell M, Olive C (2010) Advances in the design and delivery of peptide subunit vaccines with a focus on toll-like receptor agonists. Expert Rev Vaccines 9: 157-173.
29. 29. Purcell AW, McCluskey J, Rossjohn J (2007) More than one reason to rethink the use of peptides in vaccine 9 design. Nat Rev Drug Discov 6: 404-414.
30. 30. Zhu Q, McLellan JS, Kallewaard NL, Ulbrandt ND, Palaszynski S, et al. (2017) A highly potent extended half-life antibody as a potential RSV vaccine surrogate for all infants. Sci Transl Med 9.
31. 31. Samuel D, Xing W, Anita Niedziela-Majka, Jinny S Wong, Magdeleine Hung, et al. (2015) GS-5806 inhibits pre- to post fusion conformational changes of the respiratory syncytial virus fusion protein. Antimicrob Agents Chemother 59: 7109-7112.
32. 32. Mejias A, Garcia-Maurino C, Rodriguez-Fernandez R, Peeples ME, Ramilo O (2017) Development and clinical applications of novel antibodies for prevention and treatment of respiratory syncytial virus infection. Vaccine 35: 496-502.
33. 33. Villafana T, Falloon J, Griffin MP, Zhu Q, Esser MT (2017) Passive and active immunization against respiratory syncytial virus for the young and old. Expert Rev Vaccines 16: 1-13.
34. 34. Griffin MP, Khan AA, Esser MT, Jensen K, Takas T, et al. (2017) Safety, Tolerability, and Pharmacokinetics of MEDI8897, the Respiratory Syncytial Virus Prefusion F-Targeting Monoclonal Antibody with an Extended Half-Life, in Healthy Adults. Antimicrob Agents Chemother 61.
35. 35. McLellan JS, Yang Y, Graham BS, Kwong PD (2011) Structure of respiratory syncytial virus fusion glycoprotein in the post fusion conformation reveals preservation of neutralizing epitopes. J Virol 85: 7788-7796.
36. 36. Hause AM, Henke DM, Avadhanula V, Shaw CA, Tapia LI, et al. (2017) Sequence variability of the respiratory syncytial virus (RSV) fusion gene among contemporary and historical genotypes of RSV/A and RSV/B. Plos One 12: e0175792.
37. 37. Huang K, Incognito L, Cheng X, Ulbrandt ND, Wu H (2010) Respiratory syncytial virus-neutralizing monoclonal antibodies motavizumab and palivizumab inhibit fusion. J Virol 84: 8132-8140.
38. 38. Raghunandan R, Lu H, Zhou B, Xabier MG, Massare MJ, et al. (2014) An insect cell derived respiratory syncytial virus (RSV) F nanoparticle vaccine induces antigenic site II antibodies and protects against RSV challenge in cotton rats by active and passive immunization. Vaccine 32: 6485-6492.
39. 39. Palomo C, Mas V, Vazquez M, Cano O, Luque D, et al. (2014) Polyclonal and monoclonal antibodies specific for the six-helix bundle of the human respiratory syncytial virus fusion glycoprotein as probes of the protein post-fusion conformation. Virology 460: 461-119.
40. 40. Black M, Trent A, Tirrell M, Olive C (2010) Advances in the design and delivery of peptide subunit vaccines with a focus on toll-like receptor agonists. Expert Rev Vaccines 9: 157-173.
41. 41. Pang H, Y Liu, X Han, Y Xu, F Jiang, et al. (2004) Protective humoral responses to severe acute respiratory syndrome-associated coronavirus: implications for the design of an effective protein-based vaccine. J Gen Virol 85: 3109-3131.